

Геометрические методы в математической физике

Катанаев Михаил Орионович¹

Математический институт имени В. А. Стеклова
Российской Академии Наук

и

Казанский федеральный университет, г. Казань.

22 ноября 2016 г.

¹Любые замечания, указания на ошибки, неточности и опечатки прошу отправлять на e-mail: katanaev@mi.ras.ru

Оглавление

Предисловие	xii
Основные обозначения и соглашения	xiv
1 Введение	1
1.1 Множества	1
1.2 Поле вещественных чисел $\mathbb{R}$ и прямая	4
1.3 Евклидово пространство $\mathbb{R}^n$	12
1.3.1 $\mathbb{R}^n$ как метрическое пространство	12
1.3.2 $\mathbb{R}^n$ как топологическое пространство	18
1.3.3 $\mathbb{R}^n$ как векторное пространство	30
1.3.4 $\mathbb{R}^n$ как аффинное пространство	66
1.4 Отображения	71
1.4.1 Отображения топологических пространств	75
1.5 Преобразования координат	80
1.6 Группа двумерных вращений $\mathbb{O}(2)$	84
1.7 Группа двумерных преобразований Лоренца $\mathbb{O}(1, 1)$	89
1.8 Трехмерное евклидово пространство	97
1.9 Пространство Минковского	106
1.9.1 Группа Пуанкаре	109
1.9.2 Группа Галилея	115
1.9.3 Группа конформных преобразований	116
1.9.4 Трехмерное пространство Минковского	119
1.9.5 Четырехмерное пространство Минковского	122
1.10 Специальная теория относительности	127
1.10.1 Нерелятивистские модели	128
1.10.2 Релятивистские модели	131
1.10.3 Замедление времени и лоренцево сокращение	135
1.10.4 Сложение скоростей и эффект Доплера	139
1.10.5 Равноускоренное движение	142
2 Многообразия и тензорные поля	147
2.1 Многообразия	147
2.2 Разбиение единицы	156
2.3 Многообразия с краем	158
2.4 Расслоения	160
2.5 Скалярные поля и плотности	163
2.6 Векторные поля и 1-формы	166
2.6.1 Локальное определение	166

2.6.2	Глобальное определение векторных и ковекторных полей	169
2.6.3	Кокасательные векторные поля и ростки	173
2.6.4	Векторные поля и дифференцирования	174
2.6.5	Векторные поля и интегральные кривые	177
2.6.6	1-формы и гиперповерхности	189
2.6.7	Алгебра Ли векторных полей	190
2.7	Тензорные поля	192
2.8	Полностью антисимметричные тензоры	198
2.9	Отображения многообразий	200
2.10	Подмногообразия	204
2.11	Теорема Фробениуса	209
2.12	Слоения	215
2.13	Бесконечно малые преобразования координат	220
2.14	Производная Ли	225
3	Дифференциальные формы и интегрирование	230
3.1	Внешняя алгебра	230
3.2	Дифференциальные формы	237
3.3	Внешнее дифференцирование	240
3.4	Теорема Дарбу	245
3.5	Оператор Лапласа–Бельтрами	251
3.6	Разложение Ходжа	256
3.7	Интегрирование дифференциальных форм	258
3.7.1	Форма объема	261
3.7.2	Формула Стокса	264
4	Метрика	268
4.1	Определение и свойства	268
4.2	Метрика на лоренцевых многообразиях	272
4.3	Векторные поля и вложения	278
4.4	Выбор системы координат	281
5	Связность на векторном расслоении и расслоении реперов	284
5.1	Векторные расслоения	284
5.2	Связность на векторном расслоении	289
5.3	Аффинная связность	295
5.4	Связность на расслоении реперов	297
5.5	Критерий локальной тривиальности	305
6	Аффинная геометрия. Локальное рассмотрение	308
6.1	Локальное определение аффинной связности	308
6.2	Кручение и неметричность	311
6.3	Ковариантная производная тензорных плотностей	315
6.4	Параллельный перенос	317
6.5	Геометрический смысл кручения	322
6.6	Свойства аффинной связности	324
6.7	Локальное определение тензора кривизны	327
6.8	Свойства тензора кривизны	329
6.9	Неголономный базис	334

6.10	Тождества Бианки	339
7	Криволинейные координаты в $\mathbb{R}^3$	340
7.1	Сферические координаты	340
7.2	Цилиндрические координаты	345
8	Группы Ли	348
8.1	Группы Ли и локальные группы Ли	348
8.2	Действие группы слева	354
8.3	Действие группы справа	362
8.4	Присоединенное представление	363
8.5	Группы Ли как (псевдо)римановы пространства	366
8.6	Группы Ли как пространства Римана–Картана	370
8.7	Группа аффинных преобразований прямой	371
8.8	Гомоморфизмы групп Ли	375
8.9	Экспоненциальное отображение для групп Ли	379
8.10	Интегрирование на группах Ли	381
8.11	Некоторые общие свойства групп Ли	386
8.12	Полупрямое произведение групп	389
8.13	Алгебры Ли	391
8.13.1	Операции над алгебрами Ли	394
8.13.2	Простые и полупростые алгебры Ли	399
8.13.3	Квадратичные формы	402
8.14	Группа Ли $\mathrm{GL}(n, \mathbb{C})$	403
8.14.1	Алгебра Ли $\mathfrak{gl}(n, \mathbb{C})$	404
8.14.2	Группа Ли $\mathrm{GL}(n, \mathbb{C})$	411
8.15	Универсальная накрывающая $\widetilde{\mathrm{SL}}(2, \mathbb{R})$	416
8.16	Классификация простых алгебр и групп Ли	419
9	Группы преобразований	424
9.1	Действие групп преобразований	424
9.2	Инфинитезимальные преобразования	431
9.3	Инвариантные структуры	434
9.4	Отображения групп преобразований	437
10	Гомотопии и фундаментальная группа	439
10.1	Пути	439
10.2	Гомотопия непрерывных отображений	444
10.3	Фундаментальная группа	447
10.4	Фундаментальная группа и ориентируемость	453
11	Накрытия	459
11.1	Определения и примеры	459
11.2	Фундаментальная группа пространства орбит	463
11.3	Группа скольжений и существование накрытий	467
12	Главные и ассоциированные расслоения	469
12.1	Главные расслоения	469
12.2	Ассоциированные расслоения	478
12.3	Отображение расслоений	486

13 Связности на главных и ассоциированных расслоениях	491
13.1 Связность на главном расслоении	491
13.2 Форма кривизны и структурное уравнение	502
13.2.1 Связность на $\mathbb{P}(\mathbb{R}^2, \pi, \mathbb{R}) = \mathbb{R}^3$	511
13.3 Параллельный перенос	514
13.4 Группы голономии	515
13.5 Петля Вильсона	518
13.6 Отображение связностей	522
13.7 Связность на ассоциированном расслоении	524
13.8 Свойства групп голономий	527
13.9 Плоские связности	529
13.10 Локальные и инфинитезимальные группы голономии	532
13.11 Инвариантные связности	536
14 Приложения в квантовой механике	542
14.1 Адиабатическая теорема	543
14.1.1 Двухуровневая система	551
14.2 Фаза Берри	554
14.2.1 Абелев случай: невырожденное состояние	554
14.2.2 Частица со спином $1/2$ в магнитном поле	556
14.2.3 Неабелев случай: вырожденное состояние	559
14.3 Эффект Ааронова–Бома	562
14.3.1 Электрический потенциал	563
14.3.2 Магнитный потенциал	566
15 Векторные поля Киллинга	569
15.1 Изометрии и инфинитезимальные изометрии	569
15.2 Свойства векторных полей Киллинга	574
15.3 Однородные и изотропные многообразия	577
15.4 Симметричные тензоры на пространстве постоянной кривизны	583
15.5 Пространства с максимально симметричными подпространствами	585
15.6 Лоренц-инвариантные метрики	589
15.7 Автомодельные метрики	593
16 Геодезические и экстремали	596
16.1 Геодезические	596
16.2 Экстремали	603
16.3 Интегрирование уравнений для экстремалей и геодезических	608
16.4 Вторая вариация уравнений для экстремалей	609
16.5 Уравнение Гамильтона–Якоби для экстремалей	613
16.6 Волновое уравнение	616
16.7 Приближение эйконала	619
16.8 Гармонические координаты	622
16.9 Нормальные, геодезические или римановы координаты	623
16.9.1 Нормальные координаты. Локальное рассмотрение.	624
16.9.2 Нормальные координаты в (псевдо)римановом пространстве	630
16.9.3 (Псевдо)римановы пространства постоянной кривизны	632
16.9.4 Нормальные координаты и экспоненциальное отображение	636

16.10	Полнота римановых многообразий	638
16.11	Формулы Френе	641
17	Разложения кривизны и кручения	645
17.1	Разложение тензора кривизны	645
17.2	Разложение тензора кручения	651
17.3	Конформные преобразования	653
17.4	Преобразование Вейля в (псевдо)римановой геометрии	654
17.5	Вейлевская эквивалентность (псевдо)римановых пространств	656
17.6	Преобразование Вейля в аффинной геометрии	659
17.7	Алгебраически независимые компоненты тензора кривизны	661
17.8	Классификация Петрова	662
17.9	Преобразование Эйнштейна–Кауфман	673
18	Симметрические пространства	675
18.1	Определения и общие свойства	675
18.2	(Псевдо)римановы симметрические пространства	680
18.3	Псевдосферы	685
19	Симплектические и пуассоновы многообразия	690
19.1	Симплектические группы	690
19.2	Симплектические многообразия	696
19.3	Пуассоновы многообразия	700
19.4	Структура Ли–Пуассона	706
19.5	Отображения пуассоновых многообразий	707
20	Принцип наименьшего действия	713
20.1	Постановка вариационных задач	713
20.1.1	Задача с заданными граничными условиями	714
20.1.2	Задача со свободными граничными условиями	716
20.1.3	Задача с подвижной границей	717
20.1.4	Задача на условную стационарную точку	720
20.1.5	Другие задачи и терминология	721
20.2	Первая теорема Нетер	723
20.2.1	Тензор энергии-импульса	725
20.2.2	Тензор момента количества движения	727
20.3	Вторая теорема Нетер	729
20.4	Эффективное действие	732
20.5	Редуцированное действие	733
21	Канонический формализм в механике точечных частиц	736
21.1	Преобразование Лежандра	736
21.2	Гамильтонова динамика точечных частиц	738
21.3	Потенциальное движение точечной частицы	744
21.4	Комплексный язык в гамильтоновом формализме	747
21.5	Лемма Стокса	748
21.6	Канонические уравнения Гамильтона	749
21.7	Принцип Мопертюи	752
21.8	Уравнение Гамильтона–Якоби	755
21.9	Принцип Гюйгенса	758

21.10	Переменные действие-угол	760
21.11	Канонические преобразования	764
21.12	Производящие функции канонических преобразований	771
21.13	Разделение переменных в уравнении Гамильтона–Якоби	779
22	Гамильтонова динамика частиц со связями	783
22.1	Связи в гамильтоновом формализме	783
22.2	Гамильтонова динамика частиц со связями II рода	785
22.3	Гамильтонова динамика частиц со связями I рода	793
22.4	Смешанный набор связей	800
22.5	Переход от калибровочно инвариантного лагранжиана к гамильтониану	802
22.6	Гамильтонова динамика в теории поля	808
22.7	Калибровочная модель нерелятивистской частицы	812
22.8	Частица в псевдоримановом пространстве	815
22.8.1	Временн́ая калибровка	822
22.8.2	Калибровка светового конуса	825
22.9	Граничные слагаемые в калибровочных моделях	827
22.9.1	Лагранжева формулировка	827
22.9.2	Гамильтонова формулировка	830
22.10	О каноническом квантовании	833
23	Основы общей теории относительности	836
23.1	Пространство-время, метрика и гравитация	836
23.2	О постановке задач в теории гравитации	840
23.3	Действие Гильберта–Эйнштейна	842
23.4	Вариация действия Гильберта–Эйнштейна	848
23.5	Зависимость уравнений Эйнштейна	850
23.6	Действие для полей материи в обобщенных моделях гравитации	851
23.7	Скалярно-тензорные модели	854
23.8	Полиномиальная форма действия Гильберта–Эйнштейна	856
23.9	Точечные частицы в теории гравитации	862
23.9.1	Нерелятивистский предел для точечной частицы	868
23.9.2	Теория гравитации Ньютона	871
23.9.3	Свойства тензора энергии-импульса точечных частиц	874
23.10	Ньютонов предел	878
23.11	Гравитационные волны	881
23.12	Сплошная среда в общей теории относительности	885
23.12.1	Акустические фононы в нерелятивистской гидродинамике	891
23.13	Выбор системы координат	896
23.13.1	Сопутствующая система координат	897
23.13.2	Временн́ая калибровка	902
23.13.3	Калибровка светового конуса	912
24	Гамильтонова формулировка общей теории относительности	920
24.1	Лагранжиан Гильберта–Эйнштейна	920
24.2	АДМ параметризация метрики и репера	921
24.3	Геометрия гиперповерхностей	927
24.4	Кривизна в АДМ параметризации метрики	933

24.5	Гамильтониан	937
24.6	Вторичные связи	939
24.7	Полиномиальная гамильтонова форма	942
24.8	Проблема энергии в теории гравитации	948
24.8.1	Тензор энергии-импульса полей материи	949
24.8.2	Псевдотензор энергии-импульса для гравитации	950
24.8.3	Законы сохранения и векторы Киллинга	951
24.8.4	Полная гравитационная энергия асимптотически плоского пространства-времени	952
25	Скалярные и калибровочные поля	956
25.1	Действительное скалярное поле	956
25.1.1	Скалярное поле в пространстве Минковского	956
25.1.2	Скалярное поле в аффинной геометрии	960
25.1.3	Скалярное поле в общей теории относительности	962
25.2	Комплексное скалярное поле	965
25.2.1	Комплексное скалярное поле в пространстве Минковского	965
25.2.2	Комплексное скалярное поле в аффинной геометрии	969
25.2.3	Комплексное скалярное поле в общей теории относительности	970
25.3	Электромагнитное поле	971
25.3.1	Лагранжева формулировка	976
25.3.2	Законы сохранения	981
25.3.3	Гамильтонова формулировка	982
25.3.4	Скалярная электродинамика	990
25.3.5	Электромагнитное поле в аффинной геометрии	994
25.3.6	Электромагнитное поле в общей теории относительности	997
25.4	Поле Прок'а	1002
25.4.1	Гамильтонова формулировка	1003
25.5	Поля Янга–Миллса	1008
25.5.1	Лагранжева формулировка	1014
25.5.2	Гамильтонова формулировка	1016
25.5.3	Поля Янга–Миллса в аффинной геометрии	1020
25.5.4	Поле Янга–Миллса в общей теории относительности	1022
26	Алгебры Клиффорда	1024
26.1	Определение алгебр Клиффорда	1025
26.2	Алгебры Клиффорда низших размерностей	1037
26.3	Классификация алгебр Клиффорда	1044
26.4	Представления алгебр Клиффорда	1053
26.5	Свойства алгебр Клиффорда	1056
26.6	Группы (псевдо)вращений	1059
26.7	Накрытия групп вращений	1062
26.8	Накрытия в низших размерностях	1074
26.9	Алгебра Клиффорда $\mathfrak{cl}(p, q)$	1083
27	Спиноры	1087
27.1	Спинорное представление группы Лоренца $\mathbb{O}(1, 3)$	1088
27.1.1	γ -матрицы Дирака	1088
27.1.2	Алгебра матриц Дирака	1095

27.1.3	Трансформационные свойства спиноров	1097
27.1.4	Дискретные преобразования P , C и T	1103
27.1.5	Уравнение Дирака	1107
27.1.6	Спиноры в геометрии Римана–Картана	1115
28	Геометрия поверхностей	1123
28.1	Геометрия Римана–Картана на поверхности	1123
28.2	Неголономный базис в двух измерениях	1126
28.3	Выбор системы координат	1127
28.3.1	Диагональная калибровка	1128
28.3.2	Конформная калибровка для римановых поверхностей	1128
28.3.3	Конформная калибровка для лоренцевых поверхностей	1130
28.4	Координаты светового конуса	1132
29	Поверхности постоянной кривизны	1135
29.1	Сфера S^2	1135
29.2	Двуполостный гиперболоид H^2	1139
29.3	Однополостный гиперболоид L^2	1144
29.4	Уравнение Лиувилля	1151
30	Лоренцевы поверхности с одним вектором Киллинга	1157
30.1	Локальный вид лоренцевой метрики	1159
30.2	Конформные блоки	1163
30.3	Экстремали	1165
30.3.1	Форма экстремалей	1165
30.3.2	Асимптотика экстремалей	1170
30.3.3	Полнота экстремалей	1171
30.4	Построение глобальных решений	1173
30.5	Примеры	1176
30.5.1	Решение Шварцшильда	1176
30.5.2	Решение Рейснера–Нордстрема	1178
30.5.3	Экстремальная черная дыра	1179
30.5.4	Плоскость Минковского	1180
30.5.5	Поверхности постоянной кривизны	1182
30.6	Координаты Эддингтона–Финкельштейна	1185
30.7	Дифференцируемость метрики в седловой точке	1189
31	Римановы поверхности с одним вектором Киллинга	1191
31.1	Локальный вид римановой метрики	1191
31.2	Экстремали	1193
31.3	Построение глобальных решений	1197
31.4	Решение Шварцшильда	1202
32	Двумерная гравитация	1204
32.1	Двумерная гравитация с кручением	1204
32.2	Действие и уравнения движения	1207
32.3	Каноническая формулировка	1210
32.3.1	Каноническое преобразование	1214
32.4	Дилатонизация	1215
32.5	Размерности полей	1217

32.6	Общее решение без полей материи	1218
32.6.1	Локальное решение	1218
32.6.2	Глобальные решения	1224
32.7	Квадратичная модель	1226
32.8	Пуассоновы сигма модели	1228
32.8.1	Гамильтонова формулировка	1231
32.8.2	Общее решение уравнений движения	1232
32.8.3	Двумерные поля Янга–Миллса	1234
32.8.4	Двумерная гравитация как пуассонова сигма модель	1235
33	Сплетенные решения в общей теории относительности	1238
33.1	Сплетенное произведение	1238
33.2	Двумерная редукция	1239
33.3	Произведение поверхностей постоянной кривизны	1243
33.4	Пространственно симметричные решения	1245
33.4.1	Сферически симметричные решения $K = 1$	1250
33.4.2	Планарные решения $K = 0$	1257
33.4.3	Гиперболические глобальные решения $K = -1$	1260
33.5	Лоренц-инвариантные решения	1262
33.5.1	Лоренц-инвариантные решения $K = 1$	1265
33.5.2	Решения с плоскостью Минковского $K = 0$	1270
33.6	Итоги главы	1271
34	Решение Шварцшильда	1273
34.1	Координаты Шварцшильда	1273
34.2	Координаты Эддингтона–Финкельштейна	1281
34.3	Координаты Крускала–Секереша	1282
34.4	Мост Эйнштейна–Розена	1286
34.5	Изотропные координаты	1289
34.6	Координаты Пенлеве–Гулстранда	1293
34.7	Гравитационное поле точечной частицы	1294
34.8	Проход по мосту Эйнштейна–Розена	1304
34.9	Вложения решения Шварцшильда	1310
34.10	Эвристический “вывод” решения Шварцшильда	1322
34.10.1	Радиус Шварцшильда	1324
35	Космология	1327
35.1	Сплошная среда в космологии	1327
35.1.1	Сплошная среда в классической механике	1328
35.1.2	Сплошная среда в релятивистской механике	1330
35.1.3	Уравнение Райчаудхури	1333
35.1.4	Большой взрыв и большое сжатие	1336
35.2	Трехмерные пространства постоянной кривизны	1337
35.2.1	Трехмерная сфера S^3	1337
35.2.2	Двуполостный гиперболоид H^3	1344
35.3	Вселенная Фридмана	1348
35.3.1	Метрика однородной и изотропной вселенной	1349
35.3.2	Пример	1355
35.3.3	Космологическое красное смещение	1356

35.3.4	Уравнения движения	1359
35.3.5	Вселенная Эйнштейна	1364
35.3.6	Линейное уравнение состояния	1365
35.4	Вакуумные решения	1369
35.4.1	Пространство-время Минковского	1369
35.4.2	Пространство-время де Ситтера dS	1370
35.4.3	Пространство-время анти-де Ситтера AdS	1376
35.4.4	Вложения вселенной Фридмана	1380
35.4.5	Экстремали во вселенной Фридмана	1383
35.4.6	Космологические горизонты	1387
35.5	Однородные вселенные	1388
35.5.1	Классификация Бианки	1389
35.5.2	Уравнения движения для однородной вселенной	1395
35.5.3	Вселенная Казнера	1401

36 Геометрическая теория дефектов 1406

36.1	Упругие деформации	1407
36.2	Дислокации	1410
36.3	Дисклинации	1414
36.4	Свободная энергия	1418
36.5	Фиксирование калибровки	1421
36.6	Асимметричная теория упругости	1425
36.7	Дислокации в теории упругости	1427
36.7.1	Винтовая дислокация	1427
36.7.2	Клиновое дислокация	1429
36.7.3	Краевая дислокация	1432
36.8	Линейные дислокации в геометрической теории дефектов	1433
36.8.1	Параллельные клиновые дислокации	1434
36.8.2	Клиновое дислокация в геометрическом подходе	1437
36.8.3	Конформные отображения	1441
36.8.4	Непрерывное распределение линейных дислокаций	1449
36.8.5	Разрезание и склеивание	1452
36.9	Рассеяние фононов на клиновых дислокациях	1455
36.9.1	Рассеяние на одной клиновой дислокации	1457
36.9.2	Космические струны и множественные образы	1462
36.9.3	Общий подход к анализу параллельных клиновых дислокаций	1465
36.10	Цилиндрическая дислокация	1469
36.10.1	Цилиндрическая дислокация в теории упругости	1469
36.10.2	Цилиндрическая дислокация в геометрической теории дефектов	1474
36.10.3	Цилиндрическая дислокация в общей теории относительности	1479
36.10.4	Коническая цилиндрическая дислокация	1480
36.10.5	Асимптотически плоская клиновое дислокация	1484
36.10.6	Непрерывное распределение цилиндрических дислокаций	1487
36.10.7	Двустенная трубка	1488
36.11	Упругие колебания в среде с дислокациями	1493
36.11.1	Крутильные волны в волноводе с клиновой дислокацией	1496
36.11.2	Крутильные волны в двустенной трубке	1500
36.12	Примесь в поле клиновой дислокации	1503
36.13	О динамической теории	1505

37 Дополнение	1506
37.1 Матрицы	1506
37.2 Матрицы Паули	1514
37.3 Кватернионы	1516
37.4 Полностью антисимметричные тензоры	1519
37.5 δ -функция	1522
37.6 Функции Грина	1525
37.7 Ряды и интегралы Фурье	1527
37.8 Интегралы	1529
Библиография	1531

Предисловие

Вашему вниманию предлагается расширенный вариант лекций “Геометрические методы в математической физике”, которые автор читал в течении 2008-2016 годов в научно-образовательном центре при МИАН им. В.А. Стеклова. Основная цель курса состоит в изложении некоторых аспектов современной дифференциальной геометрии и ее приложений в математической физике. Большая часть материала посвящена изложению тех разделов геометрии, которые уже нашли применение в математической физике. При этом изложение ведется параллельно как на бескоординатном языке, принятом в современных курсах, так и в координатах, что необходимо для приложений. Делается упор на рассмотрение аффинной связности общего вида с кручением и неметричностью, поскольку соответствующие модели довольно широко распространены в современной математической физике. В качестве приложений рассмотрены два примера из квантовой механики: фаза Берри и эффект Ааронова-Бома, а также общая теория относительности и геометрическая теория дефектов.

Дифференциальная геометрия изучает объекты, не зависящие от выбора системы координат, такие как векторные поля, метрика и связность. Компоненты этих полей зависят от выбранной системы отсчета, но сами объекты – нет. То же происходит в физике. Любое физическое явление не зависит от выбора системы координат. При этом могут меняться некоторые его характеристики, например, скорость, которые измеряются относительно выбранной системы координат. Но само явление инвариантно, например, притяжение массивных тел. По этой причине геометрия является основным математическим инструментом при построении моделей физики.

Занятия наукой чем-то похожи на гонки в разные стороны: стараешься, спешишь, и, хотя цель одна: познать Природу, никто заранее не знает правильной дороги. Поэтому иногда следует останавливаться и смотреть вокруг.

Настоящая монография не является учебником, она, скорее, во многих местах больше похожа на справочник с комментариями. Многие утверждения приводятся без доказательств, но со ссылками на книги и статьи, где они доказаны. От читателя требуется достаточная математическая подготовка. Предполагается, что он знаком с основами общей топологии, линейной алгебры, теории групп и некоторых других разделов математики.

Кроме того, при изложении делаются ссылки на определения и утверждения, которые объясняются только в дальнейшем тексте. Это сделано для подготовленного читателя, который хотя бы слышал об их существовании, чтобы подчеркнуть то место, где рассматриваемое понятие будет использовано в дальнейшем. Если читатель с этим не знаком, то ссылки на дальнейшее рассмотрение можно пропустить.

Некоторые ссылки в тексте отмечены знаками ???. Это означает, что идет ссылка на те формулы и разделы, которые находятся в стадии подготовки, но не включены в представленный текст.

Курс, который автор читал в НОЦ, был бы невозможен без поддержки и критиче-

ских замечаний сотрудников Отдела математической физики МИАН им. В. А. Стеклова. Автор выражает искреннюю благодарность В. С. Владимирову, И. В. Волвичу, А. К. Гущину, Ю. Н. Дрожжину, В. В. Жаринову, Б. И. Завьялову, В. П. Михайлову и А. Г. Сергееву за многочисленные обсуждения вопросов дифференциальной геометрии и математической физики.

Основные обозначения и соглашения

Дифференцирование выполняется раньше алгебраических операций.

$\mathbb{N} := \{1, 2, \dots\}$	– множество натуральных чисел,
$\mathbb{Z} := \{\dots - 1, 0, 1, \dots\}$	– группа целых чисел по сложению,
$\mathbb{Q}, \mathbb{R}, \mathbb{C}$	– поле рациональных, вещественных, комплексных чисел,
$\mathbb{R}_+$	– множество положительных вещественных чисел,
$\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$	– группа вещественных чисел по умножению,
$\dagger$	– комплексное или эрмитово сопряжение,
$\ltimes$	– полупрямое произведение групп,
$\dot{+}$	– полупрямая сумма алгебр,
$:=$	– равно по определению,
$\equiv$	– тождественно равно,
$\Leftrightarrow$	– эквивалентно,
x и $(x^1, \dots, x^n)$	– точка многообразия и ее координаты,
$\{x^\alpha\} = \{x^0, x^\mu\} = \{x^0, \mathbf{x}\}$	– декартовы координаты в пространстве Минковского, или координаты на псевдоримановом многообразии,
$\partial_\alpha := \frac{\partial}{\partial x^\alpha}$	– частная производная,
$\partial_{\alpha\beta}^2 := \frac{\partial^2}{\partial x^\alpha \partial x^\beta}$	– частная производная второго порядка,
∇_α	– ковариантная производная,
$\triangle$	– оператор Лапласа или Лапласа–Бельтрами,
$\square$	– оператор Даламбера, конец доказательства, примера или определения,
$(\partial f)^2 := g^{\alpha\beta} \partial_\alpha f \partial_\beta f$	– квадрат градиента функции f ,
$g_{\alpha\beta}$	– компоненты метрики,
$g := \det(g_{\alpha\beta})$	– определитель метрики,
e_α^a	– компоненты репера,
$\sqrt{ g } = \det(e_\alpha^a)$	– элемент объема (опредетель репера),
$v := dx^1 \wedge \dots \wedge dx^n \sqrt{ g }$	– форма объема на (псевдо)римановом многообразии $\mathbb{M}$,
e	– единица группы,
e	– основание натурального логарифма,
$\Gamma_{\alpha\beta}^\gamma$	– компоненты аффинной связности,
$\omega_{\alpha a}^b$	– компоненты линейной или лоренцевой связности,
$\text{sgn } \sigma(\alpha_1, \dots, \alpha_n)$	– знак перестановки σ индексов $\alpha_1, \dots, \alpha_n$.

Некоторые многообразия и классы объектов имеют специальные обозначения:

$\mathbb{R}^n$	– n -мерное евклидово пространство,
$\mathbb{R}_+^n$	– подпространство в $\mathbb{R}^n$, определяемое условием $x^n > 0$,
$\mathbb{R}^{1,n-1}$	– n -мерное пространство Минковского,
$\mathbb{S}_r^n$	– n -мерная сфера радиуса r ,
$\mathbb{B}_r^n$	– n -мерный шар радиуса r ,
$\mathcal{C}^k(\mathbb{M})$	– класс функций на многообразии $\mathbb{M}$, непрерывных вместе с производными вплоть до k -го порядка,
$\mathcal{X}(\mathbb{M})$	– множество гладких векторных полей на многообразии $\mathbb{M}$,
$\mathcal{T}(\mathbb{M})$	– множество гладких тензорных полей на многообразии $\mathbb{M}$,
$\mathbb{M} \approx \mathbb{N}$	– многообразие $\mathbb{M}$ диффеоморфно (гомеоморфно) многообразию $\mathbb{N}$,
$a \approx b$	– a примерно равно b ,
$a \Rightarrow b$	– b следует из a ,
$\mathbb{G} \simeq \mathbb{H}$	– группа (алгебра, векторное пространство, ...) $\mathbb{G}$ изоморфна группе (алгебре, векторному пространству, ...) $\mathbb{H}$,
$\mathbb{M} \rightarrow \mathbb{N}$	– отображение множеств,
$\mathbb{M} \ni a \mapsto b \in \mathbb{N}$	– отображение элементов множеств,
Антисимметризация по индексам обозначается квадратными скобками:	

$$A^{[ab]} := \frac{1}{2}(A^{ab} - A^{ba}),$$

$$A^{[abc]} := \frac{1}{6}(A^{abc} + A^{bca} + A^{cab} - A^{bac} - A^{acb} - A^{cba}).$$

В общем случае, когда имеется n индексов, сумма берется по всем $n!$ перестановкам и делится на $n!$. При этом четные перестановки индексов входят со знаком плюс, а нечетные – со знаком минус.

Симметризация индексов обозначается круглыми скобками:

$$A^{(ab)} := \frac{1}{2}(A^{ab} + A^{ba}),$$

$$A^{(abc)} := \frac{1}{6}(A^{abc} + A^{bca} + A^{cab} + A^{bac} + A^{acb} + A^{cba}).$$

Символ Кронекера δ_a^b является тождественным оператором, действующим в векторном пространстве, и также равен единичной матрице. Например, в $\mathbb{R}^4$

$$\delta_a^b := \text{diag}(++++) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (1)$$

Для краткости, произведение символов Кронекера обозначается одним символом

$$\delta_{ef\dots h}^{ab\dots d} := \delta_e^a \delta_f^b \dots \delta_h^d. \quad (2)$$

Иногда используется обобщенный символ Кронекера, помеченный шляпкой,

$$\hat{\delta}_{ef\dots h}^{ab\dots d} := n! \delta_e^{[a} \delta_f^b \dots \delta_h^{d]} = n! \delta_{[e}^a \delta_f^b \dots \delta_{h]}^d, \quad (3)$$

который получается из произведения (2) антисимметризацией по верхним или нижним индексам (что эквивалентно).

Евклидова метрика δ_{ab} имеет два нижних индекса и равна единичной матрице. Например, в четырехмерном евклидовом пространстве $\mathbb{R}^4$

$$\delta_{ab} := \text{diag}(++++) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4)$$

Метрика Минковского в четырехмерном пространстве-времени $\mathbb{R}^{1,3}$ имеет вид

$$\eta_{ab} := \text{diag}(+---) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (5)$$

Каноническая симплектическая форма в евклидовом пространстве $\mathbb{R}^{2n}$ имеет вид

$$\varpi := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad (6)$$

где $\mathbb{1}$ – единичная $n \times n$ матрица.

Готический шрифт

$\mathfrak{A}$ a	$\mathfrak{B}$ b	$\mathfrak{C}$ c	$\mathfrak{D}$ d	$\mathfrak{E}$ e	$\mathfrak{F}$ f	$\mathfrak{G}$ g	$\mathfrak{H}$ h	$\mathfrak{I}$ i
$\mathfrak{J}$ j	$\mathfrak{K}$ k	$\mathfrak{L}$ l	$\mathfrak{M}$ m	$\mathfrak{N}$ n	$\mathfrak{O}$ o	$\mathfrak{P}$ p	$\mathfrak{Q}$ q	$\mathfrak{R}$ r
$\mathfrak{S}$ s	$\mathfrak{T}$ t	$\mathfrak{U}$ u	$\mathfrak{V}$ v	$\mathfrak{W}$ w	$\mathfrak{X}$ x	$\mathfrak{Y}$ y	$\mathfrak{Z}$ z	

Используемый греческий шрифт

α	β	Γ γ	Δ δ	ϵ	ε	ζ	η
Θ θ	ϑ	ι	κ	Λ λ	μ	ν	Ξ ξ
o	Π π	ϖ	ρ	ϱ	Σ σ	ς	τ
Υ υ	Φ ϕ	φ	χ	Ψ ψ	Ω ω		

Глава 1

Введение

После напоминания основных понятий теории множеств и введения обозначений, достаточно подробно рассмотрено евклидово пространство с различных точек зрения. Это сделано по двум причинам. Во-первых, чтобы подчеркнуть, что в евклидовом пространстве можно задавать различные структуры, которые в дальнейшем будут обобщаться в аффинной геометрии. Во-вторых, чтобы при изложении аффинной геометрии не напоминать относительно сложные понятия геометрии Евклида. Далее рассматриваются евклидовы пространства и пространства Минковского низших размерностей, которые играют важную роль в приложениях. В заключительном разделе кратко изложены основы специальной теории относительности. При написании Введения, которое содержит хорошо известный материал, использованы, в основном, монографии [1–17].

1.1 Множества

В математике некоторые исходные понятия не имеют определения. Эти понятия основаны на интуиции и служат для определения других, более сложных, конструкций. Такими интуитивными понятиями являются *множество* и *элемент множества*. Под множеством понимают произвольную совокупность объектов, которые называются элементами множества. Говорят, что множество состоит из своих элементов. Один из способов задания множества состоит просто в перечислении его элементов.

Пример 1.1.1. Натуральные числа являются множеством, которое будем обозначать следующим образом:

$$\mathbb{N} := \{1, 2, \dots\}.$$

Множество натуральных чисел не определяется и лежит в основе всей математики. □

Пример 1.1.2. Множество целых чисел, включающее нуль и все отрицательные числа, будем обозначать

$$\mathbb{Z} := \{\dots - 2, -1, 0, 1, 2, \dots\}$$

Как правило, мы рассматриваем множество целых чисел $\mathbb{Z}$ как абелеву группу по отношению к сложению. □

Пример 1.1.3. Пусть m, n – произвольные целые числа, причем $n \neq 0$. Тогда множество чисел вида m/n называется множеством рациональных чисел и обозначается $\mathbb{Q}$. □

Пример 1.1.4. Вещественные и комплексные числа также являются множествами. Для них приняты обозначения $\mathbb{R}$ и $\mathbb{C}$. $\square$

Множества натуральных, целых и рациональных чисел являются бесконечными *счетными множествами*, т.е. их элементы можно пронумеровать натуральными числами. Множества вещественных $\mathbb{R}$ и комплексных чисел $\mathbb{C}$ представляют собой примеры *несчетных множеств*, поскольку их элементы невозможно пронумеровать натуральными числами. Множество называется *конечным*, если оно состоит из конечного числа элементов.

Пример 1.1.5. Множество поворотов евклидовой плоскости $\mathbb{R}^2$ вокруг начала координат на угол $\pi/2$ является конечным и состоит из четырех элементов. Это – циклическая группа четвертого порядка. $\square$

В большинстве случаев для обозначения множеств, таких как многообразия, группы, и др. мы будем употреблять латинские буквы, напечатанные ажурным шрифтом, $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$. Иногда будут использоваться также заглавные буквы, напечатанные курсивом $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$.

Синонимами понятий множества и элемента множества являются *пространство* и *точка пространства*. Разница в употреблении терминов множество и пространство сложилась исторически. О множестве векторов говорят, как о векторном пространстве. Множество, снабженное топологией, называют топологическим пространством. Часто, говоря о пространстве, подразумевают, что оно, в отличие от множества, снабжено какой-либо дополнительной структурой, будь то топология или структура векторного пространства. Множество элементов называется также *семейством*. Как правило, термин семейство употребляется тогда, когда его элементами являются некоторые множества.

Если элемент x принадлежит множеству $\mathcal{U}$, то мы пишем $x \in \mathcal{U}$. В противном случае применяется обозначение $x \notin \mathcal{U}$.

Определение. Если каждый элемент множества $\mathcal{U}$ принадлежит также множеству $\mathcal{V}$, то $\mathcal{U}$ является *подмножеством* $\mathcal{V}$. Это обозначается $\mathcal{U} \subset \mathcal{V}$ или $\mathcal{V} \supset \mathcal{U}$. При этом совпадение или равенство множеств $\mathcal{U} = \mathcal{V}$ значит, что $\mathcal{U} \subset \mathcal{V}$ и $\mathcal{V} \subset \mathcal{U}$. Это записывается в виде

$$\mathcal{U} = \mathcal{V} \quad \Leftrightarrow \quad \mathcal{U} \subset \mathcal{V} \quad \text{и} \quad \mathcal{V} \subset \mathcal{U},$$

где стрелка $\Leftrightarrow$ обозначает утверждение “тогда и только тогда”. Множество, не содержащее ни одного элемента, называется *пустым* и обозначается символом $\emptyset$. Оно, по определению, является подмножеством любого множества, $\emptyset \subset \mathcal{U}$, $\forall \mathcal{U}$. Подмножество $\mathcal{U} \subset \mathcal{V}$ называется *собственным*, если оно не пусто и не совпадает со всем $\mathcal{V}$: $\mathcal{U} \neq \emptyset$ и $\mathcal{U} \neq \mathcal{V}$. Мы также пишем $\mathcal{U} \subseteq \mathcal{V}$, если либо $\mathcal{U} \subset \mathcal{V}$, либо $\mathcal{U} = \mathcal{V}$. $\square$

Пример 1.1.6. Множество целых чисел $\mathbb{Z}$ имеет много подмножеств. Среди них есть подмножества чисел, которые делятся без остатка на натуральное число $n \in \mathbb{N}$, отличное от нуля. Это подмножество обозначается $n\mathbb{Z} \subset \mathbb{Z}$ и получается умножением каждого целого числа на n . $\square$

Подмножество элементов $x \in \mathcal{U}$, удовлетворяющих некоторому свойству P обозначается следующим образом

$$\{x \in \mathcal{U} : P\},$$

где свойство P , как правило, задается некоторым уравнением или неравенством для элементов x .

Пример 1.1.7. Множество положительных и отрицательных вещественных чисел определяется равенствами:

$$\mathbb{R}_+ := \{x \in \mathbb{R} : x > 0\},$$

$$\mathbb{R}_- := \{x \in \mathbb{R} : x < 0\}.$$

□

Определение. Объединение, пересечение и разность двух множеств $\mathbb{U}$ и $\mathbb{V}$ обозначаются, соответственно, через

$$\mathbb{U} \cup \mathbb{V} := \{x \in \mathbb{U} \text{ или } x \in \mathbb{V}\},$$

$$\mathbb{U} \cap \mathbb{V} := \{x \in \mathbb{U} \text{ и } x \in \mathbb{V}\},$$

$$\mathbb{U} \setminus \mathbb{V} := \{x \in \mathbb{U} : x \notin \mathbb{V}\}.$$

Если $\mathbb{U} \cap \mathbb{V} = \emptyset$, т.е. множества $\mathbb{U}$ и $\mathbb{V}$ не содержат общих элементов, то они называются *непересекающимися*. Если множество $\mathbb{V}$ является подмножеством в $\mathbb{U}$, $\mathbb{V} \subset \mathbb{U}$, то разность $\mathbb{U} \setminus \mathbb{V}$ называется *дополнением к $\mathbb{V}$ в $\mathbb{U}$* . □

Из определения, очевидно, следует

Предложение 1.1.1. Пусть $\mathbb{V} \subset \mathbb{U}$ и $\mathbb{W} = \mathbb{U} \setminus \mathbb{V}$, тогда $\mathbb{W} \cup \mathbb{V} = \mathbb{U}$ и $\mathbb{W} \cap \mathbb{V} = \emptyset$.

Замечание. Операции объединения, пересечения и дополнения соответствуют логическим связкам “или”, “и”, “нет”. □

Предложение 1.1.2. Операции объединения, пересечения и дополнения множеств удовлетворяют тождествам:

$$\begin{aligned} \mathbb{M} \setminus (\mathbb{U} \cup \mathbb{V}) &= (\mathbb{M} \setminus \mathbb{U}) \cap (\mathbb{M} \setminus \mathbb{V}), \\ \mathbb{M} \setminus (\mathbb{U} \cap \mathbb{V}) &= (\mathbb{M} \setminus \mathbb{U}) \cup (\mathbb{M} \setminus \mathbb{V}) \end{aligned} \tag{1.1}$$

и

$$\begin{aligned} \mathbb{M} \cap (\mathbb{U} \cup \mathbb{V}) &= (\mathbb{M} \cap \mathbb{U}) \cup (\mathbb{M} \cap \mathbb{V}), \\ \mathbb{M} \cup (\mathbb{U} \cap \mathbb{V}) &= (\mathbb{M} \cup \mathbb{U}) \cap (\mathbb{M} \cup \mathbb{V}). \end{aligned}$$

□

Доказательство. Выполняется простой проверкой. Рис.1.1 иллюстрирует формулы (1.1). □

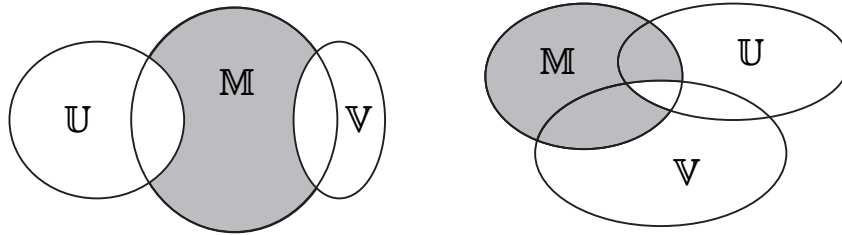


Рис. 1.1: Объединение, пересечение и дополнение множеств. Слева и справа затемнены те области, которые соответствуют равенствам (1.1).

Определение. Пусть $\mathbb{M}_1$ и $\mathbb{M}_2$ – два непустых множества, тогда их *прямым* или *декартовым* произведением $\mathbb{M}_1 \times \mathbb{M}_2$ называется множество упорядоченных пар (x_1, x_2) , где $x_1 \in \mathbb{M}_1$ и $x_2 \in \mathbb{M}_2$:

$$\mathbb{M}_1 \times \mathbb{M}_2 := \{(x_1, x_2) : x_1 \in \mathbb{M}_1, x_2 \in \mathbb{M}_2\}.$$

Если одно из множеств $\mathbb{M}_1$ или $\mathbb{M}_2$ пусто, то их прямое произведение пусто. $\square$

Для того, чтобы понятие множества было более содержательным, на нем вводятся различные структуры. Важным примером является упорядочение.

Определение. Множество $\mathbb{M}$ называется *частично упорядоченным*, если на нем задано бинарное отношение $\leq$, удовлетворяющее следующим условиям:

$$\begin{array}{ll} x \leq x, & \text{– рефлексивность,} \\ x \leq y \text{ и } y \leq z \Rightarrow x \leq z, & \text{– транзитивность,} \\ x \leq y \text{ и } y \leq x \Rightarrow x = y, & \text{– антисимметричность.} \end{array}$$

В общем случае может оказаться, что для некоторой пары элементов соотношение $\leq$ не определено. Если для любой пары элементов $x, y \in \mathbb{M}$ либо $x \leq y$, либо $y \leq x$, то множество $\mathbb{M}$ называется *линейно упорядоченным* или *цепью*. $\square$

Пример 1.1.8. Множество всех подмножеств $\mathcal{P}(\mathbb{M})$ множества $\mathbb{M}$ является частично упорядоченным по отношению включения $\subseteq$. Если множество состоит из двух элементов $\mathbb{M} := \{x, y\}$, то его подмножествами являются $\emptyset$, $\{x\}$, $\{y\}$ и $\{x, y\}$. Оно не является линейно упорядоченным, т.к. для подмножеств $\{x\}$ и $\{y\}$ отношение $\subseteq$ не определено. $\square$

Пример 1.1.9. Множество вещественных чисел $\mathbb{R}$, где символ $\leq$ означает “меньше или равно”, является линейно упорядоченным. $\square$

1.2 Поле вещественных чисел $\mathbb{R}$ и прямая

В геометрии к неопределяемым понятиям относится *прямая линия* или, короче, *прямая*, которую будем обозначать буквой $\mathbb{R}$. При этом ее представляют как отрезок, начерченный по линейке и мысленно продолженный до бесконечности в обе стороны. Прямая линия находится во взаимно однозначном соответствии с полем вещественных (действительных) чисел. Другими словами, прямая является наглядным изображением поля вещественных чисел. На ней выбирается произвольная точка, *начало отсчета*, которой ставится в соответствие число нуль. Затем каждому положительному числу $x \in \mathbb{R}_+$ ставится в соответствие точка, лежащая справа от нуля на расстоянии, равном этому числу. Каждому отрицательному числу $x \in \mathbb{R}_-$ ставится в соответствие точка, лежащая слева от начала отсчета на расстоянии, равном модулю этого числа. При этом расстояние измеряется с помощью линейки, а число x называется координатой точки. Таким образом между точками прямой и вещественными числами устанавливается взаимно однозначное соответствие. Суммируя, можно сказать, что прямая это не более, чем наглядный образ вещественных чисел.

Расстояние l между двумя точками $a \in \mathbb{R}$ и $b \in \mathbb{R}$, по определению, равно модулю разности двух вещественных чисел $l(a, b) = |b - a|$. Его можно записать в виде

интеграла

$$l(a, b) := \left| \int_a^b dxg \right| = |b - a|, \quad g := 1. \quad (1.2)$$

Здесь введена функция g , равная единице, которая называется *метрикой*. Метрика в (1.2) всюду равна единице и называется *евклидовой*. В дифференциальной геометрии понятие расстояния обобщается за счет расширения класса рассматриваемых метрик. Понятие же евклидовой метрики чрезвычайно важно, т.к. служит той точкой отсчета, с которой сравниваются все остальные метрики.

Координаты точек складываются и умножаются так же, как и обычные вещественные числа. Для вещественных чисел определены две операции: сложение и умножение. По отношению к сложению вещественные числа образуют абелеву группу. Напомним общие определения и основные сведения из теории групп.

Определение. Непустое множество $\mathbb{G}$ называется *группой*, если выполнены четыре условия:

- 1) *Закон композиции.* Каждой паре элементов $a, b \in \mathbb{G}$ сопоставляется третий элемент этого же множества, называемый *произведением* элементов и обозначаемый $a \cdot b$ или ab . Закон композиции называется также *бинарной операцией*.
- 2) *Закон ассоциативности.* Для любых трех элементов $a, b, c \in \mathbb{G}$ имеет место равенство

$$(ab)c = a(bc).$$

- 3) В $\mathbb{G}$ существует левая *единица* e :

$$ea = a, \quad \forall a \in \mathbb{G}.$$

- 4) Для каждого элемента $a \in \mathbb{G}$ существует по крайней мере один левый *обратный элемент* $a^{-1} \in \mathbb{G}$:

$$a^{-1}a = e.$$

Множество элементов с одной бинарной операцией, которая удовлетворяет только условию ассоциативности, называется *полугруппой*. Полугруппа с единичным элементом называется *моноидом*. Если для любых двух элементов $ab = ba$, то группа (или полугруппа) называется *коммутативной* или *абелевой*. В противном случае группа (или полугруппа) называется *неабелевой*. $\square$

Пример 1.2.1. Пусть $n\mathbb{Z}$ – множество целых чисел, делящихся на n , где n – произвольное натуральное число. Это множество содержит число 0 при всех n , и в нем определены операции сложения (+) и умножения ($\cdot$). Пара $(n\mathbb{Z}, +)$ является коммутативной группой, где роль единицы выполняет число 0. Обозначим через $n\mathbb{Z}_+$ все неотрицательные числа из $n\mathbb{Z}$. Тогда пара $(n\mathbb{Z}_+, +)$ будет коммутативным моноидом. Если $n \geq 2$, то пара $(n\mathbb{Z}, \cdot)$ является коммутативной полугруппой без единицы. $\square$

Пример 1.2.2. Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовыми координатами x, y, z . Пусть $\mathbb{G}$ – множество всех отображений $\mathbb{R}^3$ в себя. Под композицией двух отображений мы понимаем их последовательное выполнение. В предложении 1.4.1 будет доказано, что композиция отображений произвольного множества является ассоциативной операцией. Множество отображений $\mathbb{G}$ содержит единицу, которой является тождественное отображение, оставляющее точки евклидова пространства неподвижными. Ясно, что множество всех отображений представляет собой некоммутативный моноид. В общем случае множество отображений $\mathbb{G}$ группу

не образует, т.к. содержит, например, проекцию на плоскость x, y , которая не имеет обратного отображения. $\square$

В дифференциальной геометрии изучаются чаще всего такие преобразования многообразий, которые образуют группу. Поэтому опишем некоторые свойства групп более подробно.

Предложение 1.2.1. *Каждая левая единица $e \in \mathbb{G}$ является одновременно и правой единицей, $ae = a$ для всех $a \in \mathbb{G}$. Единица в группе единственна. Каждый левый обратный элемент $a^{-1} \in \mathbb{G}$ одновременно является и правым обратным элементом, $aa^{-1} = e$. Обратный элемент a^{-1} для всех $a \in \mathbb{G}$ единственен. Справедливо правило $(ab)^{-1} = b^{-1}a^{-1}$.*

Доказательство. Приведено в большинстве учебников по теории групп. $\square$

Определение. Группа, состоящая из конечного числа элементов, называется *конечной*. При этом количество элементов в группе называется *порядком* группы. Группа, состоящая из степеней одного элемента a называется *циклической*. Если циклическая группа имеет порядок n , то она обозначается $\mathbb{C}_n$, и $a^n = e$. Циклические группы обязательно абелевы. *Кручением* конечно порожденной абелевой группы $\mathbb{G}$ (т.е. когда любой элемент группы $\mathbb{G}$ представим в виде конечного произведения некоторых элементов и их обратных), называется ее подгруппа, состоящая из всех элементов конечного порядка. Говорят, что конечно порожденная абелева группа не имеет кручения, если в ней нет элементов конечного порядка, т.е. не существует элемента конечная положительная степень которого равна единице. $\square$

Замечание. В дифференциальной геометрии термин кручение имеет совершенно другой смысл (см. раздел 5.3). $\square$

Пример 1.2.3. Множество невырожденных (с отличным от нуля определителем) квадратных $n \times n$ матриц над полем вещественных или комплексных чисел представляют собой группы по отношению к умножению матриц, которые обозначаются $\mathrm{GL}(n, \mathbb{R})$ и $\mathrm{GL}(n, \mathbb{C})$. Эти группы при $n > 1$ некоммутативны (неабелевы). $\square$

Пример 1.2.4. Группа вращений трехмерного евклидова пространства $\mathbb{O}(3)$ и группа Лоренца $\mathbb{O}(1, 3)$ являются неабелевыми. $\square$

Пример 1.2.5. Два целых числа 1 и -1 с обычным умножением образуют циклическую группу $\mathbb{Z}_2$. Ее также отождествляют с группой одномерных вращений, $\mathbb{Z}_2 = \mathbb{O}(1)$. $\square$

Пример 1.2.6. Множество всех целых чисел $\mathbb{Z}$ образует бесконечномерную циклическую группу по отношению к сложению. Она порождена одним элементом – числом 1 – и не имеет кручения. $\square$

Пример 1.2.7. Дробно-линейные преобразования расширенной комплексной плоскости (см. раздел ??)

$$\overline{\mathbb{C}} \ni z \mapsto \frac{az + b}{cz + d} \in \overline{\mathbb{C}}, \quad a, b, c, d \in \mathbb{C}, \quad ad - bc \neq 0 \quad (1.3)$$

образуют группу Ли, которая называется *группой Мёбиуса*. Эта группа шести параметрическая, т.к. числитель и знаменатель можно разделить на произвольное отличное от нуля комплексное число, и изоморфна группе невырожденных комплексных матриц $\mathrm{SL}(2, \mathbb{C})$. $\square$

Определение. Пусть $\mathbb{H} \subset \mathbb{G}$ – подгруппа группы $\mathbb{G}$, т.е. элементы $\mathbb{H}$ сами по себе образуют группу. При этом единица подгруппы $\mathbb{H}$ совпадает с единицей группы $\mathbb{G}$. Обозначим через $a\mathbb{H}$ и $\mathbb{H}a$ множество элементов вида ah и ha , где a – некоторый фиксированный элемент группы $\mathbb{G}$, а элемент $h \in \mathbb{H}$ пробегает всю подгруппу $\mathbb{H}$. Множества элементов $a\mathbb{H}$ и $\mathbb{H}a$, где $a \in \mathbb{G}$, называются *левым* и *правым смежными классами* группы $\mathbb{G}$ по подгруппе $\mathbb{H}$. Если $a \in \mathbb{H}$, то левый и правый смежный класс совпадает с $\mathbb{H}$. $\square$

Нетрудно проверить, что два левых смежных класса по $\mathbb{H}$ либо совпадают, либо не имеют ни одного общего элемента. Поэтому любой элемент группы принадлежит одному и только одному левому смежному классу, и его можно рассматривать как представитель этого класса. Под произведением двух левых смежных классов $a\mathbb{H}$ и $b\mathbb{H}$ понимается множество всех элементов вида cd , где $c \in a\mathbb{H}$ и $d \in b\mathbb{H}$.

Сказанное выше верно и для правых смежных классов.

Определение. Два элемента a' и a группы $\mathbb{G}$ называются *сопряженными*, если они связаны преобразованием *подобия*

$$a' = bab^{-1},$$

где b – некоторый элемент из $\mathbb{G}$. Сопряженность элементов является отношением эквивалентности (см. раздел 1.4) и определяет разбиение группы $\mathbb{G}$ на классы сопряженных элементов. Подгруппа $\mathbb{H}' \subset \mathbb{G}$ называется *сопряженной* подгруппе $\mathbb{H}$, если

$$\mathbb{H}' = b\mathbb{H}b^{-1},$$

для некоторого $b \in \mathbb{G}$. Очевидно, что единичные элементы сопряженных подгрупп $\mathbb{H}$ и $\mathbb{H}'$ совпадают. Подгруппа $\mathbb{H}$ отображается на себя для всех $b \in \mathbb{G}$ тогда и только тогда, когда подгруппа $\mathbb{H}$ содержит все элементы, сопряженные с ее элементами, что можно записать в виде $b\mathbb{H} = \mathbb{H}b$. Такая подгруппа называется *нормальной* подгруппой или *нормальным делителем*. Ее также называют *инвариантной подгруппой*. Левые и правые смежные классы по нормальному делителю совпадают и образуют группу по отношению к операции умножения смежных классов. Действительно, пусть $\mathbb{H}$ – нормальный делитель группы $\mathbb{G}$. Тогда для любых смежных классов $a\mathbb{H}$ и $b\mathbb{H}$ определено умножение

$$a\mathbb{H}b\mathbb{H} = ab\mathbb{H}\mathbb{H} = ab\mathbb{H}.$$

То есть множество смежных классов образует группу само по себе, при этом роль единицы выполняет нормальная подгруппа $\mathbb{H}$. Эта группа называется *факторгруппой* и обозначается $\mathbb{G}/\mathbb{H}$. $\square$

Пример 1.2.8. Если в качестве подгруппы выбрать саму группу, $\mathbb{H} = \mathbb{G}$, то факторгруппа состоит из одного элемента – единицы, $\mathbb{G}/\mathbb{G} = e$. $\square$

Определение. Множество всех элементов группы $\mathbb{G}$, перестановочных с любым элементом этой группы, является нормальным делителем и называется *центром* группы. Центр всегда является абелевой подгруппой группы $\mathbb{G}$.

Для абелевых групп композицию двух элементов часто называют *сложением* и пишут $a + b$. Тогда группу $\mathbb{G}$ называют *аддитивной группой* или *модулем*. Вместо единичного элемента здесь фигурирует *нулевой элемент*:

$$0 + a = a, \quad \forall a \in \mathbb{G},$$

а обратный элемент обозначают $-a$:

$$-a + a = 0. \quad \square$$

Очевидно, что любая подгруппа абелевой группы является нормальным делителем.

Пример 1.2.9. Множество натуральных чисел $\mathbb{N}$ группы по отношению к сложению не образует, т.к. не содержит ни нулевого, ни обратных элементов. $\square$

Пример 1.2.10. Множества целых $\mathbb{Z}$ и вещественных чисел $\mathbb{R}$ образуют аддитивные группы по отношению к сложению. По сути дела, отсюда и пошло название “сложение” для групповой композиции в абелевых группах. $\square$

Пример 1.2.11. По отношению к умножению вещественные числа группу не образуют, т.к. у нуля обратного элемента не существует. В то же время множество вещественных чисел без нуля образует абелеву группу по отношению к умножению, которую мы обозначим $\mathbb{R}^\times$. Эта группа является одномерной группой Ли и состоит из двух компонент связности: отрицательных и положительных чисел $\mathbb{R}_- \cup \mathbb{R}_+$. При этом $e = 1 \in \mathbb{R}_+$. $\square$

Пример 1.2.12. Множество целых чисел с операцией сложения и отношением эквивалентности $n + p \sim n$, где натуральное число $p \in \mathbb{N}$ фиксировано, и $n \in \mathbb{Z}$ любое, представляет собой конечную циклическую группу $\mathbb{Z}_p$, состоящую из p элементов. В качестве элементов группы можно выбрать числа $0, 1, \dots, p-1$. В этом случае мы пишем $n + p = n \bmod p$.

Группу $\mathbb{Z}_p$ можно описать также следующим образом. Обозначим через $p\mathbb{Z}$ подгруппу целых чисел, состоящую из тех чисел, которые делятся на p без остатка. Множество $p\mathbb{Z}$ является нормальной подгруппой в $\mathbb{Z}$. Тогда множества левых и правых смежных классов совпадают и образуют факторгруппу

$$\mathbb{Z}_p := \mathbb{Z}/p\mathbb{Z}.$$

При $p = 2$ эта группа изоморфна группе, рассмотренной в примере 1.2.5. $\square$

Определение. *Прямым произведением* двух групп $\mathbb{G}_1$ и $\mathbb{G}_2$ называется группа $\mathbb{G} = \mathbb{G}_1 \times \mathbb{G}_2$, образованная всеми упорядоченными парами (a_1, a_2) , где $a_1 \in \mathbb{G}_1$ и $a_2 \in \mathbb{G}_2$, с умножением, определяемым формулой

$$(a_1, a_2)(b_1, b_2) := (a_1 b_1, a_2 b_2),$$

для всех $a_1, b_1 \in \mathbb{G}_1$ и $a_2, b_2 \in \mathbb{G}_2$. $\square$

Если группы $\mathbb{G}_1$ и $\mathbb{G}_2$ конечны, то порядок группы $\mathbb{G} = \mathbb{G}_1 \times \mathbb{G}_2$ равен произведению порядков групп $\mathbb{G}_1$ и $\mathbb{G}_2$. Единицей группы $\mathbb{G}$ является пара (e_1, e_2) , где e_1 и e_2 – единицы, соответственно, в группах $\mathbb{G}_1$ и $\mathbb{G}_2$. При этом элементы вида (a_1, e_2) , где a_1 пробегает всю группу $\mathbb{G}_1$, образуют подгруппу в прямом произведении $\mathbb{G}_1 \times \mathbb{G}_2$, которая изоморфна $\mathbb{G}_1$. Аналогично, элементы вида (e_1, a_2) образуют подгруппу в $\mathbb{G}_1 \times \mathbb{G}_2$, изоморфную $\mathbb{G}_2$.

Если у групп $\mathbb{G}_1$ и $\mathbb{G}_2$ нет общих элементов, то элемент их прямого произведения (a_1, a_2) можно записывать просто $a_1 a_2$, причем $a_1 e_2 = a_1$ и $e_1 a_2 = a_2$.

Пример 1.2.13. Любая размерная величина в физике является прямым произведением числа и единицы измерения. Выражения вида “сила = масса $\times$ ускорение” также являются прямыми произведениями. $\square$

Вещественные числа $\mathbb{R}$ представляют собой множество с двумя бинарными операциями. А именно, для двух произвольных чисел $a, b \in \mathbb{R}$ однозначно определена их сумма $a + b$ и произведение ab . Вещественные числа представляют собой частный случай множества, которое называется в алгебре кольцом.

Определение. Непустое множество с двумя бинарными операциями называется *кольцом*, если выполняются следующие условия.

Законы сложения.

- 1) Закон ассоциативности: $a + (b + c) = (a + b) + c$.
- 2) Закон коммутативности: $a + b = b + a$.
- 3) Разрешимость уравнения $a + x = b$ для всех a и b .

Закон умножения.

- 1) Закон ассоциативности: $a(bc) = (ab)c$.

Законы дистрибутивности.

- 1) $a(b + c) = ab + ac$.
- 2) $(b + c)a = ba + ca$.

Если умножение в кольце коммутативно, то говорят о *коммутативном кольце*. $\square$

В современной математике часто рассматриваются неассоциативные кольца, в определении которых отброшено условие ассоциативности. В настоящей монографии такие кольца не рассматриваются.

Три закона сложения означают в совокупности, что элементы кольца образуют абелеву группу (модуль) по отношению к сложению. Если кольцо обладает правым и левым единичным элементом одновременно, $ea = ae = a$, то он называется *единицей*, и говорят о *кольце с единицей*. В общем случае левые и правые единицы у кольца могут различаться, а единиц может быть несколько.

Пример 1.2.14. Минимальное конечное кольцо состоит из двух элементов 0 и 1 со следующей таблицей сложения и умножения

$$\begin{array}{lll} 0 + 0 = 0, & 0 + 1 = 1, & 1 + 1 = 0, \\ 0 \cdot 0 = 0, & 0 \cdot 1 = 0, & 1 \cdot 1 = 1. \end{array}$$

Это кольцо коммутативно, т.к.

$$1 \cdot 0 = 1 \cdot (1 + 1) = 1 \cdot 1 + 1 \cdot 1 = 1 + 1 = 0. \quad \square$$

Пример 1.2.15. Множество целых чисел с операцией сложения и умножения образуют коммутативное кольцо с единицей. $\square$

Определение. Кольцо, для каждого элемента которого $a^2 = a$, называется *булевым*. $\square$

Определение. Кольцо называется *телом*, если:

- 1) в нем есть по крайней мере один элемент, отличный от нуля;
- 2) уравнения $ax = b$ и $xa = b$ при $a \neq 0$ разрешимы.

Коммутативное тело называется *полем* или *рациональным кольцом*. $\square$

Замечание. Тело, как и кольцо, является аддитивной группой (модулем) по отношению к сложению, однако не является группой по отношению к умножению, т.к. у нуля в общем случае нет обратного элемента. Операции сложения и умножения в поле связаны дистрибутивными законами. $\square$

Пример 1.2.16. Множества рациональных $\mathbb{Q}$ и действительных $\mathbb{R}$ чисел с естественными операциями сложения и умножения являются полями. При этом $\mathbb{Q} \subset \mathbb{R}$. Целые числа образуют подкольцо поля рациональных чисел $\mathbb{Z} \subset \mathbb{Q}$. $\square$

Следующий пример показывает, что из одного поля можно построить другие поля.

Пример 1.2.17. В дальнейшем мы будем рассматривать не только вещественные многообразия, но и комплексные. Поэтому напомним основные определения для комплексных чисел. Рассмотрим упорядоченную пару вещественных чисел (a, b) , где $a, b \in \mathbb{R}$. Будем считать две пары (a, b) и (c, d) равными, если $a = c$ и $b = d$. Введем на множестве пар операции сложения и умножения следующим образом:

$$\begin{aligned}(a, b) + (c, d) &:= (a + c, b + d), \\ (a, b)(c, d) &:= (ac - bd, ad + bc).\end{aligned}$$

Нетрудно проверить, что все аксиомы поля для множества пар выполнены.

Определение. Множество пар (a, b) с введенными выше операциями сложения и умножения называется *полем комплексных чисел* и обозначается $z := (a, b) \in \mathbb{C}$. $\square$

Множество комплексных чисел вида $(a, 0) \in \mathbb{C}$ можно отождествить с множеством вещественных чисел $a \in \mathbb{R}$. Тогда поле вещественных чисел является подполем поля комплексных чисел, $\mathbb{R} \subset \mathbb{C}$.

Определим умножение пар $(a, b) \in \mathbb{C}$ на вещественные числа:

$$c(a, b) := (a, b)c := (ca, cb), \quad c \in \mathbb{R}.$$

Множество комплексных чисел с операцией сложения и умножения на вещественные числа образует двумерное векторное пространство. В качестве базиса этого векторного пространства выберем пары:

$$e_1 := (1, 0), \quad e_2 := (0, 1).$$

Теперь произвольное комплексное число можно представить в виде

$$(a, b) = ae_1 + be_2.$$

Базисный вектор e_1 при умножении ведет себя как единица: $ze_1 = e_1z = z$. Поэтому его можно отождествить с единицей. Для второго элемента базиса принято обозначение $i := (0, 1)$. Его называют *мнимой единицей*. Нетрудно проверить, что $i^2 = -1$. Обычно комплексные числа записывают в виде

$$z = a + ib, \quad a, b \in \mathbb{R}, \quad z \in \mathbb{C}.$$

Определение. Множество комплексных чисел $z = (a, b)$ отождествляется с евклидовой плоскостью $\mathbb{C} \approx \mathbb{R}^2$, где вещественные числа a и b рассматриваются в качестве декартовых координат. В этом случае множество комплексных чисел называется *комплексной плоскостью*. При этом ось абсцисс называется *действительной осью*, а ось ординат – *мнимой осью*. $\square$

Линейная структура евклидовой плоскости $\mathbb{R}^2$, рассматриваемой как векторное пространство (см. раздел 1.3.3), совпадает с линейной структурой комплексных чисел $\mathbb{C}$. Однако скалярное умножение векторов в $\mathbb{R}^2$ не имеет никакого отношения к умножению комплексных чисел.

Пусть $z = a + ib \in \mathbb{C}$. Тогда $|z| := \sqrt{a^2 + b^2}$ – это евклидово расстояние от начала координат до точки z . Пусть φ – угол между положительным направлением оси абсцисс и радиус-вектором точки z определяется уравнениями:

$$\cos \varphi = \frac{a}{\sqrt{a^2 + b^2}}, \quad \sin \varphi = \frac{b}{\sqrt{a^2 + b^2}}, \quad |z| \neq 0, \quad \varphi \in [0, 2\pi).$$

Тогда приняты следующие обозначения и названия:

$$\begin{aligned} a &:= \operatorname{re} z && \text{– действительная часть } z, \\ b &:= \operatorname{im} z && \text{– мнимая часть } z, \\ \bar{z} &:= a - ib && \text{– число, комплексно сопряженное к } z, \\ |z| &:= \sqrt{a^2 + b^2} && \text{– модуль } z, \\ \arg z &:= \varphi + 2\pi k, \quad k \in \mathbb{Z}, && \text{– аргумент } z. \end{aligned} \tag{1.4}$$

Функция $\arg z$ определена с точностью до прибавления целого кратного 2π .

Для комплексных чисел часто используют тригонометрическую запись:

$$z = r e^{i\varphi} = r \cos \varphi + ir \sin \varphi, \quad r := |z|.$$

При этом

$$\bar{z} = r e^{-i\varphi} = r \cos \varphi - ir \sin \varphi.$$

Определение. В теории функций комплексного переменного часто используют *расширенную комплексную плоскость* $\overline{\mathbb{C}}$, которая получается из комплексной плоскости $\mathbb{C}$ добавлением бесконечно удаленной точки $z = \infty$. На этом множестве вводится следующая топология. Если подмножество $U \subset \overline{\mathbb{C}}$ не содержит бесконечно удаленной точки ∞ , то оно считается открытым, если оно открыто в $\mathbb{C}$. Если подмножество U содержит точку ∞ , то оно считается открытым в $\overline{\mathbb{C}}$, если его дополнение является компактом в $\mathbb{C}$. Как правило, открытыми окрестностями точки ∞ мы будем считать дополнения $\overline{\mathbb{C}} \setminus \overline{\mathbb{B}}_R$ замкнутых дисков $\overline{\mathbb{B}}_R$ радиуса R до всей расширенной комплексной плоскости $\overline{\mathbb{C}}$. Расширенную комплексную плоскость называют также *комплексной сферой* или *сферой Римана*. $\square$

Комплексная плоскость $\mathbb{C}$ с естественной топологией евклидова пространства является некомпактным многообразием. При добавлении бесконечной точки мы меняем топологию комплексной плоскости (теперь, например, точки $(-\infty, 0)$ и $(\infty, 0)$ близки). В результате получаем компактное топологическое пространство – сферу Римана $\overline{\mathbb{C}}$ (см. раздел 1.3.2). Комплексная сфера является компактификацией комплексной плоскости (см. раздел 1.4.1), которая осуществляется с помощью стереографической проекции.

На сфере Римана $\overline{\mathbb{C}}$ мы определили топологию, которая отличается от топологии комплексной плоскости $\mathbb{C}$. Эта топология является метрической. Соответствующую метрику можно определить следующим образом. С помощью стереографической проекции (см. раздел 29.1) устанавливаем взаимно однозначное соответствие между точками расширенной комплексной плоскости и обычной сферой, вложенной

в трехмерное евклидово пространство, $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. Затем определяем расстояние между точками расширенной комплексной плоскости как обычное евклидово расстояние между точками сферы $\mathbb{S}^2$. Это расстояние определяет на расширенной комплексной плоскости метрику, которая называется *хордовой* и отличается от исходной евклидовой метрики. Явные выражения для хордовой метрики нам не понадобятся, и мы не будем их приводить (см., например, [18], глава 1, §1). Заметим, что при работе с расширенной комплексной плоскостью в ТФКП (теории функций комплексного переменного) используют, как правило, не хордову метрику, которая определяет топологию, а обычную евклидову метрику как более наглядную.

Очевидно, что сфера Римана не несет структуры поля или векторного пространства. Эти структуры можно определить только на подмножестве $\mathbb{C} \subset \overline{\mathbb{C}}$. Тем не менее для каждой точки $p \in \mathbb{C}$ отображение $z \mapsto \frac{1}{z-p}$, определенное обычным образом при $z \in \mathbb{C} \setminus \{p\}$, можно продолжить до биекции множества $\overline{\mathbb{C}}$ на себя, если положить $\frac{1}{0} := \infty$ и $\frac{1}{\infty} := 0$. $\square$

Определение. Говорят, что поле $\mathbb{F}$ имеет *характеристику* p , если существует такое простое число p , что выполнено равенство

$$\underbrace{a + \dots + a}_p = 0, \quad \forall a \in \mathbb{F}.$$

В этом случае поле $\mathbb{F}$ содержит подполе $\mathbb{Z}_p$. Если такого числа p не существует, то говорят, что поле $\mathbb{F}$ имеет характеристику нуль. В последнем случае оно содержит подполе, состоящее из рациональных чисел. $\square$

Пример 1.2.18. Поля рациональных $\mathbb{Q}$, действительных $\mathbb{R}$ и комплексных $\mathbb{C}$ чисел имеют характеристику нуль. $\square$

1.3 Евклидово пространство $\mathbb{R}^n$

Основным понятием дифференциальной геометрии является дифференцируемое многообразие, которое является обобщением евклидова пространства. Евклидово пространство наделено различными структурами: евклидовой метрикой, топологией, структурой векторного и аффинного пространства. Поэтому в настоящем разделе мы изучим это пространство с различных точек зрения, чтобы в дальнейшем было ясно, что именно и как обобщается в дифференциальной геометрии.

1.3.1 $\mathbb{R}^n$ как метрическое пространство

Пусть $\mathbb{R}$ – поле вещественных чисел. В геометрии поле вещественных чисел наглядно изображается в виде прямой линии на рисунках. При этом каждая точка прямой $x \in \mathbb{R}$ находится во взаимно однозначном соответствии с вещественным числом, которое обозначается той же буквой x и называется координатой точки. На прямой выбирается произвольная точка – *начало отсчета*, которой ставится в соответствие число нуль $0 \in \mathbb{R}$. Обычно предполагают, что на рисунках координаты точек упорядочены и возрастают слева направо, что отмечают стрелкой.

Евклидово расстояние l между двумя точками $x, y \in \mathbb{R}$, по определению, равно модулю разности двух вещественных чисел $l(x, y) = |x - y|$.

Определение. Обозначим через $\mathbb{R}^n$ прямое произведение n прямых:

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_n, \quad (1.5)$$

где $n \in \mathbb{N}$ – произвольное натуральное число, которое называется *размерностью* пространства $\mathbb{R}^n$. Точкой $x \in \mathbb{R}^n$ является упорядоченный набор n вещественных чисел $x^\alpha \in \mathbb{R}$, $\alpha = 1, \dots, n$, которые называются *декартовыми координатами* данной точки. Мы записываем координаты точки в виде строки,

$$x = \{x^\alpha\} = (x^1, \dots, x^n) \in \mathbb{R}^n.$$

Каждый из сомножителей, входящих в определение пространства $\mathbb{R}^n$ (1.5), называется *координатной прямой*, а точка с нулевыми координатами $(0, \dots, 0) \in \mathbb{R}^n$ – *началом координат*.

Под 0-мерным пространством $\mathbb{R}^0$ понимают одну точку – число нуль. Одномерное пространство представляет собой прямую, $\mathbb{R}^1 = \mathbb{R}$. Двумерное пространство $\mathbb{R}^2$ называется *плоскостью*. $\square$

Замечание. Номер координаты обозначается с помощью верхнего индекса так же, как и показатель степени. Как правило, различие в значении индексов ясно из контекста. $\square$

В определении пространства $\mathbb{R}^n$ точку и ее координаты можно отождествить. Однако, определив таким образом $\mathbb{R}^n$, мы можем затем перейти в другую систему координат (см. раздел 1.5). Тогда той же точке пространства $\mathbb{R}^n$ будет соответствовать другой набор вещественных чисел. Поэтому следует различать точку пространства $\mathbb{R}^n$ и ее координаты, которые зависят от выбора системы координат.

Расстояние l , т.е. отображение $l : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, между двумя произвольными точками $p, q \in \mathbb{R}^n$ с декартовыми координатами p^α и q^α определяется следующей формулой

$$l(p, q) := |q - p| = \sqrt{(q^1 - p^1)^2 + \dots + (q^n - p^n)^2}. \quad (1.6)$$

Между двумя бесконечно близкими точками x^α и $x^\alpha + dx^\alpha$ расстояние задается *интервалом*, который представляет симметричная квадратичная форма,

$$ds^2 := l^2(x, x + dx) = dx^\alpha dx^\beta g_{\alpha\beta}, \quad (1.7)$$

где компоненты метрики $g_{\alpha\beta}$ в декартовой системе координат не зависят от точки пространства $\mathbb{R}^n$ и представляют собой единичную матрицу, которую будем обозначать следующим образом

$$g_{\alpha\beta} := \delta_{\alpha\beta} = \underbrace{\text{diag}(1, \dots, 1)}_n. \quad (1.8)$$

В формуле (1.7) и в дальнейшем по повторяющимся индексам, один из которых пишется сверху, а другой – снизу, производится суммирование, если не оговорено противное. Это правило называется *правилом суммирования Эйнштейна*. Матрица (1.8) называется *евклидовой метрикой* и имеет специальное обозначение $\delta_{\alpha\beta}$.

В дифференциальной геометрии роль индексов чрезвычайно важна. Поэтому отметим ряд общих правил, которые всюду используются в дальнейшем. Эти правила связаны с группами преобразований, которые действуют на геометрические объекты.

- 1) Каждое слагаемое может содержать некоторый индекс один или два раза. В первом случае он называется *свободным*, а во втором случае – *немой*.
- 2) Если некоторое выражение состоит из суммы нескольких слагаемых, то каждое слагаемое должно содержать один и тот же набор свободных индексов. При этом значения этих индексов во всех слагаемых должно фиксироваться одновременно.
- 3) Немой индекс обязательно встречается один раз сверху и один раз снизу в каждом слагаемом. Значение этого индекса в каждом слагаемом не может быть зафиксировано, т.к. по нему проводится суммирование. В разных слагаемых немые индексы можно обозначать различными буквами, а число их пар может различаться.

Иногда мы все же будем писать повторяющиеся индексы как нижние или верхние. Например, запись $g_{\alpha\alpha}$ обозначает диагональную компоненту метрики, стоящий на α -том месте. При этом суммирование не проводится.

Выше мы ввели евклидову метрику в декартовой системе координат, с помощью которой было определено пространство $\mathbb{R}^n$. С помощью расстояния, т.е. отображения $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, понятие декартовой системы координат можно обобщить следующим образом. Существуют преобразования координат пространства $\mathbb{R}^n$, которые сохраняют вид расстояния (1.6) и, следовательно, интервала (1.7). Эти преобразования образуют неоднородную группу вращений $\mathbb{IO}(n)$, состоящую из вращений пространства $\mathbb{R}^n$,

$$x'^\alpha = x^\beta S_\beta^\alpha, \quad S_\beta^\alpha \in \mathbb{O}(n),$$

и сдвигов,

$$x'^\alpha = x^\alpha + a^\alpha, \quad a^\alpha = \text{const.}$$

Поэтому в дальнейшем под декартовой системой координат мы будем понимать любую систему координат, в которой расстояние между точками пространства $\mathbb{R}^n$ имеет вид (1.8).

Определение. Система координат x^α , $\alpha = 1, \dots, n$ пространства $\mathbb{R}^n$, в которой расстояние между двумя произвольными точками $a, b \in \mathbb{R}^n$ имеет вид (1.6), называется *декартовой системой координат*. $\square$

После того, как пространство $\mathbb{R}^n$ определено, в нем можно строить произвольные криволинейные, например, сферические или цилиндрические системы координат в зависимости от специфики той или иной задачи. В таких системах координат метрика $g_{\alpha\beta}(x)$ в (1.7) будет зависеть от точки пространства $\mathbb{R}^n$.

Определение. Кривой $\gamma = x(t) = \{x^\alpha(t)\}$ в пространстве $\mathbb{R}^n$ называется отображение замкнутого единичного отрезка $[0, 1]$ в пространство $\mathbb{R}^n$,

$$\gamma : [0, 1] \ni t \mapsto x(t) = \{x^\alpha(t)\} \in \mathbb{R}^n, \quad (1.9)$$

где t – вещественный параметр вдоль кривой. Все функции $x^\alpha(t)$ предполагаются достаточно гладкими. Говорят, что кривая соединяет две точки p и q , где $x(0) = p$, $x(1) = q$. Если граничные точки кривой совпадают, $p = q$, то кривая называется *замкнутой*. Совокупность функций $\dot{x}^\alpha$, где точка обозначает дифференцирование по параметру t , определяет касательный вектор к кривой,

$$X_{(\gamma)} = \{X_{(\gamma)}^\alpha := \dot{x}^\alpha\},$$

который называется *вектором скорости* кривой. Кривая γ называется также *путем*, при этом точка p является началом, а q – концом пути. $\square$

Не следует смешивать понятие кривой с множеством точек, через которые она проходит. Согласно определению понятие кривой включает также порядок прохождения точек данного множества.

В общем случае кривая может иметь точки самопересечения.

Определение. Кривая, не имеющая точек самопересечения, называется *простой кривой*. Совпадение начала и конца замкнутой кривой мы не будем рассматривать как точки самопересечения. Поэтому можно говорить о *простой замкнутой кривой*. $\square$

Определение. Кривая γ называется *гладкой (дифференцируемой)*, если все координатные функции $x^\alpha(t)$ являются гладкими (дифференцируемыми). Мы предполагаем, что вектор скорости дифференцируемой кривой отличен от нуля, т.е. отлична от нуля по крайней мере одна из компонент $\dot{x}^\alpha$, что соответствует погружению отрезка $[0, 1]$ в $\mathbb{R}^n$. Кривая γ называется *кусочно гладкой (дифференцируемой)*, если ее можно представить в виде объединения конечного числа простых гладких (дифференцируемых) кривых. $\square$

Согласно данным определениям не всякая гладкая кривая является кусочно гладкой, т.к. может иметь бесконечное число точек самопересечения.

Единичный отрезок в определении кривой выбран для удобства определения произведения путей и фундаментальной группы многообразия (см. раздел 10.1). Выбор другого замкнутого интервала соответствует *перепараметризации кривой*. Под этим понимается замена параметра $t \mapsto t' = t'(t)$, где $t'(t)$ – произвольная достаточно гладкая монотонная функция такая, что $dt'/dt \neq 0$. При этом вектор скорости кривой преобразуется по правилу дифференцирования сложных функций:

$$X_{(\gamma)}^\alpha := \frac{dx^\alpha}{dt} = \frac{dx^\alpha}{dt'} \frac{dt'}{dt}.$$

Определение. *Длиной* дифференцируемой кривой γ , соединяющей точки p и q , называется интеграл

$$l(\gamma) := \int_p^q ds = \int_0^1 dt \sqrt{\dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}}. \quad (1.10)$$

Если кривая является кусочно дифференцируемой, то ее длиной называется сумма длин каждой дифференцируемой части. В евклидовом пространстве $\mathbb{R}^n$ длину кривой $s(t)$ от начала p до текущей точки $x(t) \in \gamma$ всегда можно выбрать в качестве параметра вдоль кривой. В этом случае s называют *каноническим параметром*. Он однозначно определяется уравнением

$$ds = dt \sqrt{\dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}} = dt \sqrt{X_{(\gamma)}^\alpha X_{(\gamma)}^\beta g_{\alpha\beta}},$$

с начальным условием $s(0) = 0$. $\square$

Из формулы (1.10) следует, что определение длины кривой не зависит от выбора ее параметризации.

Если кривая параметризована каноническим параметром s , то вектор скорости $X_{(\gamma)}^\alpha = dx^\alpha/ds$ имеет единичную длину:

$$X_{(\gamma)}^\alpha X_{(\gamma)}^\beta g_{\alpha\beta} = 1.$$

Можно доказать, что расстояние (1.6) является точной нижней гранью интегралов (1.10) по всем возможным путям, соединяющим точки p и q ,

$$l(p, q) = \inf_{\gamma} \int_p^q ds. \quad (1.11)$$

Кривая, вдоль которой интеграл (1.11) принимает наименьшее значение называется отрезком *прямой линии*, соединяющим точки p и q .

Точная нижняя грань (1.11) по всем кривым, соединяющим две точки $p, q \in \mathbb{R}^n$, или явная формула (1.6) задают расстояние или метрику пространства $\mathbb{R}^n$ в топологическом смысле. Напомним определение метрики для произвольного множества.

Определение. Метрикой на множестве $\mathbb{M}$ называется функция

$$\mathbb{M} \times \mathbb{M} \ni p, q \mapsto l(p, q) \in \mathbb{R},$$

определенная для любых двух точек $p, q \in \mathbb{M}$, которая удовлетворяет следующим условиям:

- 1) $l(p, p) = 0$; $l(p, q) > 0$, $(p \neq q)$ – положительная определенность
- 2) $l(p, q) = l(q, p)$ – симметричность,
- 3) $l(p, q) \leq l(p, r) + l(r, q)$, $\forall r \in \mathbb{M}$ – неравенство треугольника.

Значение $l(p, q)$ называется *расстоянием* между точками p и q . Пара $(\mathbb{M}, l)$, т.е. множество $\mathbb{M}$ с заданной метрикой l , называется *метрическим пространством*. $\square$

Замечание. Эту метрику мы будем называть *топологической метрикой*, чтобы отличать ее от дифференциально-геометрической метрики (1.8) и ее обобщения, которое будет сделано в главе 4. $\square$

Понятие топологической метрики очень важно и может быть использовано для определения сходимости последовательностей. Пусть $(\mathbb{M}, l)$ – метрическое пространство, и $\{x_i\}$, $i \in \mathbb{N}$ – последовательность точек в $\mathbb{M}$.

Определение. Назовем $\{x_i\}$ *фундаментальной последовательностью* или *последовательностью Коши*, если для любого $\epsilon > 0$ существует такое натуральное число $N_\epsilon \in \mathbb{N}$, что $l(x_i, x_k) < \epsilon$ при всех $i > N_\epsilon$ и $k > N_\epsilon$. Фундаментальная последовательность, по определению, называется сходящейся к некоторой точке, которая называется *пределом* последовательности. Пространство $\mathbb{M}$ называется *метрически полным*, если любая фундаментальная последовательность в $(\mathbb{M}, l)$ сходится к некоторой точке из $\mathbb{M}$. $\square$

Другими словами полнота означает, что предел любой последовательности из $\mathbb{M}$, если он существует, тоже принадлежит $\mathbb{M}$. Нетрудно показать, что если предел существует, то он единственен.

На одном и том же множестве $\mathbb{M}$ можно задавать различные топологические метрики. Метрика l на множестве $\mathbb{M}$ называется *полной*, если метрическое пространство $(\mathbb{M}, l)$ полное.

Пример 1.3.1. Все пространство $\mathbb{R}^n$, а также любое замкнутое подмножество в нем с метрикой (1.11) является метрически полным. Открытые подмножества $\mathbb{R}^n$, отличные от всего пространства, метрически неполны, т.к. фундаментальные последовательности могут сходиться к граничным точкам, которые не принадлежат данным подмножествам. $\square$

В пространстве $\mathbb{R}^n$, используя понятие расстояния (топологической метрики) (1.6), можно задавать различные подмножества, которые широко используются в дальнейшем и играют большую роль в приложениях. Эти подмножества мы определим в декартовой системе координат.

Пример 1.3.2 (Шар). Подмножество $\mathbb{B}_r^n(p) \subset \mathbb{R}^n$, определяемое неравенством:

$$\mathbb{B}_r^n(p) := \{x \in \mathbb{R}^n : |x - p| < r\}, \quad r = \text{const} > 0, \quad (1.12)$$

называется n -мерным открытым *шаром* радиуса r с центром в точке p . В одномерном случае, $n = 1$, шар называется *интервалом*¹. При $n = 2$ шар $\mathbb{B}_r^2$ называется *диском* или *кругом*.

Если пара $(\mathbb{M}, l)$ – произвольное метрическое пространство, то в нем также можно определить шар радиуса r с центром в точке p :

$$\mathbb{B}_r(p) := \{x \in \mathbb{M} : l(x, p) < r\}. \quad (1.13)$$

Последнее определение не зависит от того имеет ли множество точек $\mathbb{M}$ размерность или нет. $\square$

Пример 1.3.3 (Сфера). Подмножество $\mathbb{S}_r^{n-1}(p) \subset \mathbb{R}^n$, определяемое равенством:

$$\mathbb{S}_r^{n-1}(p) := \{x \in \mathbb{R}^n : |x - p| = r\}, \quad r = \text{const} > 0, \quad (1.14)$$

называется $(n - 1)$ -мерной *сферой* радиуса r с центром в точке p . При $n = 1$ сфера вырождается в две точки, являющиеся концами интервала $(p - r, p + r) \subset \mathbb{R}$. В дальнейшем для сферы единичного радиуса нижний индекс мы будем опускать, $\mathbb{S}^n := \mathbb{S}_1^n$. $\square$

Пример 1.3.4 (Куб). Подмножество $\mathbb{U}_a^n(p) \subset \mathbb{R}^n$, определяемое неравенством:

$$\mathbb{U}_a^n(p) := \{x \in \mathbb{R}^n : |x^\alpha - p^\alpha| < a/2, \alpha = 1, \dots, n\}, \quad a = \text{const} > 0, \quad (1.15)$$

называется n -мерным открытым *кубом* со стороной a и центром в точке p . Куб представляет собой прямое произведение n интервалов $(-a/2, a/2) \subset \mathbb{R}$. При $n = 1$ куб представляет собой интервал, а при $n = 2$ – квадрат. $\square$

Замечание. Условия (1.12)–(1.15) определяют только множества точек шара, сферы и куба, ничего не говоря о том, как устроена топология на этих множествах. Мы всегда предполагаем, что топология шара, сферы, куба, а также других подмножеств евклидова пространства индуцирована их вложением в $\mathbb{R}^n$ (см. следующий раздел). $\square$

¹Этот термин употребляется в дифференциальной геометрии также для обозначения квадрата расстояния между двумя бесконечно близкими точками (1.7)

Пример 1.3.5. Обозначим через $\mathbb{R}_+^n \subset \mathbb{R}^n$ подпространство в $\mathbb{R}^n$, определяемое условием $x^n > 0$. Его замыкание $\overline{\mathbb{R}_+^n}$ имеет край, который является гиперплоскостью $\mathbb{R}^{n-1}$ и определяется уравнением $x^n = 0$. Мы считаем, что топологии на $\mathbb{R}_+^n$ и $\overline{\mathbb{R}_+^n}$ индуцированы вложениями $\mathbb{R}_+^n, \overline{\mathbb{R}_+^n} \hookrightarrow \mathbb{R}^n$. Заметим, что открытые подмножества $\overline{\mathbb{R}_+^n}$ могут содержать точки края. $\square$

Определение. Подмножество $U \subset \mathbb{R}^n$ называется *ограниченным*, если существует шар $\mathbb{B}_r^n$ конечного радиуса r , целиком содержащий U . $\square$

Пример 1.3.6. Все $\mathbb{R}^n$, а также $\mathbb{R}_+^n, \overline{\mathbb{R}_+^n}$ является неограниченными подмножествами в $\mathbb{R}^n$. $\square$

1.3.2 $\mathbb{R}^n$ как топологическое пространство

В предыдущем разделе мы использовали понятия открытого и замкнутого множеств, не дав им определений. Ниже мы восполним этот пробел и приведем необходимые сведения из общей топологии, которые необходимы для определения основного понятия дифференциальной геометрии – многообразия.

Начнем с общего определения топологического пространства и топологии.

Определение. *Топологическим пространством* называется пара $(M, \mathcal{I})$, состоящая из множества точек $x \in M$ и некоторого семейства $\{U_i\}_{i \in I} = \mathcal{I}$ своих подмножеств $U_i \subset M$, которые удовлетворяют следующим условиям:

- 1) пустое множество $\emptyset$ и все множество M принадлежат $\mathcal{I}$;
- 2) пересечение *любой пары* подмножеств из $\mathcal{I}$ принадлежит $\mathcal{I}$;
- 3) объединение *любого семейства* подмножеств из $\mathcal{I}$ принадлежит $\mathcal{I}$.

Элементы семейства $\mathcal{I}$ называются *открытыми множествами* пространства M , а семейство открытых множеств $\mathcal{I}$ – *топологией* пространства M . Некоторое семейство $\{B_j\}_{j \in J} = \mathcal{J}$ открытых множеств называется *базой* топологии пространства M , если каждое множество из $\mathcal{I}$ есть объединение каких-либо множеств из $\mathcal{J}$. $\square$

Ясно, что $M = \cup_{i \in I} U_i$. Множество индексов I может быть произвольным, в том числе несчетным.

Из условия 2) следует, что пересечение любого, но конечного числа открытых множеств также является открытым. В примере 1.3.14 показано, что пересечение бесконечного числа открытых множеств может не быть открытым. В условии 3) допускается объединение бесконечного числа открытых множеств.

Одна и та же топология $\mathcal{I}$ на множестве M может иметь много баз.

На любом множестве можно задать топологию, и не одну. Исключение составляют пустое множество $\emptyset$ (топология состоит из одного открытого множества – самого $\emptyset$) и множество, состоящее из одного элемента. В последнем случае топология единственна и состоит из двух открытых множеств: пустого множества и самого элемента.

Пример 1.3.7. Пусть M – множество. Будем считать, что каждая точка $x \in M$ является открытым множеством и их совокупность образует базу топологии M . Тогда любое подмножество $U \subset M$ будет открытым. Такую топологию называют *дискретной*. $\square$

Пример 1.3.8. Пусть $\mathbb{M}$ – множество. Будем считать, что вся топология $\mathbb{M}$ состоит из двух множеств: пустого множества и всего $\mathbb{M}$. Это – пример другой крайности, и поэтому такую топологию называют *антидискретной*. Антидискретную топологию называют также *тривиальной*, потому что она слишком бедна. $\square$

Определение. Если на множестве $\mathbb{M}$ задано две топологии $\mathcal{I}_1$ и $\mathcal{I}_2$, причем $\mathcal{I}_1 \subset \mathcal{I}_2$, то говорят, что топология $\mathcal{I}_1$ *слабее (грубее)* топологии $\mathcal{I}_2$, или что топология $\mathcal{I}_2$ *сильнее (тоньше)* топологии $\mathcal{I}_1$. $\square$

Очевидно, что дискретная топология является наиболее тонкой, а антидискретная топология – наиболее грубой.

Замечание. Явное описание всех открытых множеств, т.е. топологии на множестве точек $\mathbb{M}$ в общем случае является сложной задачей. Поэтому на практике, чтобы определить топологию в пространстве $\mathbb{M}$, сначала задают базу топологии, а затем объявляют все возможные объединения элементов базы открытыми множествами. Поэтому понятие базы топологии помогает конструктивно подойти к заданию топологии на множестве точек $\mathbb{M}$. $\square$

Пример 1.3.9. Пусть $\mathbb{R}$ – вещественная прямая. Семейство всех открытых интервалов $(a, b) \subset \mathbb{R}$ является базой некоторой топологии на $\mathbb{R}$, которая занимает промежуточное положение между дискретной и антидискретной топологией. Такая топология на вещественной прямой называется *естественной*. Она лежит в основе математического анализа. $\square$

Определение. Пусть $(\mathbb{M}, \mathcal{I})$ и $(\mathbb{N}, \mathcal{J})$ – два топологических пространства. На прямом произведении $\mathbb{M} \times \mathbb{N}$, образованном всеми упорядоченными парами (x, y) , где $x \in \mathbb{M}$ и $y \in \mathbb{N}$, можно ввести топологию. А именно, множество всех пар подмножеств $(\mathbb{U}, \mathbb{V})$, где $\mathbb{U}$ и $\mathbb{V}$ открыты, соответственно, в $\mathbb{M}$ и $\mathbb{N}$, образует базу некоторой топологии в прямом произведении $\mathbb{M} \times \mathbb{N}$. Такое произведение называется *топологическим*. $\square$

Определение. Евклидовым пространством $\mathbb{R}^n$ размерности n называется топологическое произведение n вещественных прямых $\mathbb{R}$:

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_n, \quad (1.16)$$

каждая из которых снабжена естественной топологией. Топологию $\mathbb{R}^n$ также будем называть *естественной*. Будем писать $\dim \mathbb{R}^n = n$. $\square$

Для топологического произведения мы будем использовать тот же символ, что и для прямого, подразумевая, что на прямом произведении задается топология, которая определяется топологией сомножителей. Поэтому для евклидова пространства мы сохранили общепринятое прежнее обозначение $\mathbb{R}^n$.

Базой естественной топологии в евклидовом пространстве $\mathbb{R}^n$ является множество всех кубов (1.15) с произвольными сторонами и центрами.

Аналогично определяется комплексное пространство

$$\mathbb{C}^n := \underbrace{\mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}}_n, \quad (1.17)$$

где $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ – комплексная плоскость с естественной топологией, совпадающей с топологией евклидовой плоскости $\mathbb{R}^2$. Очевидно, что, как топологическое пространство, комплексное пространство $\mathbb{C}^n$ диффеоморфно евклидову пространству вдвое большего числа измерений, $\mathbb{C}^n \approx \mathbb{R}^{2n}$.

Определение. Рассмотрим некоторое подмножество $\mathbb{D} \subset \mathbb{M}$ топологического пространства $(\mathbb{M}, \mathcal{I})$. Будем говорить, что на $\mathbb{D}$ задана *индуцированная топология* $(\mathbb{D}, \mathcal{J})$, если подмножество $\mathbb{V} \in \mathcal{J}$ открыто тогда и только тогда, когда $\mathbb{V} = \mathbb{U} \cap \mathbb{D}$, где $\mathbb{U} \in \mathcal{I}$ – некоторое открытое подмножество в $\mathbb{M}$. $\square$

Пример 1.3.10. Пусть $\mathbb{D} \subset \mathbb{R}^n$ – конечное множество точек евклидова пространства $\mathbb{R}^n$. Тогда топология на $\mathbb{D}$, индуцированная естественной топологией в $\mathbb{R}^n$, является дискретной. $\square$

Замечание. При рассмотрении подмножеств евклидова пространства $\mathbb{R}^n$ как топологических пространств мы будем предполагать, что они снабжены индуцированной топологией, если не оговорено противное. В частности, говоря о шарах, сферах и других подмножествах в $\mathbb{R}^n$, мы предполагаем, что на них задана естественная топология, индуцированная вложением в евклидово пространство. $\square$

Обсудим связь между метрическими пространствами, рассмотренными в предыдущем разделе, и топологическими пространствами. В произвольном метрическом пространстве $(\mathbb{M}, l)$ можно выбрать в качестве базы топологии семейство шаров, определенных формулой (1.12). Тем самым любое метрическое пространство превращается в топологическое. Обратное утверждение неверно. Не на всяком топологическом пространстве можно ввести метрику такую, чтобы она определяла заданную топологию.

Определение. Топологическое пространство $(\mathbb{M}, \mathcal{I})$ называется *метризуемым*, если на множестве его точек $\mathbb{M}$ можно ввести такую метрику $l(p, q)$, что множество всех шаров (1.13) является базой топологии $\mathcal{I}$. $\square$

Пример 1.3.11. Топологическое пространство $\mathbb{M}$ с дискретной топологией является метризуемым. Действительно, введем метрику

$$l(p, q) := \begin{cases} 0, & p = q, \\ 1, & p \neq q, \end{cases}$$

где $p, q \in \mathbb{M}$. Тогда открытый шар, радиус которого меньше или равен единице, состоит из одной точки. Если радиус открытого шара больше единицы, то он состоит из всех точек $\mathbb{M}$. Нетрудно также проверить неравенство треугольника.

Открытый шар единичного радиуса с центром в точке p состоит из одной точки $p = \bar{p}$, а замкнутый – из всех точек $\mathbb{M}$. Таким образом, замыкание открытого шара может существенно отличаться от замкнутого шара того же радиуса и с тем же центром. $\square$

Если на множестве $\mathbb{M}$ заданы две различные топологические метрики (расстояния) l_1 и l_2 , то они могут индуцировать одно и то же топологическое пространство $(\mathbb{M}, \mathcal{I})$.

Наличие евклидова расстояния (1.6) в евклидовом пространстве $\mathbb{R}^n$ позволяет определить шар (1.12). Совокупность всех шаров можно выбрать в качестве базы

некоторой топологии евклидова пространства. Любой куб можно представить в виде объединения бесконечного числа шаров. Поэтому множество всех шаров также является базой естественной топологии евклидова пространства $\mathbb{R}^n$. Таким образом, евклидово пространство является метризуемым. На нем определена евклидова метрика (1.6). В дальнейшем, мы, как правило, предполагаем, что на евклидовом пространстве $\mathbb{R}^n$ задана также евклидова метрика.

Можно показать, что в качестве базы естественной топологии евклидова пространства $\mathbb{R}^n$ можно выбрать шары с центрами в рациональных точках и с рациональными радиусами ([9], глава 1, задача П). Эта база топологии является счетной.

В общем случае, если топологическое пространство $(M, \mathcal{I})$ имеет счетную базу, то говорят, что M удовлетворяет *второй аксиоме счетности*.

Определение. В топологическом пространстве M *окрестностью* точки $x \in M$ называют любое открытое множество в M , содержащее x . Точка x есть *предельная точка* подмножества $U \subset M$, если каждая окрестность точки x содержит точки множества U , отличные от x . $\square$

Обычно в качестве окрестности точки евклидова пространства $x \in \mathbb{R}^n$ выбирается некоторый шар $B^n(x) \subset \mathbb{R}^n$ с центром в этой точке.

Предложение 1.3.1. *Подмножество $U \subset M$ открыто тогда и только тогда, когда каждая точка $x \in U$ имеет окрестность V_x , целиком содержащуюся в U .*

Доказательство. Если подмножество U открыто, то оно же является окрестностью каждой своей точки. Обратно. Если U содержит окрестность V_x каждой своей точки, то подмножество $U = \bigcup_{x \in U} V_x$ открыто, как объединение открытых подмножеств. $\square$

Дадим два эквивалентных определения замкнутого множества.

Определение. Подмножество $U \subset M$ топологического пространства $(M, \mathcal{I})$ *замкнуто*, если

- 1) U есть дополнение в M некоторого открытого множества,
- или
- 2) U содержит все свои предельные точки. $\square$

Нетрудно доказать, что объединение двух и, следовательно, произвольного конечного числа замкнутых множеств является замкнутым множеством. Кроме того, пересечение любого семейства, в том числе бесконечного, замкнутых множеств есть замкнутое множество.

Пустое множество включается в определение топологического пространства и, по определению, является открытым. Отсюда вытекает, что все топологическое пространство, как дополнение пустого множества, является одновременно открытым и замкнутым. Аналогично, само пустое множество также является открытым и замкнутым одновременно. Отдельная точка евклидова пространства $\mathbb{R}^n$ является замкнутым подмножеством как дополнение открытого подмножества.

Замечание. Топологию пространства можно определить в терминах замкнутых подмножеств. Определение аналогично приведенному выше, только открытые множества надо заменить на замкнутые, а операции пересечения и объединения поменять местами. $\square$

Замечание. Отрезок $[0, 1] \subset \mathbb{R}$ является замкнутым подмножеством в $\mathbb{R}$. Однако, если рассматривать отрезок $[0, 1]$ как самостоятельное топологическое пространство с топологией, индуцированной из $\mathbb{R}$, то он будет одновременно и замкнутым, и открытым. $\square$

Определение. *Дискретным подмножеством* топологического пространства $\mathbb{M}$ называется такое его подпространство $\mathbb{U} \subset \mathbb{M}$, каждое подмножество которого замкнуто в $\mathbb{M}$. $\square$

Пример 1.3.12. Набор изолированных точек в евклидовом пространстве $\mathbb{R}^n$ является дискретным подмножеством. Он всегда конечен или счетен. $\square$

Рассмотрим произвольное собственное подмножество $\mathbb{U} \subset \mathbb{M}$, которое не совпадает со всем $\mathbb{M}$. В общем случае оно может не быть ни открытым, ни замкнутым. *Замыканием* некоторого подмножества $\mathbb{U}$ топологического пространства $\mathbb{M}$ называется объединение самого множества $\mathbb{U}$ и всех его предельных точек или, что эквивалентно, пересечение всех замкнутых подмножеств $\mathbb{M}$, содержащих $\mathbb{U}$. Оно обозначается через $\bar{\mathbb{U}}$. Из определения следует, что всегда $\mathbb{U} \subseteq \bar{\mathbb{U}}$.

Предложение 1.3.2. *Точка $x \in \bar{\mathbb{U}}$ тогда и только тогда, когда любая окрестность $\mathbb{V}_x$ точки x имеет непустое пересечение с $\mathbb{U}$, $\mathbb{V}_x \cap \mathbb{U} \neq \emptyset$.*

Доказательство. См., например, [19], лемма 2.7. $\square$

Определение. Если некоторое множество совпадает со своим замыканием, $\mathbb{U} = \bar{\mathbb{U}}$, то оно является замкнутым. *Внутренностью* множества $\mathbb{U}$ называется наибольшее открытое множество $\text{int } \mathbb{U}$, целиком содержащееся в $\mathbb{U}$, или, что эквивалентно, объединение всех открытых множеств, содержащихся в $\mathbb{U}$. Точка $x \in \mathbb{U}$ называется внутренней, если $x \in \text{int } \mathbb{U}$. *Границей* множества $\mathbb{U}$ называется разность

$$\partial \mathbb{U} := \bar{\mathbb{U}} \setminus \text{int } \mathbb{U}.$$

$\square$

Пример 1.3.13. Шар, определенный равенством (1.12), является *открытым* шаром. Замыкание шара $\bar{\mathbb{B}}_r^n$ включает в себя точки $(n - 1)$ -мерной сферы (1.14). При этом сфера $\mathbb{S}_r^{n-1}$ является границей как открытого, так и замкнутого шара. $\square$

Пример 1.3.14. Покажем, что пересечение бесконечного числа открытых множеств может быть замкнутым. Рассмотрим последовательность открытых шаров переменного радиуса $\mathbb{B}_{r_i}^n(0)$, $i \in \mathbb{N}$, в евклидовом пространстве $\mathbb{R}^n$, $n \geq 2$, с центром в начале координат. Пусть все радиусы больше единицы, $r_i > 1$, и их последовательность сходится к единице. Тогда пересечением этой последовательности шаров является замкнутый шар единичного радиуса $\bar{\mathbb{B}}_1^n(0)$, включающий граничную окружность $\mathbb{S}_1^{n-1}(0)$. $\square$

Пример 1.3.15. Объединение бесконечного числа замкнутых подмножеств может быть открыто. Действительно, рассмотрим последовательность замкнутых шаров переменного радиуса $\bar{\mathbb{B}}_{r_i}^n(0)$, $i \in \mathbb{N}$, в евклидовом пространстве $\mathbb{R}^n$, $n \geq 2$, с центром в начале координат. Пусть все радиусы ограничены, $0 < r_i < 1$, и их последовательность сходится к единице. Тогда объединением этой последовательности шаров является открытый шар единичного радиуса $\mathbb{B}_1^n(0)$. $\square$

Предложение 1.3.3. Граница ∂U любого подмножества $U \subset M$ является замкнутым подмножеством.

Доказательство. Поскольку $M \setminus \bar{U}$ и $\text{int } U$ открыты в M , то граница ∂U всегда замкнута, как дополнение открытых множеств. $\square$

Дадим эквивалентное определение внутренних и граничных точек.

Определение. Пусть $U \subset M$ – собственное подмножество. Точка $x \in U$ называется *внутренней*, если подмножество U содержит некоторую окрестность U_x , содержащую точку x . Точка $y \in M$ называется *внешней* по отношению к подмножеству $U \subset M$, если существует окрестность $U_y \ni y$, которая не имеет общих точек с U . Точка $x \in M$ называется *граничной* точкой подмножества $U \subset M$, если любая окрестность $U_x \ni x$ содержит как внутренние, так и внешние точки подмножества $U \subset M$. $\square$

Пример 1.3.16. Если на множестве M определена дискретная топология, то каждое его подмножество является одновременно и открытым, и замкнутым (как дополнение открытого подмножества). Граница каждого подмножества пуста. $\square$

Дадим еще несколько менее наглядных определений.

Определение. Точка $x \in M$ называется *изолированной*, если у нее есть окрестность, не содержащая других точек M . Ясно, что эта окрестность состоит только из одной точки x . Подмножество $U \subset M$ называется *всюду плотным* в M , если его замыкание $\bar{U}$ совпадает со всем M . Эквивалентно, множество U *всюду плотно* в M , если в каждом открытом множестве пространства M содержится хотя бы одна точка из U . $\square$

Пример 1.3.17. Множество рациональных чисел $\mathbb{Q}$ всюду плотно в $\mathbb{R}$. $\square$

Определение. Семейство $\{U_i\}_{i \in I}$ подмножеств пространства M называется *покрытием* множества M , если

$$M = \bigcup_{i \in I} U_i.$$

При этом множество значений индекса $i \in I$ не обязано быть счетным. Если M – топологическое пространство, то покрытие называется *открытым*, если все множества U_i открыты. Мы говорим, что некоторое покрытие имеет *подпокрытие* $\{U_i\}_{i \in J}$, где $J \subset I$, если оно является покрытием само по себе. Если множество значений индекса J является конечным, то подпокрытие будет конечным. Покрытие $\{V_i\}_{i \in I}$ называется *измельчением* покрытия $\{U_i\}_{i \in I}$, если для всех значений индекса $V_i \subset U_i$. При этом некоторые из множеств V_i могут быть пустыми. $\square$

Пример 1.3.18. База топологии как и сама топология на M являются покрытиями. $\square$

Пример 1.3.19. Совокупность интервалов $U_1 = (0, 1)$, $U_2 = (0, 3/4)$, $U_3 = (1/4, 1)$ является покрытием единичного интервала $(0, 1)$. Это покрытие имеет много измельчений, в том числе $V_1 = \emptyset$, $V_2 = (0, 2/3)$, $V_3 = (1/3, 1)$. $\square$

Замечание. В дифференциальной геометрии важную роль играют счетные измельчения покрытий. $\square$

В предыдущем разделе мы определили сходимость последовательностей в метрических пространствах $(\mathbb{M}, l)$. Дадим определение сходимости последовательностей в топологических пространствах $(\mathbb{M}, \mathcal{I})$, которое не опирается на метрическую сходимость.

Определение. Последовательность точек $\{x_i\}$, $i = 1, 2, \dots$, топологического пространства $\mathbb{M}$ сходится к точке $x \in \mathbb{M}$, если для каждой окрестности $\mathbb{U}_x$ точки x существует такое натуральное число $N_{\mathbb{U}}$, что $x_i \in \mathbb{U}_x$ для всех $i > N_{\mathbb{U}}$. $\square$

Теорема 1.3.1. Пусть топологическое пространство $(\mathbb{M}, \mathcal{I})$ метризуемо, т.е. на $\mathbb{M}$ существует метрика l , которая индуцирует данную топологию. Тогда последовательность точек $\{x_i\}$ сходится к точке $x \in \mathbb{M}$ в метрике l тогда и только тогда, когда она сходится в топологии $\mathcal{I}$.

Доказательство. См., например, [20], раздел 4.1. $\square$

Из этой теоремы следует, что для метрических пространств определения сходимости по метрике и индуцированной топологии эквивалентны. В то же время определение сходимости последовательностей в топологических пространствах является более общим, т.к. не всякое топологическое пространство метризуемо. С другой стороны, в метрических пространствах можно ввести важные понятия фундаментальной последовательности и полноты.

Продолжим общее рассмотрение.

Определение. Топологическое пространство $\mathbb{M}$ называется *компактным*, если каждое его открытое покрытие содержит конечное подпокрытие. Хаусдорфово компактное топологическое пространство называется *компактом*. Подмножество $\mathbb{U}$ топологического пространства $\mathbb{M}$ называется *компактным*, если оно компактно в индуцированной топологии. Подмножество $\mathbb{U}$ топологического пространства $\mathbb{M}$ называется *относительно компактным*, если его замыкание $\bar{\mathbb{U}}$ компактно.

Компактные топологические пространства обладают рядом замечательных свойств.

Теорема 1.3.2. Топологическое пространство $\mathbb{M}$ компактно тогда и только тогда, когда каждая направленность в $\mathbb{M}$ имеет предельную точку.

Доказательство. См., например, [20], теорема 3.1.23. $\square$

Напомним

Определение. *Направленностью* в топологическом пространстве $\mathbb{M}$ называется произвольная функция, заданная на непустом направленном множестве Σ , со значениями в $\mathbb{M}$:

$$\Sigma \ni \sigma \mapsto x_\sigma \in \mathbb{M}.$$

Направленность обозначается символом $S = \{x_\sigma, \sigma \in \Sigma\}$, где x_σ – точка в $\mathbb{M}$, соответствующая $\sigma \in \Sigma$. Отношение, направляющее Σ , обозначается знаком неравенства $\leq$ и обладает следующими свойствами:

- 1) если $\sigma_1 \leq \sigma_2$ и $\sigma_2 \leq \sigma_3$, то $\sigma_1 \leq \sigma_3$ (транзитивность);
- 2) для всех $\sigma \in \Sigma$ выполнено неравенство $\sigma \leq \sigma$ (рефлексивность);
- 3) для всех $\sigma_{1,2} \in \Sigma$ существует такой элемент $\sigma \in \Sigma$, что выполнены неравенства $\sigma_1 \leq \sigma$ и $\sigma_2 \leq \sigma$.

Для $\sigma_1, \sigma_2 \in \Sigma$ мы пишем также $\sigma_2 \geq \sigma_1$ вместо $\sigma_1 \leq \sigma_2$.

Точка x называется *пределом направленности* $S = \{x_\sigma, \sigma \in \Sigma\}$, если для каждой окрестности $\mathbb{U}_x$ точки x существует такое σ_0 , что для каждого $\sigma \geq \sigma_0$ выполнено включение $x_\sigma \in \mathbb{U}_x$. В этом случае говорят, что направленность S сходится к точке x . Множество всех пределов направленности S обозначается $\lim S$ или $\lim_{\sigma \in \Sigma} x_\sigma$. В общем случае направленность может иметь несколько пределов. Если направленность S имеет один предел, то мы пишем

$$x = \lim S = \lim_{\sigma \in \Sigma} x_\sigma$$

Точка x называется *предельной точкой направленности*, если для каждой окрестности $\mathbb{U}_x$ точки x и каждого $\sigma_0 \in \Sigma$ существует такое $\sigma \geq \sigma_0$, что выполнено включение $x_\sigma \in \mathbb{U}_x$. $\square$

Направленное множество Σ играет роль индексов.

Пример 1.3.20. Множество, состоящее из конечного числа N элементов, можно направлять. Для этого достаточно пронумеровать все элементы $i = 1, \dots, N$ и упорядочить их как натуральные числа. $\square$

Пример 1.3.21. Пусть $\Sigma = \mathbb{N}$ – множество натуральных чисел $i = 1, 2, \dots \in \mathbb{N}$, направленное обычным отношением $\leq$. Предположим, что задана некоторая функция $\mathbb{N} \ni i \mapsto x_i \in \mathbb{M}$, отображающая натуральные числа в произвольное топологическое пространство $\mathbb{M}$. Ясно, что последовательность точек x_i является направленностью. Таким образом, любая последовательность является направленностью. Тем самым понятие направленности обобщает понятие последовательности. $\square$

Пример 1.3.22. Пусть $\Sigma = \mathbb{N}$ и $x_i := -1/i \in \mathbb{R}$. То есть задана неубывающая последовательность отрицательных чисел $S = \{-1/i, i \in \mathbb{N}\}$, сходящаяся к нулю. $\square$

Пример 1.3.23. Пусть Σ – множество всех отрицательных рациональных чисел, направленное обычным отношением $\leq$, и пусть $x_\sigma := \sigma$ для всех $\sigma \in \Sigma$. Тогда $S = \{x_\sigma, \sigma \in \Sigma\}$ – направленность на вещественной прямой $\mathbb{R}$, сходящаяся к 0. Заметим, что 0 – единственный предел направленности S . $\square$

Пример 1.3.24. Пусть $\mathbb{M}$ – топологическое пространство. Тогда совокупность всех открытых окрестностей точки $x \in \mathbb{M}$ есть направленное множество, если направляющее отношение отождествить с включением $\leq \sim \subseteq$. $\square$

Теорема 1.3.3. Для того, чтобы метризуемое топологическое пространство $\mathbb{M}$ было компактно, необходимо и достаточно, чтобы любая последовательность $\{x_i \in \mathbb{M}, i = 1, 2, \dots\}$ имела хотя бы одну предельную точку.

Доказательство. См., например, [6], утверждение 2.11. $\square$

Теорема 1.3.4. Всякая непрерывная вещественная функция на компактном пространстве ограничена и достигает на нем своего минимального и максимального значения.

Доказательство. См., например, [20], следствие 3.2.9. $\square$

Теорема 1.3.5. Если $\mathbb{U}$ – компактное подмножество хаусдорфова пространства $\mathbb{M}$ и $x \in \mathbb{M} \setminus \mathbb{U}$, то у точки x и подмножества $\mathbb{U}$ существуют непересекающиеся окрестности.

Доказательство. См., например, [9], глава 5, теорема 7. $\square$

Следствие. Каждое компактное подмножество $\mathbb{U}$ хаусдорфова пространства $\mathbb{M}$ замкнуто: $\mathbb{U} = \overline{\mathbb{U}}$.

Теорема 1.3.6. У любых двух непересекающихся компактных подмножеств $\mathbb{U}$ и $\mathbb{V}$ хаусдорфова пространства $\mathbb{M}$ существуют непересекающиеся окрестности.

Доказательство. См., например, [9], глава 5, теорема 9. $\square$

Теорема 1.3.7. Пусть $\{\mathbb{U}_i\}$ – открытое покрытие компактного подмножества $\mathbb{V}$ метрического пространства $(\mathbb{M}, l)$. Тогда существует такое положительное число r , что открытый шар радиуса r с центром в произвольной точке подмножества $\mathbb{U}$ содержится в некотором элементе покрытия $\mathbb{U}_i$.

Доказательство. См., например, [9], глава 5, теорема 26. $\square$

Следствие (Лемма Лебега). Для любого открытого покрытия $\{\mathbb{U}_i\}$ замкнутого интервала вещественных чисел существует такое положительное число r , что если $|x - y| < r$, то в покрытии найдется такой элемент $\mathbb{U}_i$, который содержит обе точки x и y .

Пример 1.3.25. Все евклидово пространство $\mathbb{R}^n$ является некомпактным, т.к. можно выбрать последовательность точек, которая не содержит ни одной предельной точки. Например, множество натуральных чисел, лежащих на любой из координатных прямых не содержит предельной точки. $\square$

Пример 1.3.26. Открытый шар $\mathbb{B}_r^n \subset \mathbb{R}^n$ является некомпактным множеством, т.к. последовательность точек, сходящихся к некоторой граничной точке, не имеет в нем предельной точки. В то же время замкнутый шар $\overline{\mathbb{B}}_r^n$ будет уже компактным множеством. $\square$

Пример 1.3.27. Множество с дискретной топологией компактно тогда и только тогда, когда оно конечно. $\square$

Теорема 1.3.8 (Тихонов). Топологическое произведение произвольного множества непустых топологических пространств является компактным тогда и только тогда, когда каждый сомножитель является компактным пространством.

Доказательство. См. [21]. $\square$

Замечание. Количество сомножителей в сформулированной теореме может быть несчетным. $\square$

Теорема Тихонова позволяет дать критерий компактности подмножеств евклидова пространства с индуцированной топологией.

Теорема 1.3.9 (Гейне–Борель–Лебег). Подпространство $\mathbb{M}$ n -мерного евклидова пространства $\mathbb{R}^n$ является компактом в том и только в том случае, если множество $\mathbb{M}$ замкнуто и ограничено.

Доказательство. См., например, [9], глава 8, теорема 14. $\square$

Пример 1.3.28. Единичный отрезок $[0, 1] \subset \mathbb{R}$ компактен. $\square$

Пример 1.3.29. Окружность $\mathbb{S}^1 \hookrightarrow \mathbb{R}^2$ компактна. $\square$

Определение. Топологическое пространство $\mathbb{M}$ называется *локально компактным*, если каждая точка $x \in \mathbb{M}$ имеет окрестность, замыкание которой компактно. $\square$

Конечно, каждое компактное топологическое пространство является локально компактным. Обратное утверждение неверно.

Пример 1.3.30. Евклидово $\mathbb{R}^n$ и комплексное $\mathbb{C}^n$ пространства некомпактны, но локально компактны. $\square$

Пример 1.3.31. Бесконечномерное гильбертово пространство не является локально компактным. $\square$

В дальнейшем при рассмотрении многообразий необходимо будет воспользоваться существованием разбиения единицы. Для этой цели нам понадобится понятие паракомпактности, которое является слабейшим требованием, достаточным для существования разбиения единицы.

Определение. Топологическое пространство $\mathbb{M}$ называется *паракомпактным*, если любое его покрытие открытыми множествами $\{\mathbb{U}_i\}_{i \in I}$ имеет локально конечное извлечение $\{\mathbb{V}_i\}_{i \in I}$. *Локальная конечность* означает, что для каждой точки $x \in \mathbb{M}$ существует окрестность $\mathbb{W}_x \subset \mathbb{M}$ такая, что $\mathbb{V}_i \cap \mathbb{W}_x \neq \emptyset$ только для конечного числа индексов. Или, любое компактное подмножество $\mathbb{M}$ пересекается с конечным числом открытых множеств $\mathbb{V}_i$. $\square$

Пример 1.3.32. Любое метризуемое топологическое пространство паракомпактно. В частности, евклидово пространство $\mathbb{R}^n$ является паракомпактным. $\square$

Пример 1.3.33. Любое компактное хаусдорфово пространство паракомпактно. $\square$

Перечислим некоторые топологические свойства $\mathbb{R}^n$. Часть этих свойств, например, сепарабельность и хаусдорфовость, наследуется всеми многообразиями. Другие же свойства, такие как связность и односвязность, различны для различных многообразий.

Определение. Топологическое пространство называется *сепарабельным*, если оно содержит счетное всюду плотное подмножество. $\square$

Пример 1.3.34. Евклидово пространство $\mathbb{R}^n$ является сепарабельным, т.к. в качестве счетного всюду плотного подмножества можно выбрать, например, множество точек с рациональными координатами, которое, как известно, счетно при конечном n и всюду плотно в $\mathbb{R}^n$. $\square$

Теорема 1.3.10. Пространство, топология которого обладает счетной базой, сепарабельно.

Доказательство. См., например, [9], глава 1, теорема 14. $\square$

Определение. Топологическое пространство $\mathbb{M}$ называется *хаусдорфовым*, если для любой пары различных точек $x, y \in \mathbb{M}$ существуют открытые подмножества $\mathbb{U}_x$ и $\mathbb{U}_y$, содержащие, соответственно, точки x и y , такие, что $\mathbb{U}_x \cap \mathbb{U}_y = \emptyset$. $\square$

Пример 1.3.35. Евклидово пространство является хаусдорфовым пространством. $\square$

Хаусдорфовость топологического пространства важна при определении предела последовательностей. Если мы хотим, чтобы у любой последовательности точек мог существовать не более, чем один предел, необходимо потребовать, чтобы пространство было хаусдорфовым.

Теорема 1.3.11. *Топологическое пространство является хаусдорфовым тогда и только тогда, когда никакая последовательность в этом пространстве не сходится к двум различным точкам.*

Доказательство. См., например, [9], глава 2, теорема 3. $\square$

Пример 1.3.36. Дадим пример нехаусдорфова топологического пространства. Рассмотрим множество точек, состоящее из объединения вещественной прямой $y = 0$ и точки $(0, 1) \in \mathbb{R}^2$ на евклидовой плоскости, изображенных на рис.1.2,а. Определим

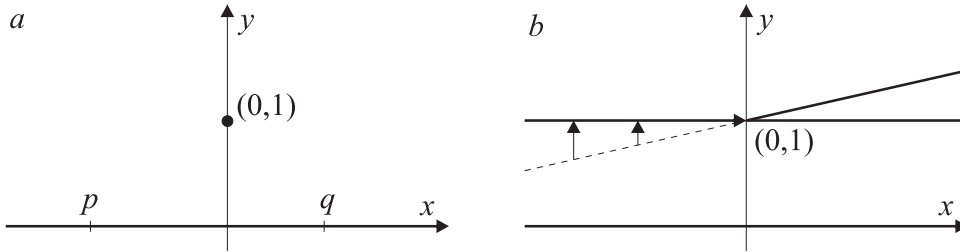


Рис. 1.2: Примеры 1.3.36 и 1.3.37 нехаусдорфовых топологических пространств.

топологию на рассматриваемом множестве следующим образом. Совокупность открытых интервалов (p, q) на оси x , а также множества, состоящие из объединения точки $(0, 1)$ с интервалами (p, q) , $p < 0$, $q > 0$ оси x с выколотым началом координат, будем считать базой топологии. По построению, любые окрестности начала координат $(0, 0) \in \mathbb{R}^2$ и точки $(0, 1) \in \mathbb{R}^2$ имеют непустое пересечение, и, следовательно, построенное топологическое пространство не является хаусдорфовым. Любая последовательность точек, сходящаяся к началу координат $(0, 0)$, сходится также и к точке $(0, 1)$, и наоборот. Другими словами, с точки зрения сходящихся последовательностей две точки $(0, 0)$ и $(0, 1)$ неразличимы. $\square$

Пример 1.3.37. Рассмотрим две прямые на плоскости с естественной топологией, проходящие через точку $(0, 1) \in \mathbb{R}^2$, и отождествим их точки с отрицательными абсциссами, $x < 0$, как показано на рис.1.2,б. В результате получим некоторое топологическое пространство $\mathbb{M}$. Тогда одна точка плоскости $(0, 1) \in \mathbb{R}^2$ соответствует двум различным точкам в $\mathbb{M}$, которые лежат на разных прямых и не отождествлены. Эти точки различны, и в тоже время не имеют непересекающихся окрестностей. $\square$

Замечание. В приведенных примерах нехаусдорфовых пространств их топология не является топологией, индуцированной из $\mathbb{R}^2$. $\square$

Теорема 1.3.12. *Любое метризуемое топологическое пространство $\mathbb{M}$ является хаусдорфовым.*

Доказательство. Выберем две произвольные различные точки $x \neq y$ в $\mathbb{M}$ и построим два непересекающихся шара с центрами в x и y . $\square$

Дадим два эквивалентных определения связности топологических пространств.

Определение. Топологическое пространство $\mathbb{M}$ называется *связным*,

- 1) если его нельзя представить в виде объединения двух непересекающихся множеств, каждое из которых не содержит предельных точек другого; или
- 2) если единственными подмножествами $\mathbb{M}$, открытыми и замкнутыми одновременно, являются $\emptyset$ и $\mathbb{M}$.

Компонентой топологического пространства называется максимальное связное подпространство. $\square$

Если пространство связно, то оно состоит из одной компоненты.

Предложение 1.3.4. Если топологическое пространство является объединением двух непустых непересекающихся открытых подмножеств, $\mathbb{M} = \mathbb{U}_1 \cup \mathbb{U}_2$, где $\mathbb{U}_1 \cap \mathbb{U}_2 = \emptyset$, то оно не является связным.

Доказательство. От противного. Пусть $x \in \mathbb{U}_1$ – предельная точка для $\mathbb{U}_2$. Выберем окрестность $\mathbb{U}_x \subset \mathbb{U}_1$, содержащую точку x . Тогда она содержит точки множества $\mathbb{U}_2$, и, следовательно, подмножества $\mathbb{U}_1$ и $\mathbb{U}_2$ пересекаются. $\square$

Определение. Произвольное открытое связное собственное подмножество $\mathbb{U} \subset \mathbb{R}^n$ называется *областью*. $\square$

В настоящей монографии под областью в дифференцируемом многообразии, которое является основным понятием дифференциальной геометрии (см. раздел 2.1), мы понимаем, как правило, произвольное открытое связное подмножество, которое диффеоморфно открытому шару и, следовательно, покрывается одной системой координат. Такие области диффеоморфны всему евклидову пространству.

Пример 1.3.38. Рассмотрим сферу, вложенную в евклидово пространство, $\mathbb{S}^{n-1} \hookrightarrow \mathbb{R}^n$. Будем считать, что топология на сфере индуцирована вложением. Тогда сфера – связное топологическое пространство. $\square$

Несвязные топологические пространства $\mathbb{M}$ представляют собой объединение конечного или бесконечного числа компонент. В последнем случае число компонент связности может быть счетным или несчетным. Каждая точка принадлежит только одной компоненте, которая называется компонентой данной точки. Всякая компонента является замкнутым подмножеством в $\mathbb{M}$.

Замечание. В общем случае компонента связности топологического пространства может не быть открытой. Экзотический случай описан в [20], пример 6.1.24. $\square$

В дальнейшем будут использованы следующие два утверждения, доказательство которых приведено, например, в [20], следствие 6.1.10 и теорема 6.1.15.

Теорема 1.3.13. Если произвольное семейство $\{\mathbb{U}_i\}, i \in I$, связных подпространств топологического пространства $\mathbb{M}$ имеет непустое пересечение, то его объединение $\bigcup_{i \in I} \mathbb{U}_i$ связно.

Теорема 1.3.14. *Топологическое произведение произвольного множества непустых топологических пространств связно в том и только в том случае, если все сомножители связны.*

Определение. Топологическое пространство $\mathbb{M}$ называется *локально связным*, если для каждой точки $x \in \mathbb{M}$ и любой ее окрестности $\mathbb{U}_x$ существует связное множество $\mathbb{V} \subset \mathbb{U}_x$, такое, что $x \in \text{int } \mathbb{V}$. $\square$

Пример 1.3.39. Евклидово пространство $\mathbb{R}^n$ связно и локально связно. $\square$

Пример 1.3.40. Рассмотрим подмножество точек плоскости $(x, y) \in \mathbb{R}^2$, состоящее из прямых, проходящих через начало координат и заданных уравнениями $mx + ny = 0$ с целыми коэффициентами m и n . Будем считать, что топология этого подмножества индуцирована плоскостью. Это множество связно, как объединение связных множеств, имеющих общую точку. В то же время оно не является локально связным, т.к. любое связное подпространство обязано содержать начало координат. $\square$

Теорема 1.3.15. *В локально связном топологическом пространстве $\mathbb{M}$ каждая компонента является одновременно открытой и замкнутой.*

Доказательство. См., например, [22], теорема 36₆. $\square$

1.3.3 $\mathbb{R}^n$ как векторное пространство

Если на евклидовом пространстве $\mathbb{R}^n$ задана евклидова метрика и индуцированная топология, то этого достаточно для определения дифференцируемого многообразия. Более того, метрику можно и не задавать, понимая под евклидовым пространством просто топологическое произведение n прямых. Вместе с этим на евклидовом пространстве помимо метрики и топологии можно ввести дополнительную алгебраическую структуру – структуру векторного или линейного пространства. Она понадобится для определения касательного расслоения к многообразию. Кроме этого, структура векторного пространства играет большую роль в приложениях, т.к. пространство-время в нерелятивистских моделях и специальной теории относительности имеет структуру векторного пространства.

Пусть $\mathbb{R}^n$ – евклидово пространство с декартовыми координатами x^α , $\alpha = 1, \dots, n$.

Определение. Назовем *вектором* $\overrightarrow{PQ}$ евклидова пространства $\mathbb{R}^n$ с началом в точке $P = \{x_P^\alpha\}$ и концом в точке $Q = \{x_Q^\alpha\}$ упорядоченный набор чисел $\{x_Q^\alpha - x_P^\alpha\}$, которые называются *компонентами вектора*. $\square$

В евклидовой геометрии вектор представляется в виде направленного отрезка прямой линии, соединяющего точки P и Q .

Рассмотрим множество векторов $\mathbb{V}$ с началом в начале системы координат. Для их обозначения будем использовать жирный наклонный шрифт

$$\mathbf{x} = \{x^\alpha\} \in \mathbb{V}.$$

Пространство векторов $\mathbb{V}$ находится во взаимно однозначном соответствии с точками евклидова пространства $\mathbb{R}^n$, при этом вектор $\mathbf{x} \in \mathbb{V}$ называется *радиусом-вектором* точки $\{x^\alpha\} \in \mathbb{R}^n$. Назовем *нулевым* вектором $\mathbf{0}$ вектор с нулевыми компонентами, $\mathbf{0} = (0, \dots, 0) \in \mathbb{V}$. Вектор $-\mathbf{x}$ с компонентами $\{-x^\alpha\} \in \mathbb{V}$ назовем обратным вектором к

вектору $\mathbf{x}$. Для любых двух векторов $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ определим их сумму $\mathbf{x} + \mathbf{y}$ как вектор $\mathbf{z}$ с компонентами

$$\mathbf{z} = \mathbf{x} + \mathbf{y} = \{z^\alpha := x^\alpha + y^\alpha\} \in \mathbb{V}. \quad (1.18)$$

Из свойств сложения вещественных чисел следует, что операция сложения векторов коммутативна: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$. Пространство векторов $\mathbb{V}$ с введенной операцией сложения образует абелеву группу или модуль, поскольку сложение векторов коммутативно и выполнены необходимые групповые аксиомы:

- 1) $\mathbf{x} + \mathbf{y} \in \mathbb{V}$ — замкнутость по отношению к сложению;
 - 2) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ — ассоциативность;
 - 3) $\mathbf{x} + \mathbf{0} = \mathbf{x}$ — существование единичного элемента;
 - 4) $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ — существование обратного элемента.
- (1.19)

Напомним, что для абелевых групп групповую операцию принято называть сложением, а единичный элемент — нулем.

Замечание. Понятие вектора в дифференциальной геометрии является инвариантным и не зависит от выбора системы координат. Приведенные выше определения были сделаны в декартовой системе координат сознательно, т.к. линейная структура наиболее просто задается в декартовых координатах. Векторное пространство $\mathbb{V}$ можно было бы описать и в произвольных координатах. Однако в криволинейной системе координат операция сложения векторов выглядит довольно громоздко. $\square$

Определим новую операцию на множестве векторов $\mathbb{V}$. А именно, введем операцию умножения векторов на действительные числа $a, b, \dots \in \mathbb{R}$ (элементы поля вещественных чисел), которые в данном случае принято называть *скалярами*,

$$\mathbb{R} \times \mathbb{V} \ni a, \mathbf{x} \mapsto a\mathbf{x} := \{ax^\alpha\} \in \mathbb{V}. \quad (1.20)$$

Здесь каждая компонента вектора умножается на одно и то же число a . Эта операция обладает всеми свойствами, которые перечислены в следующем общем определении векторного пространства.

Определение. Абелева группа $\mathbb{V} = \{\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots\}$ с операциями сложения (1.19) и умножения на скаляры (элементы некоторого поля), которые удовлетворяют свойствам:

- 1) $a\mathbf{x} \in \mathbb{V}$ — замкнутость по отношению к умножению на скаляры;
 - 2) $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$ — дистрибутивность по отношению к сложению векторов;
 - 3) $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$ — дистрибутивность по отношению к сложению скаляров;
 - 4) $(ab)\mathbf{x} = a(b\mathbf{x})$ — ассоциативность;
 - 5) $1 \cdot \mathbf{x} = \mathbf{x}$ — умножение на единицу.
- (1.21)

называется *линейным пространством*, *линейным векторным пространством* или просто *векторным пространством*. Элементы пространства $\mathbb{V}$ называются *векторами*. Мы также говорим, что векторное пространство — это модуль над полем вещественных чисел. $\square$

Замечание. В определении сложения векторов евклидова пространства (1.18) и их умножения на скаляры (1.20) мы складываем и умножаем компоненты как вещественные числа. В абстрактном подходе рассматривают множество векторов $\mathbb{V}$ как произвольную абелеву группу, в которой задано умножение на скаляры со свойствами (1.21). В этом случае можно забыть про исходное евклидово пространство и говорить про декартовы или криволинейные координаты не имеет смысла. Если в определении векторного пространства заменить вещественные числа на комплексные, то получим векторное пространство над полем комплексных чисел. $\square$

В произвольном векторном пространстве можно ввести важное понятие *базиса*, как такой набор линейно независимых векторов $\{\mathbf{e}_\alpha\}$, что произвольный вектор можно представить в виде линейной комбинации базисных векторов:

$$\mathbf{x} = x^\alpha \mathbf{e}_\alpha. \quad (1.22)$$

В этом разложении числа $x^\alpha \in \mathbb{R}$ называются компонентами вектора $\mathbf{x}$ по отношению к базису $\mathbf{e}_\alpha$. Мы будем говорить, что векторное пространство $\mathbb{V}$ натянута на векторы $\mathbf{e}_\alpha$, т.е. состоит из всех линейных комбинаций вида (1.22). Базис векторного пространства определен неоднозначно, но число базисных векторов от выбора базиса не зависит. Это число называется *размерностью* векторного пространства, и мы пишем

$$\dim \mathbb{V} = n, \quad n \in \mathbb{N}.$$

Другими словами, размерностью линейного пространства называется максимальное число линейно независимых векторов. Эти векторы можно выбрать в качестве базиса векторного пространства. В общем случае размерность n может быть бесконечна, однако в настоящей монографии мы рассматриваем, как правило, конечномерные векторные пространства. По построению, базисные векторы имеют следующие компоненты:

$$\mathbf{e}_\alpha = (\underbrace{0, \dots, 0}_{\alpha-1}, 1, 0, \dots, 0), \quad (1.23)$$

где на α -ом месте стоит единица, а остальные компоненты равны нулю.

Пусть задано произвольное векторное пространство $\mathbb{V}$, $\dim \mathbb{V} = n$, с фиксированным базисом. Тогда каждый вектор $\mathbf{x} \in \mathbb{V}$ взаимно однозначно задается своими компонентами в этом базисе, $\mathbf{x} \leftrightarrow \{x^\alpha\}$, $\alpha = 1, \dots, n$. Сложение векторов и умножение на скаляры в компонентах задается при этом точно так же, как и для векторов евклидова пространства $\mathbb{R}^n$. Поэтому произвольное векторное пространство $\mathbb{V}$ с фиксированным базисом изоморфно евклидову пространству $\mathbb{R}^n$ той же размерности. Тем самым на любом векторном пространстве можно ввести структуру многообразия, при этом оно становится тривиальным многообразием, диффеоморфном $\mathbb{R}^n$.

Пример 1.3.41. Произвольное поле $\mathbb{F}$ можно рассматривать как одномерное векторное пространство с базисом, состоящем из одного вектора $\mathbf{e}$. Базисным вектором в этом случае может являться любой элемент $\mathbf{e} \in \mathbb{F}$, отличный от нуля. $\square$

Пример 1.3.42. Естественным базисом евклидова пространства $\mathbb{R}^n$, которое рассматривается, как векторное пространство, является набор единичных векторов, направленных вдоль декартовых координатных осей. $\square$

Пример 1.3.43. Множество абсолютно интегрируемых функций на отрезке $[-\pi, \pi]$ с поточечным сложением и умножением на действительные числа является примером бесконечномерного векторного пространства. При этом разложение в ряд Фурье

представляет собой разложение по счетному базису, состоящему из тригонометрических функций. $\square$

Векторные пространства можно строить над любым полем, т.к. в определении линейного пространства от скаляров требуется только наличие сложения, умножения и единицы. В дальнейшем мы часто будем использовать векторные пространства над полем комплексных чисел $\mathbb{C}$. В этом случае произвольный вектор также разлагается по базису $\mathbf{x} = x^\alpha \mathbf{e}_\alpha \in \mathbb{V}$, где $x^\alpha \in \mathbb{C}$ для всех $\alpha = 1, \dots, n$. Число n называется комплексной размерностью векторного пространства $\mathbb{V}$. Мы пишем $\dim_{\mathbb{C}} \mathbb{V} = n$.

Определение. Отображение двух векторных пространств

$$A: \mathbb{V} \ni \mathbf{x} \mapsto \mathbf{y} \in \mathbb{W}$$

называется *линейным*, если

$$A(a\mathbf{x}_1 + b\mathbf{x}_2) = aA(\mathbf{x}_1) + bA(\mathbf{x}_2), \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{V}, \quad a, b \in \mathbb{R}.$$

Отображение A называется также *линейным оператором*. $\square$

Если в пространствах $\mathbb{V}$ и $\mathbb{W}$ заданы базисы $\mathbf{e}_\alpha$, $\alpha = 1, \dots, \dim \mathbb{V}$, и $\mathbf{e}_i$, $i = 1, \dots, \dim \mathbb{W}$, то произвольный линейный оператор задается матрицей A_α^i :

$$A: \mathbb{V} \ni \mathbf{x} = x^\alpha \mathbf{e}_\alpha \mapsto \mathbf{y} = y^i \mathbf{e}_i \in \mathbb{W}, \quad y^i = x^\alpha A_\alpha^i,$$

где $A(\mathbf{e}_\alpha) := A_\alpha^i \mathbf{e}_i$.

Если компоненты всех векторов (1.22) преобразовать с помощью невырожденной матрицы

$$x^\alpha \mapsto x^\beta A_\beta^\alpha, \quad \det A \neq 0, \quad (1.24)$$

то получим взаимно однозначное отображение векторного пространства на себя, при котором все линейные свойства пространства сохраняются. То есть каждая невырожденная матрица задает некоторый автоморфизм векторного пространства. Верно и обратное утверждение. Любой автоморфизм векторного пространства в компонентах задается некоторой невырожденной матрицей. Совокупность матриц A образует группу, которая называется группой линейных однородных преобразований векторного пространства и обозначается $\mathrm{GL}(n, \mathbb{R})$. При преобразовании координат (1.24) с произвольной матрицей $A \in \mathrm{GL}(n, \mathbb{R})$ нулевой вектор не меняется. Поэтому группа общих линейных преобразований действует эффективно, но не свободно (см. раздел 9.1).

Пусть в векторном пространстве $\mathbb{V}$ задан некоторый базис $\mathbf{e}_\alpha$, $\alpha = 1, \dots, n$. Вообще говоря, базис определен с точностью до произвольного $\mathrm{GL}(n, \mathbb{R})$ преобразования (вращения). Поэтому два произвольных базиса $\mathbf{e}_\alpha$ и $\mathbf{e}'_\alpha$ связаны между собой некоторой матрицей перехода:

$$\mathbf{e}_\alpha \mapsto \mathbf{e}'_\alpha = S_\alpha^\beta \mathbf{e}_\beta, \quad S \in \mathrm{GL}(n, \mathbb{R}).$$

При преобразовании базиса компоненты вектора преобразуются с помощью обратной матрицы так, чтобы сам вектор не изменился:

$$\mathbf{x} = x'^\alpha \mathbf{e}'_\alpha = x^\alpha S_\alpha^{-1\beta} S_\beta^\gamma \mathbf{e}_\gamma.$$

Нетрудно видеть, что при таком преобразовании базиса компоненты линейного оператора $A : V \rightarrow V$ преобразуются по правилу подобия:

$$A \mapsto SAS^{-1}.$$

В определении векторного пространства присутствует два типа объектов: векторы, которые образуют абелеву группу (модуль) по отношению к сложению, и скаляры. На множестве скаляров определены две операции: сложение и умножение. Если поле ($\mathbb{R}$ или $\mathbb{C}$) в определении векторного пространства заменить на произвольное кольцо $\mathbb{k}$, то получим $\mathbb{k}$ -модуль, который является одним из основных объектов изучения общей алгебры.

Определение. Пусть задано произвольное кольцо $\mathbb{k}$ и произвольная абелева группа M . Тогда левым $\mathbb{k}$ -модулем называется множество элементов M , на котором задано отображение

$$\mathbb{k} \times M \ni a, x \mapsto ax \in M,$$

которое удовлетворяет следующим условиям:

- 1) $a(x + y) = ax + ay$ – дистрибутивность по отношению к сложению векторов;
- 2) $(a + b)x = ax + bx$ – дистрибутивность по отношению к сложению скаляров;
- 3) $(ab)x = a(bx)$ – ассоциативность.

Если кольцо содержит единицу $1 \in \mathbb{k}$, то дополнительно требуют выполнения условия $1 \cdot x = x$ для всех $x \in M$. Такой $\mathbb{k}$ -модуль называется *унитарным* или *унитальным*. $\square$

Аналогично определяется правый $\mathbb{k}$ -модуль, причем условие ассоциативности записывается в виде $x(ab) = (xa)b$. Если кольцо $\mathbb{k}$ коммутативно, то левые и правые $\mathbb{k}$ -модули можно отождествить. Любой модуль можно рассматривать и над алгеброй, в которой также присутствуют сложение и умножение.

Пример 1.3.44. Пусть задана абелева группа G и кольцо целых чисел $\mathbb{Z}$. Определим произведение

$$\mathbb{Z} \times G \ni m, x \mapsto \underbrace{x + \dots + x}_m \in G.$$

Тогда любая абелева группа превращается в модуль над кольцом целых чисел. $\square$

Пример 1.3.45. Векторное пространство над полем вещественных $\mathbb{R}$ или комплексных $\mathbb{C}$ чисел является модулем, соответственно, над $\mathbb{R}$ или $\mathbb{C}$. $\square$

Пример 1.3.46. Множества гладких дифференциальных форм на многообразии $\Lambda(M)$ (см. раздел 3.2) и гладких векторных полей $\mathcal{X}(M)$ (см. раздел 2.6.2) образуют $C^\infty(M)$ -модули над алгеброй гладких функций. $\square$

В общем случае $\mathbb{k}$ -модули устроены сложнее, чем векторные пространства. Например, не в каждом $\mathbb{k}$ -модуле можно выбрать базис, и даже те, в которых это возможно, могут иметь несколько базисов с различным числом элементов (в случае некоммутативного кольца).

Вернемся к обсуждению векторных пространств. Если некоторое подмножество векторного пространства $V_1 \subset V$ само является векторным пространством, то оно называется *линейным подпространством*. Пусть V_1 и V_2 – два линейных подпространства векторного пространства V . Нетрудно проверить, что их пересечение $V_1 \cap V_2$

также является векторным пространством. В то же время их объединение в общем случае векторным пространством не будет. Тем не менее мы будем писать $\mathbb{V}_1 \cup \mathbb{V}_2$, когда понадобится необходимость сказать, что некоторый элемент принадлежит либо пространству $\mathbb{V}_1$, либо пространству $\mathbb{V}_2$.

Пример 1.3.47. Пусть $\mathbb{V} = \mathbb{R}^3$. Рассмотрим две плоскости $\mathbb{V}_1$ и $\mathbb{V}_2$ натянутые на базисные векторы $(\mathbf{e}_1, \mathbf{e}_2)$ и $(\mathbf{e}_2, \mathbf{e}_3)$, соответственно. Тогда их пересечение $\mathbb{V}_1 \cap \mathbb{V}_2$ является прямой линией с направляющим вектором $\mathbf{e}_2$. Тем самым пересечение является одномерным векторным пространством. Объединение $\mathbb{V}_1 \cup \mathbb{V}_2$ представляет собой объединение двух перпендикулярных плоскостей, что, конечно, не совпадает с трехмерным евклидовым пространством $\mathbb{R}^3$ и не является векторным пространством. $\square$

Очевидно, что

$$\dim(\mathbb{V}_1 \cap \mathbb{V}_2) \leq \min\{\dim \mathbb{V}_1, \dim \mathbb{V}_2\}.$$

Определение. Два векторных подпространства $\mathbb{V}_1$ и $\mathbb{V}_2$ конечномерного векторного пространства $\mathbb{V}$ называются *трансверсальными* друг к другу, если они порождают все $\mathbb{V}$, т.е.

$$\dim(\mathbb{V}_1 \cap \mathbb{V}_2) + \dim \mathbb{V} = \dim \mathbb{V}_1 + \dim \mathbb{V}_2. \quad \square$$

В частном случае, если подпространства не пересекаются, $\mathbb{V}_1 \cap \mathbb{V}_2 = \emptyset$, то $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2$.

Векторное пространство $\mathbb{V}$, $\dim \mathbb{V} = n$, над полем вещественных чисел образует абелеву группу по отношению к сложению или модуль. Рассмотрим его подпространство $\mathbb{H} \subset \mathbb{V}$, состоящее из всех линейных комбинаций

$$\sum_{i=1}^p a^i \mathbf{x}_i \in \mathbb{H}, \quad a^i \in \mathbb{R}, \quad \mathbf{x}_i \in \mathbb{V},$$

где $p < n$. Не ограничивая общности, можно считать, что все векторы $\mathbf{x}_i$ линейно независимы. Тем самым мы предполагаем, что $\dim \mathbb{H} = p$. Подпространство $\mathbb{H}$ образует подгруппу векторного пространства $\mathbb{V}$ или подмодуль. Эта подгруппа является нормальной, поскольку группа абелева. Отсюда следует, что существует факторгруппа или факторпространство $\mathbb{V}/\mathbb{H}$. Элементами факторпространства являются все линейные комбинации вида

$$\mathbf{x} + \sum_{i=1}^p a^i \mathbf{x}_i \in \mathbf{x} + \mathbb{H}, \quad \text{где } \mathbf{x} \in \mathbb{V},$$

со следующим отношением эквивалентности. Два элемента факторпространства $\mathbf{x} + \mathbb{H}$ и $\mathbf{y} + \mathbb{H}$ совпадают, если их разность лежит в $\mathbb{H}$:

$$\mathbf{y} - \mathbf{x} \in \mathbb{H}.$$

В факторпространстве $\mathbb{V}/\mathbb{H}$ естественным образом вводится умножение на вещественные числа, что превращает его само в векторное пространство. Очевидно, что $\dim \mathbb{V}/\mathbb{H} = \dim \mathbb{V} - \dim \mathbb{H} = n - p$.

Пример 1.3.48. Пусть подпространство $\mathbb{H} \subset \mathbb{V}$ натянуто на $p < n$ первых базисных векторов векторного пространства $\mathbb{V}$

$$\mathbb{H} = \{\mathbf{x} \in \mathbb{V} : \mathbf{x} = \sum_{i=1}^p a^i \mathbf{e}_i\}.$$

Тогда факторпространство $\mathbb{V}/\mathbb{H}$ изоморфно векторному пространству, натянутому на оставшиеся базисные векторы:

$$\frac{\mathbb{V}}{\mathbb{H}} \simeq \{\mathbf{x} \in \mathbb{V} : \mathbf{x} = \sum_{k=p+1}^n a^k \mathbf{e}_k\}. \quad \square$$

Определение. *Прямой линией*, проходящей через две различные точки векторного пространства $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ называется множество векторов

$$\{\mathbf{z} \in \mathbb{V} : \mathbf{z} = t\mathbf{x} + (1-t)\mathbf{y}, \quad \forall t \in \mathbb{R}\}. \quad (1.25)$$

Подмножество векторного пространства $\mathbb{U} \subset \mathbb{V}$ называется *плоским*, если вместе с любыми двумя точками, оно содержит прямую, проходящую через эти точки. Если параметр пробегает единичный отрезок, $t \in [0, 1]$, то мы имеем отрезок, соединяющий точки $\mathbf{x}$ и $\mathbf{y}$. Подмножество точек $\mathbb{U}$ векторного пространства называется *выпуклым*, если вместе с двумя произвольными точками $\mathbf{x}, \mathbf{y} \in \mathbb{U}$ оно содержит и отрезок, соединяющий эти точки. $\square$

Прямая сумма

Если задано два векторных пространств $\mathbb{V}_1$ и $\mathbb{V}_2$, то с их помощью можно построить новые векторные пространства: прямую сумму и тензорное произведение.

Определение. *Прямой суммой* двух векторных пространств называется векторное пространство $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2$ образованное всеми парами $\mathbf{x}_1 \oplus \mathbf{x}_2$, где $\mathbf{x}_1 \in \mathbb{V}_1$ и $\mathbf{x}_2 \in \mathbb{V}_2$, с векторным сложением и умножением на скаляры, определяемыми формулами:

$$\begin{aligned} \mathbf{x}_1 \oplus \mathbf{x}_2 + \mathbf{y}_1 \oplus \mathbf{y}_2 &:= (\mathbf{x}_1 + \mathbf{y}_1) \oplus (\mathbf{x}_2 + \mathbf{y}_2), \\ a(\mathbf{x}_1 \oplus \mathbf{x}_2) &:= (a\mathbf{x}_1) \oplus (a\mathbf{x}_2), \end{aligned}$$

где $\mathbf{y}_1 \oplus \mathbf{y}_2$ также принадлежит $\mathbb{V}_1 \oplus \mathbb{V}_2$. $\square$

Нетрудно проверить, что все аксиомы векторного пространства для прямой суммы выполнены.

Размерность прямой суммы векторных пространств равна сумме размерностей пространств $\mathbb{V}_1$ и $\mathbb{V}_2$:

$$\dim(\mathbb{V}_1 \oplus \mathbb{V}_2) = \dim \mathbb{V}_1 + \dim \mathbb{V}_2.$$

Пусть в векторных пространствах заданы базисы $\mathbf{e}_{\alpha_1}$, $\alpha_1 = 1, \dots, \dim \mathbb{V}_1$, и $\mathbf{e}_{\alpha_2}$, $\alpha_2 = 1, \dots, \dim \mathbb{V}_2$. Тогда прямые суммы $\mathbf{e}_{\alpha_1} \oplus \mathbf{0}$ и $\mathbf{0} \oplus \mathbf{e}_{\alpha_2}$ образуют базис в $\mathbb{V}_1 \oplus \mathbb{V}_2$. Компонентами прямой суммы двух векторов в этом базисе является упорядоченный набор $\dim \mathbb{V}_1 + \dim \mathbb{V}_2$ чисел, состоящий из компонент первого и второго вектора:

$$\mathbf{x}_1 \oplus \mathbf{x}_2 = \{x_1^{\alpha_1}, x_2^{\alpha_2}\}.$$

Пример 1.3.49. Вещественную прямую $\mathbb{R}$ можно естественным образом рассматривать, как одномерное векторное пространство. Тогда векторное пространство $\mathbb{R}^n$ есть прямая сумма n одномерных векторных пространств.

$$\mathbb{R}^n = \underbrace{\mathbb{R} \oplus \dots \oplus \mathbb{R}}_n. \quad \square$$

Векторные пространства $\mathbb{V}_1$ и $\mathbb{V}_2$ можно рассматривать, как подпространства $\mathbb{V}_1 \oplus \mathbf{0}$ и $\mathbf{0} \oplus \mathbb{V}_2$ в прямой сумме $\mathbb{V}_1 \oplus \mathbb{V}_2$. Поэтому они изоморфны следующим факторпространствам

$$\mathbb{V}_1 \simeq \frac{\mathbb{V}_1 \oplus \mathbb{V}_2}{\mathbb{V}_2} \quad \text{и} \quad \mathbb{V}_2 \simeq \frac{\mathbb{V}_1 \oplus \mathbb{V}_2}{\mathbb{V}_1}$$

Определение. *Проекционным оператором* или *проектором* в векторном пространстве $\mathbb{V}$ называется отображение $\Pi : \mathbb{V} \rightarrow \mathbb{V}$, которое удовлетворяет условию $\Pi^2 := \Pi \circ \Pi = \Pi$. $\square$

Если векторное пространство разлагается в прямую сумму $\mathbb{V} = \bigoplus_{A=1}^N \mathbb{V}_A$, то этому разложению естественным образом ставится в соответствие N проекторов, которые определяются следующим образом. Каждый вектор однозначно разлагается в прямую сумму $\mathbf{x} = \bigoplus_{A=1}^N \mathbf{x}_A \in \mathbb{V}$, где $\mathbf{x}_A \in \mathbb{V}_A$. Поэтому положим

$$\Pi_A(\mathbf{x}) := \mathbf{x}_A.$$

Легко проверить линейность операторов и свойство $\Pi_A^2 = \Pi_A$. Очевидно, что образом векторного пространства $\mathbb{V}$ при отображении Π_A является подпространство $\mathbb{V}_A \subset \mathbb{V}$. Кроме того из построения следует, что

$$\sum_{A=1}^N \Pi_A = \text{id}, \quad \Pi_A \Pi_B = 0 \quad \text{при} \quad A \neq B. \quad (1.26)$$

Верно также обратное построение.

Предложение 1.3.5. *Пусть задан набор проекторов $\Pi_A : \mathbb{V} \rightarrow \mathbb{V}$, удовлетворяющих условиям (1.26). Положим $\mathbb{V}_A := \Pi_A \mathbb{V}$. Тогда $\mathbb{V} = \sum_{A=1}^N \mathbb{V}_A$.*

Доказательство. См., например, [12], §5, теорема 10. $\square$

Обобщением разложения вектора по базису (1.22) служит градуировка векторного пространства.

Определение. *Градуированным векторным (линейным) пространством* называется векторное пространство $\mathbb{V}$ вместе с его разложением в прямую сумму подпространств

$$\mathbb{V} = \bigoplus_{i=1}^{\infty} \mathbb{V}_i, \quad (1.27)$$

где некоторые подпространства могут быть пустыми. Суммарная размерность подпространств в сумме должна быть равна размерности $\mathbb{V}$. В общем случае бесконечномерных векторных пространств эта сумма бесконечна, но каждый отдельный вектор $\mathbf{x} \in \mathbb{V}$ однозначно представляется в виде конечной суммы

$$\mathbf{x} = \sum_{i=1}^{\infty} \mathbf{x}_i, \quad \mathbf{x}_i \in \mathbb{V}_i.$$

где все $\mathbf{x}_i$ кроме конечного числа равны нулю. Вектор $\mathbf{x}_i$ называется *однородной компонентой* вектора $\mathbf{x}$ степени i . Если $\mathbf{x} \in \mathbb{V}_i$, то вектор $\mathbf{x}$ называется *однородным элементом* степени i . $\square$

Пример 1.3.50. Евклидово векторное пространство $\mathbb{R}^n$ представляет собой прямую сумму n прямых, что можно рассматривать как градуировку. При этом количество подпространств в разложении (1.27) является конечным. На одном и том же векторном пространстве можно задать несколько различных градуировок. Например, евклидово пространство можно представить в виде суммы двух подпространств:

$$\mathbb{R}^n = \mathbb{R}^m \oplus \mathbb{R}^{n-m}, \quad 1 \leq m \leq n-1. \quad \square$$

Определение. Одним из способов изучения множеств $\mathbb{M}$ с заданными алгебраическими структурами состоит в выделении в них последовательности подмножеств $\mathbb{M}_0 \subset \mathbb{M}_1 \subset \mathbb{M}_2 \subset \dots$ или $\mathbb{M}_0 \supset \mathbb{M}_1 \supset \mathbb{M}_2 \supset \dots$ таким образом, что переход от одного подмножества к другому устроен к какому-то смыслу просто. Общее название таких последовательностей – *фильтрации* (*возрастающая и убывающая* соответственно). В теории линейных пространств строго возрастающая последовательность подпространств $\mathbb{V}_0 \subset \mathbb{V}_1 \subset \dots \subset \mathbb{V}_n$ пространства $\mathbb{V}$ называется *флагом*. Число n называется длиной флага. Флаг $\mathbb{V}_0 \subset \mathbb{V}_1 \subset \dots \subset \mathbb{V}_i \subset \dots$ называется *максимальным*, если $\mathbb{V}_0 = \{0\}$, $\cup_i \mathbb{V}_i = \mathbb{V}$ и между $\mathbb{V}_i$ и $\mathbb{V}_{i+1}$ нельзя вставить подпространство, т.е. если $\mathbb{V}_i \subset \mathbb{U} \subset \mathbb{V}_{i+1}$, то либо $\mathbb{U} = \mathbb{V}_i$, либо $\mathbb{U} = \mathbb{V}_{i+1}$. $\square$

Замечание. Мотивировка названия: флаг $\{\text{точка } 0\} \subset \{\text{прямая}\} \subset \{\text{плоскость}\}$ – это “гвоздь”, “древко” и “полотнище”. $\square$

По всякому базису $\mathbf{e}_\alpha$, $\alpha = 1, \dots, n$, векторного пространства $\mathbb{V}$, $\dim \mathbb{V} = n$, можно построить флаг следующим образом. Положим $\mathbb{V}_0 := \{0\}$ и пусть $\mathbb{V}_i$ – линейная оболочка первых i базисных векторов $\{\mathbf{e}_1, \dots, \mathbf{e}_i\}$ при $1 \leq i \leq n$. Нетрудно проверить, что построенный флаг максимален.

Предложение 1.3.6. *Размерность векторного пространства $\mathbb{V}$ равна длине любого его максимального флага.*

Доказательство. См., например, [12], часть 1, §2, теорема 15. $\square$

В конечномерном линейном пространстве $\mathbb{V}$ любой флаг можно дополнить до максимального, и поэтому его длина всегда меньше или равна $\dim \mathbb{V}$.

Тензорная алгебра

Перейдем к описанию более сложного понятия тензорного произведения векторных пространств. Обозначим через $\mathbb{V}_1 \times \mathbb{V}_2$ прямое произведение векторных пространств. Оно не снабжено структурой линейного пространства и просто обозначает множество упорядоченных пар элементов $(\mathbf{x}_1, \mathbf{x}_2)$, где $\mathbf{x}_1 \in \mathbb{V}_1$ и $\mathbf{x}_2 \in \mathbb{V}_2$.

Определение. Пусть $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ – векторное пространство над полем вещественных чисел, свободно порожденное элементами вида $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{V}_1 \times \mathbb{V}_2$. То есть $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ состоит из всех конечных линейных комбинаций пар $(\mathbf{x}_1, \mathbf{x}_2)$. А именно, мы рассматриваем все конечные линейные комбинации

$$\sum_{ij} a_{ij}(\mathbf{x}_{1i}, \mathbf{x}_{2j}), \quad \mathbf{x}_{1i} \in \mathbb{V}_1, \quad \mathbf{x}_{2j} \in \mathbb{V}_2, \quad a_{ij} \in \mathbb{R}.$$

Построенное векторное пространство $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ является бесконечномерным, т.к. базисом этого пространства являются все элементы вида $(\mathbf{x}_1, \mathbf{x}_2)$, которых бесконечно много.

Пусть $\mathbb{H}(\mathbb{V}_1, \mathbb{V}_2)$ – линейное подпространство в $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$, свободно порожденное всеми элементами вида

$$\begin{aligned} (\mathbf{x}_1, \mathbf{x}_2 + \mathbf{y}_2) - (\mathbf{x}_1, \mathbf{x}_2) - (\mathbf{x}_1, \mathbf{y}_2), \\ (\mathbf{x}_1 + \mathbf{y}_1, \mathbf{x}_2) - (\mathbf{x}_1, \mathbf{x}_2) - (\mathbf{y}_1, \mathbf{x}_2), \\ (a\mathbf{x}_1, \mathbf{x}_2) - a(\mathbf{x}_1, \mathbf{x}_2), \\ (\mathbf{x}_1, a\mathbf{x}_2) - a(\mathbf{x}_1, \mathbf{x}_2), \end{aligned} \quad (1.28)$$

где $a \in \mathbb{R}$. Это пространство также бесконечномерно. Тогда факторпространство

$$\mathbb{V}_1 \otimes \mathbb{V}_2 := \frac{\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)}{\mathbb{H}(\mathbb{V}_1, \mathbb{V}_2)}$$

называется *тензорным произведением* векторных пространств. При этом каждой паре элементов $\mathbf{x}_1 \in \mathbb{V}_1$ и $\mathbf{x}_2 \in \mathbb{V}_2$ ставится в соответствие их *тензорное произведение* $\mathbf{x}_1 \otimes \mathbf{x}_2 \in \mathbb{V}_1 \otimes \mathbb{V}_2$. $\square$

Тензорное произведение некоммукативно, т.к. порядок сомножителей важен и фиксирован.

Из определения следует, что тензорное произведение векторов *билинейно*, т.е. линейно по каждому из сомножителей при фиксированном другом:

$$\begin{aligned} \mathbf{x}_1 \otimes (\mathbf{x}_2 + \mathbf{y}_2) &= \mathbf{x}_1 \otimes \mathbf{x}_2 + \mathbf{x}_1 \otimes \mathbf{y}_2, \\ (\mathbf{x}_1 + \mathbf{y}_1) \otimes \mathbf{x}_2 &= \mathbf{x}_1 \otimes \mathbf{x}_2 + \mathbf{y}_1 \otimes \mathbf{x}_2, \\ (a\mathbf{x}_1) \otimes \mathbf{x}_2 &= a(\mathbf{x}_1 \otimes \mathbf{x}_2), \\ \mathbf{x}_1 \otimes (a\mathbf{x}_2) &= a(\mathbf{x}_1 \otimes \mathbf{x}_2), \end{aligned} \quad (1.29)$$

т.к. элементы $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ вида (1.28) отождествлены.

Пусть $\mathbb{V}_1$ и $\mathbb{V}_2$ – конечномерные векторные пространства с заданными базисами $\mathbf{e}_\alpha$, $\alpha = 1, \dots, \dim \mathbb{V}_1$ и $\mathbf{e}_i$, $i = 1, \dots, \dim \mathbb{V}_2$. Тогда их элементы имеют вид $\mathbf{x}_1 = x_1^\alpha \mathbf{e}_\alpha$ и $\mathbf{x}_2 = x_2^i \mathbf{e}_i$. Из билинейности (1.29) вытекает, что тензорное произведение произвольных векторов имеет вид

$$\mathbf{x}_1 \otimes \mathbf{x}_2 = x^{\alpha i} \mathbf{e}_\alpha \otimes \mathbf{e}_i,$$

где $x^{\alpha i} = x_1^\alpha x_2^i$. Это означает, что векторы $\{\mathbf{e}_\alpha \otimes \mathbf{e}_i\}$ образуют базис тензорного пространства $\mathbb{V}_1 \otimes \mathbb{V}_2$, и компонентами тензорного произведения векторов $\mathbf{x}_1 \otimes \mathbf{x}_2$ относительно этого базиса являются простые произведения компонент сомножителей $\{x_1^\alpha x_2^i\}$. Отсюда следует, что размерность тензорного произведения конечна и равна произведению размерностей векторных пространств:

$$\dim(\mathbb{V}_1 \otimes \mathbb{V}_2) = \dim \mathbb{V}_1 \times \dim \mathbb{V}_2.$$

Предложение 1.3.7. *Тензорное произведение дистрибутивно, т.е. для произвольных векторных пространств $\mathbb{V}_1$, $\mathbb{V}_2$ и $\mathbb{V}_3$ справедливы формулы:*

$$\begin{aligned} (\mathbb{V}_1 \oplus \mathbb{V}_2) \otimes \mathbb{V}_3 &= (\mathbb{V}_1 \otimes \mathbb{V}_3) \oplus (\mathbb{V}_2 \otimes \mathbb{V}_3), \\ \mathbb{V}_3 \otimes (\mathbb{V}_1 \oplus \mathbb{V}_2) &= (\mathbb{V}_3 \otimes \mathbb{V}_1) \oplus (\mathbb{V}_3 \otimes \mathbb{V}_2). \end{aligned}$$

Доказательство. Прямая проверка. $\square$

Заметим, что тензорное произведение различных пар векторов может совпадать.

Пример 1.3.51.

$$\mathbf{x}_1 \otimes \mathbf{x}_2 = (a\mathbf{x}_1) \otimes \left(\frac{\mathbf{x}_2}{a}\right), \quad \forall a \neq 0. \quad \square$$

Определение. Элементы тензорного произведения векторных пространств $\mathbb{V}_1 \otimes \mathbb{V}_2$ называются *тензорами*, а элементы вида $\mathbf{x}_1 \otimes \mathbf{x}_2$ – *разложимыми тензорами*. $\square$

Поскольку элементы вида $(\mathbf{x}_1, \mathbf{x}_2)$ составляют базис $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$, то разложимые тензоры $\mathbf{x}_1 \otimes \mathbf{x}_2$ порождают все тензорное произведение $\mathbb{V}_1 \otimes \mathbb{V}_2$. Однако они не являются базисом, т.к. между ними существует много линейных зависимостей.

Определение. Каждой паре элементов $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{V}_1 \times \mathbb{V}_2 \subset \mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ мы поставили в соответствие их тензорное произведение $\mathbf{x}_1 \otimes \mathbf{x}_2$, которое определяется естественной проекцией

$$\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2) \rightarrow \mathbb{V}_1 \otimes \mathbb{V}_2.$$

Это отображение называется *каноническим билинейным отображением* векторных пространств и обозначается следующим образом

$$\mathbb{V}_1 \times \mathbb{V}_2 \ni (\mathbf{x}_1, \mathbf{x}_2) \mapsto \varphi(\mathbf{x}_1, \mathbf{x}_2) := \mathbf{x}_1 \otimes \mathbf{x}_2 \in \mathbb{V}_1 \otimes \mathbb{V}_2. \quad \square$$

Тензорное произведение обладает следующим свойством универсальности. Пусть $\mathbb{W}$ – некоторое векторное пространство и ψ – билинейное (1.29) отображение векторных пространств $\mathbb{V}_1$ и $\mathbb{V}_2$:

$$\psi : \mathbb{V}_1 \times \mathbb{V}_2 \rightarrow \mathbb{W}.$$

Тогда для любых $\mathbb{W}$ и ψ существует единственное линейное отображение

$$\phi : \mathbb{V}_1 \otimes \mathbb{V}_2 \rightarrow \mathbb{W}$$

такое, что $\psi = \phi \circ \varphi$, где φ – каноническое билинейное отображение. Это означает коммутативность диаграммы

$$\begin{array}{ccc} \mathbb{V}_1 \times \mathbb{V}_2 & \xrightarrow{\varphi} & \mathbb{V}_1 \otimes \mathbb{V}_2 \\ & \searrow \psi & \downarrow \phi \\ & & \mathbb{W} \end{array}$$

Последовательно умножая векторные пространства можно получить тензорное произведение произвольного конечного числа сомножителей. При этом выполняется закон ассоциативности $(\mathbb{V}_1 \otimes \mathbb{V}_2) \otimes \mathbb{V}_3 = \mathbb{V}_1 \otimes (\mathbb{V}_2 \otimes \mathbb{V}_3)$ и можно просто писать $\mathbb{V}_1 \otimes \mathbb{V}_2 \otimes \mathbb{V}_3$.

Два вектора из одного векторного пространства $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ можно рассматривать как элементы двух одинаковых векторных пространств, $\mathbb{V}_1 = \mathbb{V}_2 = \mathbb{V}$, и построить их тензорное произведение $\mathbf{x} \otimes \mathbf{y}$. Оно, как было отмечено выше, некоммутативно,

$$\mathbf{x} \otimes \mathbf{y} \neq \mathbf{y} \otimes \mathbf{x},$$

поскольку в тензорном произведении важен порядок сомножителей.

Определение. Рассмотрим векторное пространство $\mathbb{V}$ и его тензорное произведение на себя

$$\mathbb{V}^r := \underbrace{\mathbb{V} \otimes \dots \otimes \mathbb{V}}_r.$$

Элемент $\mathbf{x} \in \mathbb{V}^r$ называется *тензором* ранга r . Прямая сумма всех тензорных пространств

$$\otimes \mathbb{V} := \bigoplus_{r=0}^{\infty} \mathbb{V}^r, \quad (1.30)$$

где $\mathbb{V}^0$ – поле вещественных чисел, с операциями сложения, умножения на числа и тензорным умножением называется *тензорной алгеброй*. $\square$

Пример 1.3.52. Рассмотрим поле вещественных чисел как одномерное векторное пространство с элементами $a\mathbf{1} \in \mathbb{R}$, где $\mathbf{1}$ – базисный вектор. Пусть также задано n -мерное векторное пространство над полем вещественных чисел $\mathbf{x} = x^\alpha \mathbf{e}_\alpha \in \mathbb{V}$. Тогда тензорное произведение векторов из этих пространств имеет вид

$$(a\mathbf{1}) \otimes (x^\alpha \mathbf{e}_\alpha) = ax^\alpha (\mathbf{1} \otimes \mathbf{e}_\alpha) = ax^\alpha \hat{\mathbf{e}}_\alpha,$$

где $\hat{\mathbf{e}}_\alpha := \mathbf{1} \otimes \mathbf{e}_\alpha$ – базисный вектор тензорного произведения. Это значит, что для любого векторного пространства над $\mathbb{R}$ существует канонический изоморфизм $\mathbb{R} \otimes \mathbb{V} \simeq \mathbb{V}$. Аналогично, $\mathbb{V} \otimes \mathbb{R} \simeq \mathbb{V}$. $\square$

В тензорной алгебре определены три операции: сложение, умножение на числа (элементы из $\mathbb{V}^0$), а также тензорное произведение. При этом сложение двух тензоров одного ранга понимается, как сложение в соответствующем векторном пространстве $\mathbb{V}^r$. В то же время для тензоров разного ранга используется прямая сумма.

Тензорная алгебра некоммутативна и ассоциативна. Если $\mathbf{x} \in \mathbb{V}^r$, то мы пишем $\deg \mathbf{x} = r$. Тензорная алгебра (1.30) бесконечномерна, т.к. ранг тензоров неограничен. По построению она обладает естественной $\mathbb{Z}$ -градуировкой. Гомоморфизм $\mathbb{Z} \rightarrow \mathbb{Z}_2$, делящий все целые числа на четные и нечетные, индуцирует $\mathbb{Z}_2$ -градуировку тензорной алгебры.

Напомним общее определение алгебры.

Определение. *Ассоциативной алгеброй* над полем вещественных чисел называется кольцо $\mathbb{A}$, которое является векторным пространством над полем вещественных чисел $\mathbb{R}$ и удовлетворяет условию

$$a(fg) = (af)g = f(ag), \quad \forall f, g \in \mathbb{A}, \quad \forall a \in \mathbb{R}.$$

Если в определении кольца исключить условие ассоциативности, то получим общее определение алгебры. Алгебра называется *коммутативной*, если $fg = gf$. *Размерностью* алгебры называется ее размерность как векторного пространства. $\square$

Таким образом в алгебре определено три операции: сложение, умножение и умножение на числа (скаляры). Конечно, поле вещественных чисел $\mathbb{R}$ в определении алгебры можно заменить на произвольное поле $\mathbb{K}$.

Отметим, что коммутативная алгебра может не быть ассоциативной.

Пример 1.3.53. Рассмотрим множество квадратных $n \times n$ -матриц, элементами которых являются вещественные $\mathbb{R}$ или комплексные $\mathbb{C}$ числа. В этих множествах определены обычные операции сложения и умножения матриц. Вместе с умножением матриц на скаляры, в роли которых выступают, соответственно, вещественные или комплексные числа, мы получаем ассоциативные некоммутативные (при $n \geq 2$) алгебры с единицей над полем $\mathbb{R}$ или $\mathbb{C}$. Эти алгебры обозначаются, соответственно, $\mathbf{mat}(n, \mathbb{R})$ и $\mathbf{mat}(n, \mathbb{C})$. Отметим, что умножение матриц не является групповой операцией в алгебре, т.к. среди матриц есть вырожденные (с нулевым определителем), которые не имеют обратных. Размерность алгебры $\mathbf{mat}(n, \mathbb{R})$ равна числу независимых вещественных параметров, однозначно определяющих соответствующие матрицы,

$$\dim \mathbf{mat}(n, \mathbb{R}) = n^2.$$

Комплексная размерность алгебры $\mathbf{mat}(n, \mathbb{C})$ также равна n^2 . Алгебру комплексных матриц $\mathbf{mat}(n, \mathbb{C})$ можно также рассматривать над полем вещественных чисел. В этом случае ее вещественная размерность равна $2n^2$. Подмножества обратимых матриц в алгебрах $\mathbf{mat}(n, \mathbb{R})$ и $\mathbf{mat}(n, \mathbb{C})$ представляют собой общие линейные группы $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{GL}(n, \mathbb{C})$. В них сохраняется только одна бинарная операция – умножение матриц. $\square$

Пример 1.3.54. Пусть $\mathbb{k} = \mathbb{R}, \mathbb{C}$ поле вещественных или комплексных чисел. Обозначим через $\mathbb{k}[x]$ множество многочленов от переменного x с коэффициентами из поля $\mathbb{k}$. Многочлены можно складывать, умножать на элементы поля и перемножать между собой. Тем самым они образуют алгебру. Эта алгебра ассоциативна, коммутативна и бесконечномерна. В ней также есть единица $1 \in \mathbb{k}$. Рассмотрим произвольный многочлен $f \in \mathbb{k}[x]$. Если уравнение $f = 0$ разрешимо в поле $\mathbb{k}$, т.е. $f(x) = 0$ для некоторого $x \in \mathbb{k}$, то поле $\mathbb{k}$ называется *алгебраически замкнутым*. В противном случае поле $\mathbb{k}$ алгебраически не замкнуто. Поле комплексных чисел $\mathbb{C}$ является алгебраически замкнутым, а поле вещественных чисел $\mathbb{R}$ – нет (например, уравнение $x^2 + 1 = 0$ не имеет вещественных решений). Мы говорим, что поле комплексных чисел является алгебраическим расширением поля вещественных чисел.

Каждый многочлен над полем комплексных чисел $f \in \mathbb{C}[z]$ с точностью до постоянного множителя задается своими нулями $z_1, \dots, z_r$ и их кратностями $k_1, \dots, k_r$. При этом сумма $k_1 + \dots + k_r$ равна степени многочлена. Рациональная функция $\varphi(z) := f(z)/g(z)$, где $f, g \in \mathbb{C}[z]$ – многочлены без общих делителей, опять же с точностью до постоянного множителя, определяется нулями многочленов f и g , где она обращается либо в нуль, либо в бесконечность.

Если имеется конечное число переменных $x^1, \dots, x^n$, то через $\mathbb{k}[x^1, \dots, x^n]$ обозначается алгебра многочленов от n переменных с коэффициентами из поля $\mathbb{k}$. Также употребляется сокращенная векторная запись $\mathbb{k}[x] := \mathbb{k}[x^1, \dots, x^n]$, где $x = (x^1, \dots, x^n)$. $\square$

Для любых двух элементов f, g ассоциативной алгебры $\mathbb{A}$ можно ввести их *коммутатор*

$$[f, g] := fg - gf.$$

Ассоциативность необходима для того, чтобы выполнялись *тождество Якоби*

$$[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0, \quad f, g, h \in \mathbb{A},$$

которое проверяется прямой проверкой. Это означает, что на любой ассоциативной алгебре можно ввести структуру алгебры Ли (см. раздел 2.6.7). Обратное утверждение не верно. Не всякую алгебру Ли можно рассматривать, как ассоциативную алгебру.

Пример 1.3.55. На алгебре Ли векторных полей нельзя определить структуру ассоциативной алгебры. $\square$

Напомним определение идеала.

Определение. Подмножество $\mathbb{I}$ кольца $\mathbb{K}$ называется *идеалом*, если выполнены два условия:

- 1) $\mathbb{I}$ есть подгруппа $\mathbb{K}$ по отношению к сложению;
- 2) $\mathbb{I}$ содержит все произведения ab (*левый идеал*), или все произведения ba (*правый идеал*), или все произведения ab и ba (*двусторонний идеал*), где a – любой элемент кольца $\mathbb{K}$ и b – любой элемент идеала $\mathbb{I}$. $\square$

Роль двусторонних идеалов в алгебрах аналогична роли нормальных подгрупп в теории групп. В частности, можно проверить, что фактор пространство $\mathbb{A}/\mathbb{I}$ является алгеброй.

Определение. Алгебра $\mathbb{A}$ называется *фильтрованной*, если в ней выделены подпространства $\mathbb{A}_k$, индексированные элементами линейно упорядоченной группы $\mathbb{G}$ таким образом, что $\mathbb{A}_k \subseteq \mathbb{A}_l$ при $k < l$ и $\mathbb{A}_k \mathbb{A}_l \subseteq \mathbb{A}_{k+l}$.

С каждой фильтрацией данной алгебры ассоциируется *градуированная алгебра*

$$\mathrm{gr} \mathbb{A} = \bigoplus_k \bar{\mathbb{A}}_k,$$

где

$$\bar{\mathbb{A}}_k := \frac{\mathbb{A}_k}{\sum_{l < k} \mathbb{A}_l},$$

а произведение элементов $\bar{f} \in \bar{\mathbb{A}}_k$ и $\bar{g} \in \bar{\mathbb{A}}_l$ определяется по формуле $\bar{f}\bar{g} := \overline{fg}$, где f и g – представители смежных классов $\bar{f}$ и $\bar{g}$, а $\overline{fg} \in \bar{\mathbb{A}}_{k+l}$ – смежный класс, порожденный элементом $fg \in \mathbb{A}_{k+l}$. $\square$

Замечание. Чаще всего в качестве линейно упорядоченной группы выступает группа целых чисел по сложению $\mathbb{Z}$. В этом случае $\bar{\mathbb{A}}_k = \mathbb{A}_k / \mathbb{A}_{k-1}$. $\square$

Если в алгебре $\mathbb{A}$ выполняется какое-либо полилинейное тождество (например, коммутативность, ассоциативность или тождество Якоби), то в градуированной алгебре $\mathrm{gr} \mathbb{A}$ также выполняется это тождество.

Пример 1.3.56. Ассоциированной градуированной алгеброй будет внешняя алгебра $\Lambda(\mathbb{V})$, которая описана в разделе 3.1. $\square$

Пример 1.3.57. Рассмотрим алгебру Клиффорда $\mathrm{cl}(\mathbb{V}, Q)$ (см. главу 26) с образующими e^a , $a = 1, \dots, n$. Обозначим через cl_k , $k \in \mathbb{Z}$, множество элементов алгебры Клиффорда, представимых в виде многочленов степени $\leq k$ от образующих. Множества cl_k при $k < 0$ являются пустыми. Тогда подмножества $\mathrm{cl}_k \subset \mathrm{cl}(\mathbb{V}, Q)$ задают фильтрацию алгебры Клиффорда. $\square$

Нормированные векторные пространства

Векторное евклидово пространство $\mathbb{R}^n$ можно снабдить естественной топологической метрикой. Определим расстояние между двумя векторами $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ с началом в начале системы координат по следующей формуле

$$l(\mathbf{x}, \mathbf{y}) := \sqrt{(x^1 - y^1)^2 + \dots + (x^n - y^n)^2}, \quad (1.31)$$

что совпадает с евклидовым расстоянием (1.6) между точками в $\mathbb{R}^n$, соответствующими концам векторов. При этом все свойства метрики, очевидно, выполнены. Таким образом векторное пространство $\mathbb{R}^n$ становится метрическим пространством, и, следовательно, на нем можно определить естественную топологию. Можно доказать, что любая топология в евклидовом пространстве, в которой операции сложения и умножения на числа непрерывны, является естественной [6].

Метрика (1.31) обладает двумя важными свойствами:

- 1) $l(\mathbf{x}, \mathbf{y}) = l(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z})$ – инвариантность относительно сдвига на вектор $\mathbf{z} \in \mathbb{R}^n$;
 - 2) $l(a\mathbf{x}, a\mathbf{y}) = |a|l(\mathbf{x}, \mathbf{y})$, – умножение на скаляр увеличивает расстояние в $|a|$ раз.
- (1.32)

Заметим, что не каждая метрика обладает такими свойствами.

Если в векторном пространстве определена метрика, то определено отображение

$$\mathbb{V} \ni \mathbf{x} \mapsto \|\mathbf{x}\| := l(\mathbf{x}, \mathbf{0}) \in \mathbb{R}. \quad (1.33)$$

Если при этом метрика обладает свойствами (1.32), то построенное отображение имеет следующие свойства:

- 1) $\|\mathbf{0}\| = 0$, $\|\mathbf{x}\| > 0$, $\forall \mathbf{x} \neq \mathbf{0}$ – положительная определенность;
 - 2) $\|\mathbf{x}\| = \|-\mathbf{x}\|$ – четность;
 - 3) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ – неравенство треугольника;
 - 4) $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$ – умножение на скаляры.
- (1.34)

Свойство 2) является следствием 4) при $a = -1$. Оно выделено, потому что является также следствием инвариантности метрики относительно сдвига и симметрии относительно перестановки аргументов.

Определение. Векторное пространство $\mathbb{V}$ называется *нормированным*, если оно снабжено нормой, т.е. отображением

$$\mathbb{V} \ni \mathbf{x} \mapsto \|\mathbf{x}\| \in \mathbb{R}, \quad (1.35)$$

удовлетворяющим условиям (1.34). При этом число $\|\mathbf{x}\|$ называется *нормой* или *длиной* вектора $\mathbf{x} \in \mathbb{V}$. Полное (см. раздел 1.3.1) нормированное векторное пространство называется *банаховым*. □

Если в векторном пространстве $\mathbb{V}$ задана норма, то она порождает следующую метрику $l(\mathbf{x}, \mathbf{y}) := \|\mathbf{y} - \mathbf{x}\|$ и, следовательно, некоторую топологию. Нетрудно проверить, что все аксиомы метрики при этом выполнены. Конечно, не каждая метрика порождается некоторой нормой.

Наличие нормы в векторном пространстве позволяет определить сходимость последовательностей, используя соответствующую метрику. Если на некотором линейном пространстве задано две нормы, то они называются *эквивалентными*, если сходимость последовательности по любой из них влечет за собой сходимость по другой норме.

Теорема 1.3.16. *Все нормы в конечномерных вещественных или комплексных векторных пространствах эквивалентны.*

Доказательство. См., например, [23], глава 5, следствие 5.4.6. $\square$

Пример 1.3.58. В евклидовом пространстве наиболее часто используются следующие нормы:

$$\begin{aligned} \|\mathbf{x}\|_1 &:= |x^1| + |x^2| + \dots + |x^n|, & L_1 - \text{норма}, \\ \|\mathbf{x}\|_2 &:= \sqrt{|x^1|^2 + |x^2|^2 + \dots + |x^n|^2}, & L_2 - \text{норма}, \\ \|\mathbf{x}\|_p &:= \left(\sum_{\alpha=1}^n |x^\alpha|^p \right)^{1/p}, \quad p \geq 1 & L_p - \text{норма}, \\ \|\mathbf{x}\|_\infty &:= \max \{|x^1|, |x^2|, \dots, |x^n|\}, & L_\infty - \text{норма}. \end{aligned} \quad (1.36)$$

Все выражения, как нетрудно проверить, удовлетворяют свойствам 1) – 4) норм. L_1 -норма не порождается никаким скалярным произведением. Обычная длина вектора в евклидовом пространстве есть ни что иное, как L_2 -норма, которая порождается евклидовым скалярным произведением. L_p -норма называется также нормой *Гёльдера* с показателем $p \geq 1$. Она порождается скалярным произведением

$$l(\mathbf{x}, \mathbf{y}) := \left(\sum_{\alpha=1}^n |x^\alpha - y^\alpha|^p \right)^{1/p}.$$

При $0 < p < 1$ выражение $\|\mathbf{x}\|_p$ удовлетворяет всем свойствам нормы кроме неравенства треугольника. L_∞ -норма является пределом L_p -нормы. А именно,

$$\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Все приведенные выше нормы прямо обобщаются на комплексное пространство $\mathbb{C}^n$. Пространства $\mathbb{R}^n$ и $\mathbb{C}^n$ с любой из приведенных выше норм являются банаховыми. $\square$

Рассмотренный пример показывает, что в одном и том же векторном пространстве можно задавать различные нормы и скалярные произведения. Выбор той или иной нормы зависит от конкретной задачи.

Пример 1.3.59. Понятие нормы можно ввести и на бесконечномерном векторном пространстве. На множестве $\mathcal{C}[a, b]$ всех непрерывных функций $f(x)$ на отрезке $[a, b]$, которое является бесконечномерным векторным пространством, обычно рассматривают следующие нормы:

$$\begin{aligned} \|f\|_1 &= \int_a^b dx |f(x)|, & L_1 - \text{норма}, \\ \|f\|_2 &= \left[\int_a^b dx |f(x)|^2 \right]^{1/2}, & L_2 - \text{норма}, \\ \|f\|_p &= \left[\int_a^b dx |f(x)|^p \right]^{1/p}, \quad p \geq 1 & L_p - \text{норма}, \\ \|f\|_\infty &= \max \{|f(x)| : x \in [a, b]\}, & L_\infty - \text{норма}. \end{aligned} \quad (1.37)$$

При этом, как и в предыдущем примере, выполнено равенство

$$\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p, \quad \forall f \in \mathcal{C}[a, b]. \quad \square$$

Определение. Метрику (1.31) (а также любую другую метрику) можно использовать для определения базы топологии векторного пространства. В результате будет построено *топологическое векторное пространство*. Топологическое векторное пространство $\mathbb{V}$ называется *метризуемым*, если его топология может быть индуцирована трансляционно инвариантной метрикой $l(\mathbf{x}, \mathbf{y}) = l(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z})$, $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{V}$. В общем случае топологическое векторное пространство, которое является метризуемым и полным, называется *пространством Фреше*. $\square$

Пример 1.3.60. Евклидово пространство $\mathbb{R}^n$, рассматриваемое как векторное пространство с метрикой (1.31), является пространством Фреше. $\square$

Определение. Пусть $f : \mathbb{V} \rightarrow \mathbb{W}$ – отображение, возможно, нелинейное, двух нормированных векторных пространств, и пусть $\mathbf{x} \in \mathbb{V}$ – некоторая точка. Тогда, если отображение f вблизи точки $\mathbf{x}$ можно представить в виде

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + f'_x(\mathbf{h}) + o(\mathbf{h}),$$

где f'_x – линейное отображение $\mathbb{V} \rightarrow \mathbb{W}$ и $o(\mathbf{h}) \in \mathbb{W}$ – некоторый вектор, такой, что

$$\lim_{\|\mathbf{h}\| \rightarrow 0} \frac{\|o(\mathbf{h})\|}{\|\mathbf{h}\|} = 0, \quad (1.38)$$

то отображение $f'_x : \mathbb{V} \rightarrow \mathbb{W}$ называется *производной по Фреше* отображения f в точке $\mathbf{x}$. $\square$

В частности, если $\mathbb{W} = \mathbb{R}$ – вещественная прямая, то мы имеем производную по Фреше от функционала на нормированном векторном пространстве.

Нетрудно показать, что если производная по Фреше существует, то она единственна.

Пример 1.3.61. Пусть в векторных пространствах $\mathbb{V}$ и $\mathbb{W}$ заданы базисы $\mathbf{e}_\alpha$, $\alpha = 1, \dots, n$, и $\mathbf{e}_i$, $i = 1, \dots, m$, соответственно. Тогда отображение f задается m функциями $f^i(\mathbf{x})$ от n переменных, которые в общем случае могут быть нелинейными. Если функции f^i дифференцируемы по всем аргументам, то $n \times m$ -матрица Якоби, составленная из частных производных $\{\partial_\alpha f^i|_x\}$ является производной по Фреше отображения f в точке $\mathbf{x} \in \mathbb{V}$. $\square$

Для производной по Фреше справедлива теорема о дифференцировании сложной функции.

Если отображение f двух банаховых пространств непрерывно дифференцируемо по Фреше в некоторой окрестности точки $\mathbf{x}$ и в этой точке производная Фреше f'_x является гомеоморфизмом, то в окрестности данной точки существует обратное отображение, которое также является гомеоморфизмом.

Наряду с производной Фреше отображений двух линейных пространств используется также производная Гато, для которой необходимо только наличие топологии.

Определение. Пусть $f : \mathbb{V} \rightarrow \mathbb{W}$ – отображение двух топологических векторных пространств, и пусть $\mathbf{x} \in \mathbb{V}$ некоторая точка. Если отображение f двух линейных пространств вблизи точки $\mathbf{x}$ можно представить в виде

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + f'_{G,\mathbf{x}}(\mathbf{h}) + o(\mathbf{h}),$$

где $f'_{G,x}(\mathbf{h})$ – линейное отображение $\mathbb{V} \rightarrow \mathbb{W}$ и

$$\lim_{t \rightarrow 0} \frac{o(t\mathbf{h})}{t} = \mathbf{0} \in \mathbb{W}, \quad t \in \mathbb{R}, \quad (1.39)$$

где сходимость определяется топологией пространства $\mathbb{W}$, то линейное отображение $f'_{G,x} : \mathbb{V} \rightarrow \mathbb{W}$ называется *производной по Гато* отображения f в точке $\mathbf{x}$. $\square$

В частности, если $\mathbb{W} = \mathbb{R}$ – вещественная прямая, то мы имеем производную по Гато от функционала на топологическом векторном пространстве.

Отличие производной Гато от производной Фреше сводится только к определению сходимости остаточных членов (1.38) и (1.39).

Поскольку каждое нормированное векторное пространство является топологическим с естественной топологией, то производная по Фреше является одновременно и производной по Гато. Обратное утверждение в общем случае неверно, т.к. наличие топологии в векторном пространстве совсем не означает наличие нормы. По этой причине производную Фреше называют также *сильной производной*, а производную Гато – *слабой*.

Для производной Гато теорема о дифференцируемости сложной функции, вообще говоря, не верна.

Скалярное произведение в векторном пространстве

Определение. Скалярным произведением $(\mathbf{x}, \mathbf{y})$ двух элементов векторного пространства $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ называется билинейное отображение

$$\mathbb{V} \times \mathbb{V} \ni \mathbf{x}, \mathbf{y} \mapsto (\mathbf{x}, \mathbf{y}) \in \mathbb{R},$$

которое обладает следующими свойствами:

- 1) $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$ – симметричность,
- 2) $(\mathbf{x}, \mathbf{x}) > 0, \forall \mathbf{x} \neq \mathbf{0}; \quad (\mathbf{0}, \mathbf{0}) = 0$ – положительная определенность.

Если скалярное произведение двух векторов равно нулю $(\mathbf{x}, \mathbf{y}) = 0$, то они называются *ортogonalными*. $\square$

В компонентах скалярное произведение задается положительно определенной квадратичной формой

$$(\mathbf{x}, \mathbf{y}) = x^\alpha y^\beta g_{\alpha\beta}, \quad (1.40)$$

где $g_{\alpha\beta} := (\mathbf{e}_\alpha, \mathbf{e}_\beta)$ – произвольная симметричная положительно определенная матрица.

Пример 1.3.62. В евклидовом пространстве $\mathbb{R}^n$ естественное скалярное произведение векторов задается в декартовых координатах с помощью евклидовой метрики (1.8) Это скалярное произведение симметрично и положительно определено. При этом норму (длину) вектора можно записать в виде $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$. $\square$

Отметим, что скалярное произведение различных векторов $(\mathbf{x}, \mathbf{y})$ может быть отрицательно. Нулевой вектор ортогонален самому себе и всем другим векторам.

В современной математической физике скалярное произведение часто понимается более широко. А именно, отбрасывается требование положительной определенности. В дальнейшем, как правило, под скалярным произведением векторов мы будем понимать просто симметричную билинейную форму.

Предложение 1.3.8. Пусть $\mathbb{V}$ – произвольное векторное пространство с положительно определенным скалярным произведением (1.40). Тогда на $\mathbb{V}$ можно задать норму или длину вектора

$$\|\mathbf{x}\| := \sqrt{(\mathbf{x}, \mathbf{x})}. \quad (1.41)$$

Доказательство. Из линейности и положительной определенности скалярного произведения немедленно следуют свойства 1), 2) и 4) в определении нормы. Для доказательства неравенства треугольника воспользуемся неравенством Коши–Буняковского (1.42), которое будет доказано чуть позже:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2|(\mathbf{x}, \mathbf{y})| \leq \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \quad \square \end{aligned}$$

Обсудим некоторые свойства векторных пространств с положительно определенным скалярным произведением. Следующие два утверждения проверяются прямой проверкой.

Предложение 1.3.9 (Правило параллелограмма). Пусть $\mathbb{V}$ – векторное пространство с положительно определенным скалярным произведением. Тогда справедлива формула

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2,$$

где $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ – произвольные векторы.

Предложение 1.3.10 (Теорема Пифагора). Пусть $\mathbb{V}$ – векторное пространство с положительно определенным скалярным произведением. Тогда справедлива формула

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2,$$

где $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ – произвольные ортогональные векторы, $(\mathbf{x}, \mathbf{y}) = 0$.

Допустим, что в векторном пространстве $\mathbb{V}$, $\dim \mathbb{V} = n$, задан произвольный набор взаимно ортогональных векторов $\mathbf{x}_i$, $i = 1, \dots, N \leq n$:

$$(\mathbf{x}_i, \mathbf{x}_j) = 0, \quad i \neq j.$$

Тогда из теоремы Пифагора по индукции следует равенство

$$\left\| \sum_{i=1}^N \mathbf{x}_i \right\|^2 = \sum_{i=1}^N \|\mathbf{x}_i\|^2.$$

В произвольном векторном пространстве $\mathbb{V}$, $\dim \mathbb{V} = n$, со скалярным произведением можно выбрать *ортонормальный базис* $\mathbf{e}_\alpha$, $\alpha = 1, \dots, n$, состоящий из взаимно ортогональных векторов единичной длины:

$$(\mathbf{e}_\alpha, \mathbf{e}_\beta) = \delta_{\alpha\beta},$$

где $\delta_{\alpha\beta}$ – евклидова метрика (1.8). Произвольный вектор $\mathbf{x} \in \mathbb{V}$ можно разложить по ортонормальному базису

$$\mathbf{x} = x^\alpha \mathbf{e}_\alpha,$$

где $x^\alpha := \delta^{\alpha\beta}(\mathbf{x}, \mathbf{e}_\beta)$ – компоненты данного вектора в ортонормальном базисе.

Предложение 1.3.11 (Неравенство Бесселя). Пусть $\mathbf{x}_i \in \mathbb{V}$, $i = 1, \dots, N \leq n$, – произвольный набор ортонормальных векторов. Тогда справедливо неравенство

$$\sum_{i=1}^N |(\mathbf{x}, \mathbf{x}_i)|^2 \leq \|\mathbf{x}\|^2,$$

где $\mathbf{x} \in \mathbb{V}$ – произвольный вектор.

Доказательство. Из неравенства

$$0 \leq \left\| \mathbf{x} - \sum_{i=1}^N (\mathbf{x}, \mathbf{x}_i) \mathbf{x}_i \right\|^2 = \|\mathbf{x}\|^2 - 2 \sum_{i=1}^N (\mathbf{x}, \mathbf{x}_i)^2 + \sum_{i=1}^N \sum_{j=1}^N (\mathbf{x}, \mathbf{x}_i)(\mathbf{x}, \mathbf{x}_j)(\mathbf{x}_i, \mathbf{x}_j).$$

вытекает неравенство Бесселя, т.к. векторы $\mathbf{x}_i$ ортонормальны, $(\mathbf{x}_i, \mathbf{x}_j) = \delta_{ij}$. $\square$

Знак равенства в неравенстве Бесселя достигается тогда и только тогда, когда набор векторов $\mathbf{x}_i$ образует ортонормальный базис в $\mathbb{V}$, т.е. $N = n$.

Предложение 1.3.12. Пусть $\mathbf{e}_\alpha \in \mathbb{V}$, $\alpha = 1, \dots, n$, – ортонормальный базис векторного пространства. Тогда справедливо равенство

$$\sum_{\alpha=1}^n |(\mathbf{x}, \mathbf{e}_\alpha)|^2 = \|\mathbf{x}\|^2,$$

где $\mathbf{x} \in \mathbb{V}$ – произвольный вектор. То есть квадрат длины произвольного вектора равен сумме квадратов компонент относительно ортонормального базиса.

Доказательство. В доказательстве неравенства Бесселя надо заменить знак неравенства на равенство. $\square$

Предложение 1.3.13 (Неравенство Коши–Буняковского). Пусть $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ – произвольные векторы. Тогда справедливо неравенство

$$|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\| \|\mathbf{y}\|. \quad (1.42)$$

Неравенство Коши–Буняковского называют также *неравенством Шварца*.

Доказательство. Если $\mathbf{y} = \mathbf{0}$, то неравенство Коши–Буняковского (1.42) выполнено. Допустим, что $\mathbf{y} \neq \mathbf{0}$. В этом случае можно построить единичный вектор $\mathbf{y}_0 := \mathbf{y} / \|\mathbf{y}\|$, $\|\mathbf{y}_0\| = 1$. Тогда из неравенства Бесселя следует неравенство

$$|(\mathbf{x}, \mathbf{y}_0)| \leq \|\mathbf{x}\|.$$

Поскольку

$$|(\mathbf{x}, \mathbf{y}_0)| = \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{y}\|},$$

то отсюда вытекает неравенство Коши–Буняковского. $\square$

Замечание. В моделях математической физики, например, при рассмотрении лоренцевых многообразий, под скалярным произведением понимают произвольное симметричное билинейное отображение, отбрасывая требование положительной определенности. В этом случае два вектора также называются ортогональными, если их скалярное произведение равно нулю. При этом отличный от нуля вектор $\mathbf{x} \neq \mathbf{0}$ может оказаться ортогональным самому себе, если $(\mathbf{x}, \mathbf{x}) = 0$. Такой вектор называется *нулевым*. Если метрика имеет лоренцеву сигнатуру $\text{sign } g_{\alpha\beta} = (+ - \dots -)$, то нулевой вектор называется также *светоподобным*.

Из симметрии матрицы $g_{\alpha\beta} := (\mathbf{e}_\alpha, \mathbf{e}_\beta)$ следует, что все ее собственные числа вещественны. Для того, чтобы матрица была положительно определена необходимо и достаточно, чтобы все ее собственные числа были вещественны и положительны.

Если в векторном пространстве $\mathbb{V}$ задано скалярное произведение, то в нем всегда можно выбрать базис $\mathbf{e}_\alpha$, состоящий из взаимно ортогональных векторов. Эти векторы будут собственными векторами матрицы $g_{\alpha\beta}$:

$$\sum_{\beta=1}^n g_{\alpha\beta} \mathbf{e}_\beta = \lambda_\alpha \mathbf{e}_\alpha,$$

где проводится суммирование по индексу β , но отсутствует суммирование по α , и $\lambda_\alpha > 0$ – собственные значения матрицы $g_{\alpha\beta}$. В общем случае часть собственных значений может совпадать. Для метрики евклидова пространства $\mathbb{R}^n$ в декартовых координатах все собственные значения совпадают и равны единице.

Рассмотрим два векторных пространства $\mathbb{V}_1$ и $\mathbb{V}_2$ с элементами $\mathbf{x} = x^\alpha \mathbf{e}_\alpha \in \mathbb{V}_1$, $\alpha = 1, \dots, m$ и $\mathbf{y} = y^i \mathbf{e}_i \in \mathbb{V}_2$, $i = 1, \dots, n$. Выберем в каждом пространстве базис, состоящий из ортогональных векторов. Пусть в них заданы скалярные произведения матрицами $g_{\alpha\beta}$ и g_{ij} , соответственно. Тогда в тензорном произведении $\mathbb{V}_1 \otimes \mathbb{V}_2$ можно определить скалярное произведение. Действительно, пусть задано два вектора

$$X = X^{\alpha i} \mathbf{e}_\alpha \otimes \mathbf{e}_i, \quad Y = Y^{\alpha a} \mathbf{e}_\alpha \otimes \mathbf{e}_i.$$

Тогда билинейное отображение

$$(X, Y) := X^{\alpha i} Y^{\beta j} g_{\alpha\beta} g_{ij}$$

симметрично и положительно определено. Последнее утверждение следует из того, что векторы базиса $\mathbf{e}_\alpha \otimes \mathbf{e}_i$ являются собственными векторами матрицы $g_{(\alpha i)(\beta j)} := g_{\alpha\beta} g_{ij}$, определяющей скалярное произведение,

$$\sum_{\beta, j} g_{\alpha\beta} g_{ij} \mathbf{e}_\beta \otimes \mathbf{e}_j = \lambda_\alpha \lambda_i \mathbf{e}_\alpha \otimes \mathbf{e}_i,$$

с положительными собственными числами $\lambda_\alpha \lambda_i > 0$.

Как уже отмечалось, векторные пространства можно рассматривать над полем комплексных чисел $\mathbb{C}$. В этом случае говорят о комплексном векторном пространстве, которое обозначим $\mathbb{H}$.

Определение. Квадратичная форма

$$\mathbb{H} \times \mathbb{H} \ni \quad \mathbf{x}, \mathbf{y} \mapsto (\mathbf{x}, \mathbf{y}) \quad \in \mathbb{C}, \quad (1.43)$$

которая линейна по первому аргументу,

$$(a\mathbf{x} + b\mathbf{y}, \mathbf{z}) = a(\mathbf{x}, \mathbf{z}) + b(\mathbf{y}, \mathbf{z}), \quad a, b \in \mathbb{C},$$

и удовлетворяет равенству

$$(\mathbf{x}, \mathbf{y})^\dagger = (\mathbf{y}, \mathbf{x}), \quad (1.44)$$

где символ $\dagger$ обозначает комплексное сопряжение называется *эрмитовой*. Каждая эрмитова квадратичная форма определяет в комплексном векторном пространстве

скалярное произведение. Если скалярное произведение положительно определено, т.е. $(\mathbf{x}, \mathbf{x}) \geq 0$ для всех $\mathbf{x} \in \mathbb{H}$, то оно задает *норму* в $\mathbb{H}$:

$$\|\mathbf{x}\| := \sqrt{(\mathbf{x}, \mathbf{x})}.$$

Эта норма называется *строго положительной*, если из условия $\|\mathbf{x}\| = 0$ следует $\mathbf{x} = 0$. Конечномерное комплексное векторное пространство $\mathbb{H}$ с положительно определенной эрмитовой квадратичной формой называется *эрмитовым пространством*. Полное комплексное векторное пространство со строго положительной нормой называется *гильбертовым*. $\square$

Вещественным аналогом эрмитова пространства является евклидово пространство со скалярным произведением, которое задается положительно определенной симметричной билинейной формой.

Если в комплексном векторном пространстве $\mathbb{H}$, $\dim_{\mathbb{C}} \mathbb{H} = n$, задан базис $\mathbf{e}_{\alpha}$, $\alpha = 1, \dots, n$, то скалярное произведение в компонентах задается суммой

$$(\mathbf{x}, \mathbf{y}) = x^{\alpha} y^{\dagger\beta} \eta_{\alpha\beta},$$

где $\eta_{\alpha\beta} := (\mathbf{e}_{\alpha}, \mathbf{e}_{\beta})$ – некоторая эрмитова матрица. В эрмитовом пространстве $\mathbb{H}$ всегда можно выбрать такой базис, что эрмитова квадратичная форма примет канонический (диагональный) вид. Напомним, что у эрмитовой матрицы все собственные числа вещественны. Обозначим их через λ_{α} . При этом часть собственных значений может быть равна нулю. Тогда справедливо

Предложение 1.3.14. *В n -мерном комплексном векторном пространстве $\mathbb{H}$ с эрмитовым скалярным произведением существует такое унитарное преобразование базиса, что в новом базисе квадратичная форма принимает диагональный вид:*

$$(\mathbf{x}, \mathbf{y}) = \sum_{\alpha=1}^n \lambda_{\alpha} x^{\alpha} y^{\dagger\alpha}.$$

Доказательство. См., например, [11], часть II, глава 3, §3, теорема 7. $\square$

Базис, в котором квадратичная форма диагональна, является ортогональным. Конечно, этот базис определен неоднозначно. Дальнейшей растяжкой координат можно добиться, чтобы на диагонали стояли $+1$, -1 или 0 . В частности, симметричную билинейную квадратичную форму в вещественном векторном пространстве всегда можно привести к диагональному виду с помощью ортогонального преобразования.

Если в комплексном векторном пространстве задано сразу две эрмитовы квадратичные формы, то иногда их можно привести к канонической (диагональной) форме одновременно.

Предложение 1.3.15. *Пусть в n -мерном комплексном (или вещественном) векторном пространстве $\mathbb{H}$ задано две эрмитовы (вещественные симметричные) квадратичные формы, причем одна из них положительно определена. Тогда в $\mathbb{H}$ существует такой ортогональный базис, в котором обе квадратичные формы будут иметь канонический (диагональный) вид.*

Доказательство. См., например, [11], часть II, глава 3, §3, теорема 8. $\square$

Гильбертовы пространства лежат в основе квантовой механики, где каждое состояние физической системы отождествляется с вектором соответствующего гильбертова пространства. Гильбертовы пространства в квантовой механике обычно бесконечномерны.

Пример 1.3.63. В нерелятивистской квантовой механике в качестве гильбертова пространства, как правило, рассматривается пространство $\mathbb{H} = \mathcal{L}_2(\mathbb{R}^n)$ достаточно гладких квадратично интегрируемых функций в евклидовом пространстве

$$\psi : \mathbb{R}^n \rightarrow \mathbb{C}.$$

Скалярное произведение в этом пространстве определяется интегралом

$$\mathbb{H} \times \mathbb{H} \ni \psi_1, \psi_2 \mapsto (\psi_1, \psi_2) := \int_{\mathbb{R}^n} dx \psi_1 \psi_2^\dagger \in \mathbb{C}.$$

Можно доказать, что это пространство является полным. $\square$

В комплексном векторном пространстве также рассматриваются обычные симметричные квадратичные формы, для которых условие (1.44) заменяется равенством

$$(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$$

без комплексного сопряжения. Таким симметричные формы будут рассмотрены, например, в главе 26 при рассмотрении алгебр Клиффорда.

Сопряженные пространства

Определение. *Линейным функционалом* или *линейной формой* $\mathbf{f}$ на вещественном векторном пространстве $\mathbb{V}$ называется линейное отображение

$$\mathbb{V} \ni \mathbf{x} \mapsto \mathbf{f}(\mathbf{x}) \in \mathbb{R},$$

обладающее свойствами:

$$\mathbf{f}(\mathbf{x} + \mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{y}), \quad \mathbf{f}(a\mathbf{x}) = a\mathbf{f}(\mathbf{x}), \quad a \in \mathbb{R}.$$

Множество $\mathbb{V}^*$ всех линейных функционалов на $\mathbb{V}$ снабжено естественной структурой векторного пространства над полем вещественных чисел относительно сложения и умножения на числа:

$$\begin{aligned} (\mathbf{f}_1 + \mathbf{f}_2)(\mathbf{x}) &:= \mathbf{f}_1(\mathbf{x}) + \mathbf{f}_2(\mathbf{x}), \\ (a\mathbf{f})(\mathbf{x}) &:= a\mathbf{f}(\mathbf{x}). \end{aligned}$$

Векторное пространство $\mathbb{V}^*$ называется *сопряженным*, *дуальным* или *двойственным* пространством к $\mathbb{V}$. $\square$

Можно показать, что сопряженное пространство имеет ту же размерность, что и векторное пространство $\mathbb{V}$. Для *конечномерных* векторных пространств взятие сопряжения дважды приводит к векторному пространству, которое изоморфно исходному, $\mathbb{V}^{**} \simeq \mathbb{V}$.

Определение. Пусть $\mathbb{U} \subset \mathbb{V}$, тогда *ортogonalным дополнением* или *аннулятором* множества $\mathbb{U}$ (не обязательно подпространства) называется подмножество

$$\mathbb{U}^\perp := \{\mathbf{f} \in \mathbb{V}^* : \mathbf{f}(\mathbf{x}) = 0, \quad \forall \mathbf{x} \in \mathbb{U}\}. \quad \square \quad (1.45)$$

Предложение 1.3.16. Если $\mathbb{U}$ – линейное подпространство в $\mathbb{V}$ размерности m , $1 \leq m \leq n$, то его ортогональное дополнение $\mathbb{U}^\perp$ является линейным подпространством в $\mathbb{V}^*$ и имеет размерность $n - m$.

Доказательство. Выберем в подпространстве $\mathbb{U}$ m линейно независимых векторов $\mathbf{x}_i$, $i = 1, \dots, m$. Тогда система линейных уравнений $\mathbf{f}(\mathbf{x}_i) = 0$ будет иметь $n - m$ линейно независимых решений, которые можно выбрать в качестве базиса в $\mathbb{U}^\perp$. $\square$

Аналогично определяется ортогональное дополнение некоторого подмножества $\mathbb{U}^* \subset \mathbb{V}^*$:

$$\mathbb{U}^{*\perp} := \{\mathbf{x} \in \mathbb{V} : \mathbf{f}(\mathbf{x}) = 0, \quad \forall \mathbf{f} \in \mathbb{U}^*\}. \quad (1.46)$$

Это означает, что операцию ортогонального дополнения можно применить дважды. Тогда для конечномерных векторных пространств справедливо равенство

$$(\mathbb{U}^\perp)^\perp = \mathbb{U}.$$

Если $\mathbb{V}_1, \mathbb{V}_2 \subset \mathbb{V}$ – два подпространства векторного пространства, то верна формула

$$\begin{aligned} (\mathbb{V}_1 \cap \mathbb{V}_2)^\perp &= \mathbb{V}_1^\perp \oplus [\mathbb{V}_2^\perp \setminus (\mathbb{V}_1^\perp \cap \mathbb{V}_2^\perp)] = \\ &= [\mathbb{V}_1^\perp \setminus (\mathbb{V}_1^\perp \cap \mathbb{V}_2^\perp)] \oplus \mathbb{V}_2^\perp. \end{aligned} \quad (1.47)$$

Здесь из одного из слагаемых вычтено пересечение ортогональных дополнений, чтобы не учитывать его дважды.

Определение. Гиперплоскостью $\mathbb{S}$ в векторном пространстве $\mathbb{V}$, ортогональной некоторому фиксированному элементу из сопряженного пространства $\mathbf{f} \in \mathbb{V}^*$, называется множество точек

$$\mathbb{S} := \{\mathbf{x} \in \mathbb{V} : \mathbf{f}(\mathbf{x}) = a, \quad \mathbf{f} \neq 0, \mathbf{f} \in \mathbb{V}^*, \quad a \in \mathbb{R}\}. \quad \square \quad (1.48)$$

Гиперплоскости являются подпространствами размерности $n - 1$ и определены для всех ненулевых линейных функционалов.

Пример 1.3.64. В евклидовом пространстве $\mathbb{R}^n$ скалярное произведение векторов при одном фиксированном векторе задает линейный функционал. Верно и обратное утверждение: любой линейный функционал можно задать, как скалярное произведение с некоторым вектором. Это означает, что дуальное пространство $\mathbb{R}^{n*}$ изоморфно самому евклидову пространству,

$$\mathbb{R}^{n*} \simeq \mathbb{R}^n.$$

Поэтому понятия ортогонального дополнения и гиперплоскости, ортогональной заданному вектору, в евклидовом пространстве имеют наглядный геометрический смысл. $\square$

Если $\mathbf{e}_\alpha$ – базис векторного пространства $\mathbb{V}$, то множество функционалов $\mathbf{e}^\alpha$, определенное соотношением

$$\mathbf{e}^\alpha(\mathbf{e}_\beta) = \delta_\beta^\alpha, \quad (1.49)$$

определяет единственный *сопряженный (дуальный) базис* сопряженного пространства. При этом любой функционал взаимно однозначно представим в виде

$$\mathbf{f} = \mathbf{e}^\alpha f_\alpha, \quad (1.50)$$

с некоторыми компонентами $f_\alpha \in \mathbb{R}$, а значение функционала на векторе равно сумме компонент

$$\mathbf{f}(\mathbf{x}) = x^\alpha f_\alpha. \quad (1.51)$$

Метрика в евклидовом пространстве $\mathbb{R}^n$ задает двойственное пространство $\mathbb{R}^{n*}$, элементы которого определяются компонентами $x_\alpha := x^\beta g_{\beta\alpha}$. Наличие метрики позволяет установить изоморфизм $\mathbb{R}^n$ и его дуального пространства $\mathbb{R}^{n*}$, рассматриваемых как векторные пространства.

Используя понятие линейного функционала, можно дать новое эквивалентное определение тензорного произведения векторных пространств.

Определение. Пусть $\mathbb{V}_1^*$ и $\mathbb{V}_2^*$ – векторные пространства, дуальные к $\mathbb{V}_1$ и $\mathbb{V}_2$. Тогда *тензорным произведением* векторных пространств $\mathbb{V}_1 \otimes \mathbb{V}_2$ называется множество всех билинейных отображений упорядоченной пары дуальных пространств в поле вещественных чисел

$$\mathbb{V}_1 \otimes \mathbb{V}_2 : (\mathbb{V}_1^*, \mathbb{V}_2^*) \rightarrow \mathbb{R}. \quad \square \quad (1.52)$$

В компонентах отображение (1.52) записывается следующим образом. Если $\mathbf{f}_1 = f_{1\alpha} \mathbf{e}^\alpha \in \mathbb{V}_1$ и $\mathbf{f}_2 = f_{2i} \mathbf{e}^i \in \mathbb{V}_2$, то элемент тензорного произведения $x^{\alpha i} \mathbf{e}_\alpha \otimes \mathbf{e}_i \in \mathbb{V}_1 \otimes \mathbb{V}_2$ задает отображение $x^{\alpha i} f_{1\alpha} f_{2i} \in \mathbb{R}$.

Пусть задано векторное пространство $\mathbb{V}$ и сопряженное к нему пространство $\mathbb{V}^*$. Тензором типа (r, s) мы будем называть элемент тензорного произведения

$$\underbrace{\mathbb{V} \otimes \dots \otimes \mathbb{V}}_r \otimes \underbrace{\mathbb{V}^* \otimes \dots \otimes \mathbb{V}^*}_s,$$

где пространства $\mathbb{V}$ и $\mathbb{V}^*$ встречаются, соответственно, r и s раз. Для определенности векторное пространство $\mathbb{V}$ выбрано в качестве первых r сомножителей, хотя возможен и любой другой порядок множителей. Если в пространствах $\mathbb{V}$ и $\mathbb{V}^*$ заданы базисы $\mathbf{e}_\alpha$ и $\mathbf{e}^\alpha$, то тензор типа (r, s) будет иметь r верхних и s нижних индексов

$$X = X^{\alpha_1 \dots \alpha_r}{}_{\beta_1 \dots \beta_s} \mathbf{e}_{\alpha_1} \otimes \dots \otimes \mathbf{e}_{\alpha_r} \otimes \mathbf{e}^{\beta_1} \otimes \dots \otimes \mathbf{e}^{\beta_s},$$

которые называются *контравариантными* и *ковариантными*, соответственно. Поскольку тензорное произведение некоммутативно, то порядок индексов является существенным. Так $X^{\alpha\beta} \neq X^{\beta\alpha}$ и $X^\alpha{}_\beta \neq X_\beta{}^\alpha$.

Определение. Пусть задано линейное отображение двух векторных пространств

$$\varphi : \mathbb{V} \rightarrow \mathbb{W},$$

которые могут иметь разные размерности: $\dim \mathbb{V} = n$, $\dim \mathbb{W} = m$. В компонентах это отображение задается $n \times m$ -матрицей $A = (A_\alpha{}^i)$

$$\varphi : \mathbb{V} \ni \{x^\alpha\} \mapsto \{y^i := x^\alpha A_\alpha{}^i\} \in \mathbb{W}.$$

Пусть $\mathbf{f}$ и $\mathbf{g}$ – линейные формы на $\mathbb{V}$ и $\mathbb{W}$, соответственно. Тогда отображение φ индуцирует *возврат отображения* φ^* , действующий в обратную сторону

$$\varphi^* : \mathbb{W}^* \rightarrow \mathbb{V}^*$$

по следующему правилу

$$(\mathbf{f} = \varphi^* \mathbf{g})(\mathbf{x}) := \mathbf{g}(\varphi(\mathbf{x})).$$

В компонентах возврат отображения имеет вид

$$\varphi^* : \mathbb{W}^* \ni \{g_i\} \mapsto \{f_\alpha := A_\alpha{}^i g_i\} \in \mathbb{V}^*. \quad \square$$

Замечание. Матрица A_α^i в общем случае является прямоугольной, и говорить об обратной матрице не имеет смысла. $\square$

Понятие тензоров с ковариантными и контравариантными индексами будет использовано в дальнейшем при построении тензорных полей на многообразиях.

Структура линейных отображений

Рассмотрим два векторных пространства $\mathbb{V}$ ($\dim \mathbb{V} = n$), $\mathbb{W}$ ($\dim \mathbb{W} = m$) и линейное отображение $A : \mathbb{V} \rightarrow \mathbb{W}$. Базисы этих пространств обозначим, соответственно, через $\mathbf{e}_\alpha$, $\alpha = 1, \dots, n$, и $\mathbf{e}_i$, $i = 1, \dots, m$. Тогда линейное отображение взаимно однозначно задается матрицей A_α^i , которая определяется отображением базисных векторов: $A(\mathbf{e}_\alpha) = A_\alpha^i \mathbf{e}_i$. Поэтому и для отображения, и для матрицы мы будем использовать одну и ту же букву. Смысл всегда понятен из контекста.

Изучение структуры линейных отображений сводится к приведению матрицы A_α^i к возможно более простому (каноническому) виду путем выбора соответствующих базисов в векторных пространствах. Если пространства $\mathbb{V}$ и $\mathbb{W}$ никак между собой не связаны, то общая структура линейного отображения дается следующим утверждением.

Теорема 1.3.17. Пусть $A : \mathbb{V} \rightarrow \mathbb{W}$ – линейное отображение векторных пространств. Тогда справедливы следующие утверждения:

а) Существуют такие прямые разложения $\mathbb{V} = \mathbb{V}_0 \oplus \mathbb{V}_1$ и $\mathbb{W} = \mathbb{W}_1 \oplus \mathbb{W}_2$, что $\ker A = \mathbb{V}_0$ (т.е. $A(\mathbb{V}_0) = \mathbf{0} \in \mathbb{W}$) и A индуцирует изоморфизм $\mathbb{V}_1$ с $\mathbb{W}_1$.

б) Существуют такие базисы в $\mathbb{V}$ и $\mathbb{W}$, что элементы матрицы A_α^i в этих базисах имеют вид $A_a^a = 1$ (суммирование нет) при $1 \leq a \leq r$, где r – ранг матрицы A_α^i , и $A_\alpha^i = 0$ для всех остальных случаев.

в) Пусть A – некоторая матрица размера $n \times m$. Тогда существуют такие невырожденные квадратные матрицы B и C размеров $n \times n$ и $m \times m$ и такое число $r \leq \min(n, m)$, что матрица BAC имеет вид, описанный в предыдущем пункте. Число r определено однозначно и равно рангу матрицы A .

Доказательство. а) Положим $\mathbb{V}_0 := \ker A$, а в качестве $\mathbb{V}_1$ выберем прямое дополнение к $\mathbb{V}_0$. Это возможно в силу предложения 1.3.5. Теперь положим $\mathbb{W}_1 := A(\mathbb{V})$, а в качестве $\mathbb{W}_2$ выберем прямое дополнение к $\mathbb{W}_1$. По построению отображение $A : \mathbb{V}_1 \rightarrow \mathbb{W}_1$ взаимно однозначно, т.к. ядро $\mathbb{V}_0$ отображения A пересекается с $\mathbb{V}_1$ лишь по нулю.

б) Положим $r = \dim \mathbb{V}_1 = \dim \mathbb{W}_1$ и выберем в $\mathbb{V}$ базис $\{\mathbf{e}_1, \dots, \mathbf{e}_r; \mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$, где первые r векторов образуют базис в $\mathbb{V}_1$, а следующие – базис в $\mathbb{V}_0$. Далее, векторы $\mathbf{e}'_a := A(\mathbf{e}_a)$, $1 \leq a \leq r$, образуют базис в $\mathbb{W}_1$. Дополним его до базиса в $\mathbb{W}$ векторами $\mathbf{e}'_{r+1}, \dots, \mathbf{e}'_m$. По построению

$$\begin{aligned} A(\mathbf{e}_1, \dots, \mathbf{e}_r; \mathbf{e}_{r+1}, \dots, \mathbf{e}_n) &= (\mathbf{e}'_1, \dots, \mathbf{e}'_r; 0, \dots, 0) = \\ &= (\mathbf{e}'_1, \dots, \mathbf{e}'_r; \mathbf{e}'_{r+1}, \dots, \mathbf{e}'_m) \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

где $\mathbb{1}$ – единичная $r \times r$ матрица. Следовательно, матрица A в этих базисах имеет требуемый вид.

в) Построим по матрице A линейное отображение $\mathbb{R}^n \rightarrow \mathbb{R}^m$ с этой матрицей. Затем применим к этому отображению утверждение б). В новых базисах матрица

A будет иметь требуемый вид и выражаться через A в виде BAC , где B и C – невырожденные матрицы перехода. Наконец, $\text{rank } A = \text{rank}(BAC) = \dim(A(\mathbb{V}))$. $\square$

При $\mathbb{W} = \mathbb{V}$ структура линейного отображения $A : \mathbb{V} \rightarrow \mathbb{V}$ становится намного богаче и интереснее. В этом случае отображение A является линейным оператором, действующим в векторном пространстве $\mathbb{V}$, который взаимно однозначно задается некоторой квадратной $n \times n$ матрицей A_α^β .

Определение. Подпространство $\mathbb{V}_0 \subset \mathbb{V}$ называется *инвариантным* относительно действия оператора A , если $A(\mathbb{V}_0) \subseteq \mathbb{V}_0$. Оператор $A : \mathbb{V} \rightarrow \mathbb{V}$ называется *диагонализируемым*, если выполнено одно из двух равносильных условий:

- а) $\mathbb{V}$ разлагается в прямую сумму одномерных инвариантных подпространств;
- б) существует базис $\mathbb{V}$, в котором матрица оператора A диагональна. $\square$

Равносильность условий диагонализируемости операторов проверяется без труда. Если в базисе $\mathbf{e}_\alpha$ матрица оператора A диагональна, то $A(\mathbf{e}_\alpha) = \lambda_\alpha \mathbf{e}_\alpha$, и одномерные подпространства $\mathbb{V}_\alpha$, натянутые на $\mathbf{e}_\alpha$, инвариантны, а векторное пространство $\mathbb{V}$ разлагается в их прямую сумму. Наоборот, если $\mathbb{V} = \bigoplus \mathbb{V}_\alpha$ – такое разложение и $\mathbf{e}_\alpha$ – любой ненулевой вектор из $\mathbb{V}_\alpha$, то совокупность векторов $\mathbf{e}_\alpha$ образует базис в $\mathbb{V}$.

Диагонализируемые операторы образуют простейший и во многих отношениях самый важный класс. Например, над полем комплексных чисел любой оператор можно сделать диагональным, как угодно мало изменив его матрицу, так что оператор “в общем положении” диагонализируем.

Чтобы продемонстрировать, что может мешать оператору быть диагонализируемым, дадим

Определение. Одномерное подпространство $\mathbb{V}_1 \subset \mathbb{V}$ называется *собственным* для оператора A , если оно инвариантно, т.е. $A(\mathbb{V}_1) \subseteq \mathbb{V}_1$. В этом подпространстве оператор A действует умножением на скаляр $\lambda \in \mathbb{R}$, т.е. $A(\mathbf{x}) = \lambda \mathbf{x}$ для всех $\mathbf{x} \in \mathbb{V}_1$. Этот скаляр называется *собственным значением* оператора A на $\mathbb{V}_1$.

Вектор $\mathbf{x} \in \mathbb{V}$ называется *собственным* для оператора A , если линейная оболочка $\mathbb{R}\mathbf{x}$ является собственным подпространством. Иными словами, $\mathbf{x} \neq \mathbf{0}$ и $A(\mathbf{x}) = \lambda \mathbf{x}$ для подходящего $\lambda \in \mathbb{R}$.

Полином (многочлен) $P(t) := \det(t\mathbb{1} - A)$, где $t \in \mathbb{R}$, называется *характеристическим*. Множество всех корней характеристического полинома называется *спектром* оператора A . Если кратности всех собственных значений оператора A равны единице, то говорят, что оператор A имеет *простой спектр*. $\square$

Нулевой вектор $\mathbf{0} \in \mathbb{V}$ удовлетворяет уравнению $A(\mathbf{0}) = \lambda \mathbf{0}$ для всех λ . Однако он не является собственным, т.к. не определяет никакого одномерного подпространства $\mathbb{V}_1 \subset \mathbb{V}$. Поэтому собственный вектор всегда отличен от нуля. Он определен с точностью до умножения на отличную от нуля постоянную. В то же время собственное значение оператора может быть равно нулю для отличного от нуля собственного вектора. В этом случае оператор A имеет нетривиальное ядро и вырожден.

Степень характеристического полинома равна размерности векторного пространства $\dim \mathbb{V} = n$, и он имеет следующую структуру

$$P(t) = t^n - \text{tr } A \cdot t^{n-1} + \dots + (-1)^n \det A.$$

Отсюда следует, что для диагонализируемых операторов след матрицы A равен сумме собственных значений, а определитель A – их произведению.

Теорема 1.3.18. а) Характеристический полином линейного оператора A не зависит от выбора базиса, в котором представлена его матрица.
 б) Любое собственное значение λ оператора A является корнем характеристического полинома, $P(\lambda) = 0$, и любой корень, лежащий в $\mathbb{R}$, является собственным значением для A , отвечающим некоторому (не обязательно единственному) собственному подпространству в $\mathbb{V}$.

Доказательство. а) При замене базиса $\mathbf{e}_\alpha \mapsto S_\alpha^\beta \mathbf{e}_\beta$, $\det S \neq 0$, матрица оператора преобразуется по правилу подобия $A \mapsto SAS^{-1}$. Поэтому

$$\det(t\mathbb{1} - SAS^{-1}) = \det(S(t\mathbb{1} - A)S^{-1}) = \det(t\mathbb{1} - A).$$

б) Пусть $\lambda \in \mathbb{R}$ – корень характеристического полинома $P(t)$. Тогда отображение $\lambda \text{id} - A$ представлено вырожденной матрицей и, значит, имеет нетривиальное ядро. Допустим, что $\mathbf{x} \neq 0$ – вектор из ядра, тогда $A(\mathbf{x}) = \lambda \mathbf{x}$. Так что λ есть собственное значение оператора A и $\mathbf{x}$ – собственный вектор. Наоборот, если $A(\mathbf{x}) = \lambda \mathbf{x}$, то вектор $\mathbf{x}$ лежит в ядре оператора $\lambda \text{id} - A$ и, следовательно, $P(\lambda) = 0$. $\square$

Из этой теоремы следует, что оператор A не диагонализировать, если, например, характеристический полином $P(t)$ вообще не имеет вещественных корней, т.к. поле вещественных чисел $\mathbb{R}$ алгебраически не замкнуто. Если же рассматриваются комплексные векторные пространства, то характеристический полином всегда имеет n комплексных корней, т.к. поле комплексных чисел $\mathbb{C}$ алгебраически замкнуто. Напомним

Определение. Поле $\mathbb{k}$ называется *алгебраически замкнутым*, если любой полином $f(t)$ от одной переменной t с коэффициентами из поля $\mathbb{k}$ имеет корень из этого поля $\lambda \in \mathbb{k}$. $\square$

Если векторное пространство $\mathbb{V}$ рассматривается над алгебраически замкнутым полем $\mathbb{k}$, то линейный оператор A всегда имеет собственное подпространство. Однако он все равно может оказаться недиагонализированным, ибо сумма всех собственных подпространств может оказаться меньше $\mathbb{V}$.

Пример 1.3.65. Пусть $\mathbb{V}$ – двумерное комплексное векторное пространство, в котором зафиксирован какой-либо базис. Тогда линейный оператор представляется комплексной 2×2 матрицей

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Характеристический многочлен оператора A имеет вид

$$P(t) = t^2 - (a + d)t + (ad - bc).$$

Его корни суть

$$\lambda_{1,2} = -\frac{a+d}{2} \pm \sqrt{\frac{(a-d)^2}{4} + bc}.$$

Возможны два случая.

1) $\lambda_1 \neq \lambda_2$. Пусть $\mathbf{e}_1$ и $\mathbf{e}_2$ – собственные векторы для собственных значений λ_1 и λ_2 , соответственно. Они линейно независимы. Действительно, допустим, что $f\mathbf{e}_1 + g\mathbf{e}_2 = 0$, где $f, g \in \mathbb{C}$. Тогда выполнены равенства:

$$A(f\mathbf{e}_1 + g\mathbf{e}_2) = f\lambda_1\mathbf{e}_1 + g\lambda_2\mathbf{e}_2 = 0.$$

Отсюда следует

$$\lambda_1(f\mathbf{e}_1 + g\mathbf{e}_2) - (f\lambda_1\mathbf{e}_1 + g\lambda_2\mathbf{e}_2) = g(\lambda_1 - \lambda_2)\mathbf{e}_2 = 0,$$

т.е. $g = 0$. Аналогично получим $f = 0$. Следовательно, в базисе $\{\mathbf{e}_1, \mathbf{e}_2\}$ матрица A диагональна.

2) $\lambda_1 = \lambda_2 = \lambda$. В этом случае оператор A диагонализировать, только если он умножает на λ все векторы из $\mathbb{V}$. Это значит, что

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix},$$

т.е. $a = d = \lambda$ и $b = c = 0$. Если эти условия не выполнены, а выполнено более слабое условие $(a - d)^2 + 4bc = 0$, эквивалентное равенству $\lambda_1 = \lambda_2$, то оператор A имеет только один собственный вектор, с точностью до умножения на ненулевую постоянную. Поэтому оператор A в этом случае недиагонализировать.

Пример такой матрицы:

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Эта матрица называется *жордановой клеткой* размера 2×2 (или ранга 2). $\square$

В приложении 37.1 дано общее определение жордановой матрицы и сформулирована теорема о приведении произвольной матрицы к жордановой форме. Переформулируем эту теорему для описания канонической формы линейных операторов.

Теорема 1.3.19. Пусть $\mathbb{V}$ – конечномерное векторное пространство над алгебраически замкнутым полем $\mathbb{K}$ и $A : \mathbb{V} \rightarrow \mathbb{V}$ – линейный оператор. Тогда существует жорданов базис, в котором матрица A имеет жорданову форму (37.29). Жорданова матрица определена однозначно с точностью до перестановки входящих в нее жордановых клеток.

Доказательство. См., например, [12], часть 1, §9, теорема 1. $\square$

В заключение данного раздела сформулируем еще одну теорему для нужд последующего рассмотрения.

Определение. Для любого многочлена $Q(t) := \sum_{i=1}^n a_i t^i$ с коэффициентами из поля $\mathbb{K}$ имеет смысл выражение $Q(A) := \sum_{i=1}^n a_i A^i$, где A можно понимать как матрицу. Будем говорить, что многочлен $Q(t)$ *аннулирует* оператор A , если $Q(A) = 0$. Рассмотрим многочлен $M(t)$ со старшим коэффициентом единица, который аннулирует A и имеет наименьшую возможную степень. Он называется *минимальным многочленом* оператора A . $\square$

Ненулевые многочлены, аннулирующие оператор A , существуют всегда. Действительно, $n \times n$ матрица A содержит n^2 элементов, и поэтому операторы $\text{id}, A, A^2, \dots, A^{n^2}$ линейно зависимы. Отсюда следует, что существует аннулирующий A многочлен степени не выше n^2 . Ниже мы увидим, что для каждого оператора существует аннулятор степени n .

Минимальный многочлен определен однозначно. Действительно, если $M_1(t)$ и $M_2(t)$ – два таких многочлена, то их разность $M_1(t) - M_2(t)$ аннулирует A и имеет меньшую степень. Кроме того, любой многочлен, аннулирующий оператор A , делится на минимальный многочлен. Допустим, что $Q(A) = 0$. Разделим Q с остатком на минимальный многочлен: $Q(t) = X(t)M(t) + R(t)$, где степень остатка R меньше степени минимального многочлена. Тогда $R(A) = Q(A) - X(A)M(A) = 0$, что противоречит определению минимального многочлена. Поэтому $R(t) = 0$.

Теорема 1.3.20 (Гамильтона–Кэли). *Характеристический полином $P(t)$ оператора A аннулирует этот оператор.*

Доказательство. См., например, [12], часть 1, §8, теорема 12. $\square$

Комплексификация

В приложениях часто используются также комплексные векторные пространства. Если задано вещественное векторное пространство, то оно естественным образом определяет комплексное векторное пространство.

Определение. *Комплексификацией* вещественного векторного пространства $\mathbb{V}$, $\dim \mathbb{V} = n$, называется комплексное векторное пространство $\mathbb{V}^{\mathbb{C}}$ той же комплексной размерности $\dim_{\mathbb{C}}(\mathbb{V}^{\mathbb{C}}) = n$ и вдвое большей вещественной размерности $\dim_{\mathbb{R}}(\mathbb{V}^{\mathbb{C}}) = 2n$, состоящее из пар $(\mathbf{x}, \mathbf{y})$, которые обозначаются $\mathbf{z} := \mathbf{x} + i\mathbf{y}$, где $\mathbf{x}, \mathbf{y} \in \mathbb{V}$, с обычными операциями сложения и умножения на комплексные числа:

$$\begin{aligned} \mathbf{z}_1 + \mathbf{z}_2 &:= (\mathbf{x}_1 + \mathbf{x}_2) + i(\mathbf{y}_1 + \mathbf{y}_2), \\ (a + ib)\mathbf{z} &:= (a\mathbf{x} - b\mathbf{y}) + i(b\mathbf{x} + a\mathbf{y}). \end{aligned}$$

Комплексификацией линейного оператора $A : \mathbb{V} \rightarrow \mathbb{V}$ называется линейный оператор

$$A^{\mathbb{C}} : \mathbb{V}^{\mathbb{C}} \ni \mathbf{x} + i\mathbf{y} \mapsto A^{\mathbb{C}}(\mathbf{x} + i\mathbf{y}) := A\mathbf{x} + iA\mathbf{y} \in \mathbb{V}^{\mathbb{C}}. \quad \square$$

Пример 1.3.66. Рассмотрим вещественную прямую как вещественное векторное пространство, $\mathbf{x} = xe \in \mathbb{R}$. Тогда $\mathbf{z} = xe + iye$ – элемент комплексификации $\mathbb{R}^{\mathbb{C}}$, которая естественным образом отождествляется с комплексной плоскостью: $\mathbb{R}^{\mathbb{C}} = \mathbb{C}$. $\square$

Выше мы показали, как вещественное векторное пространство порождает комплексное. Наоборот, каждое комплексное векторное пространство естественным образом порождает вещественное.

Определение. Пусть задано комплексное векторное пространство $\mathbb{W}$, $\dim_{\mathbb{C}} \mathbb{W} = n$, над полем комплексных чисел $\mathbb{C}$ с базисом $\mathbf{e}_{\alpha}$, $\alpha = 1, \dots, n$. Тогда произвольный вектор представим в виде

$$\mathbf{x} = z^{\alpha} \mathbf{e}_{\alpha} = x^{\alpha} \mathbf{e}_{\alpha} + y^{\alpha} (i\mathbf{e}_{\alpha}), \quad (1.53)$$

где x^{α} и y^{α} – вещественные и мнимые части комплексных компонент z^{α} данного вектора. Примем векторы $\mathbf{e}_{\alpha}$ и $i\mathbf{e}_{\alpha}$ в качестве базиса нового вещественного векторного пространства. Оно называется *овеществлением* $\mathbb{W}$ и обозначается $\mathbb{W}^{\mathbb{R}}$, $\dim_{\mathbb{R}}(\mathbb{W}^{\mathbb{R}}) = 2n$. $\square$

По возможности, мы будем обозначать вещественное векторное пространство буквой $\mathbb{V}$, а комплексное – $\mathbb{W}$.

В определении овеществления комплексного векторного пространства мы установили взаимно однозначное соответствие между векторами комплексного пространства и векторами вещественного векторного пространства вдвое большей размерности. При этом нулевому вектору комплексного векторного пространства $\mathbb{W}$ соответствует нулевой вектор из его овеществления $\mathbb{W}^{\mathbb{R}}$, и, наоборот, нулевому вектору из $\mathbb{W}^{\mathbb{R}}$

соответствует нулевой вектор из $\mathbb{W}$. Заметим также, что последовательное овеществление и комплексификация не дает исходное комплексное векторное пространство: $(\mathbb{W}^{\mathbb{R}})^{\mathbb{C}} \neq \mathbb{W}$ (не совпадают даже размерности).

Пусть

$$A: \quad \mathbb{W}_1 \rightarrow \mathbb{W}_2 \quad (1.54)$$

– линейное отображение двух комплексных векторных пространств. Ему соответствует линейное отображение овеществленных пространств:

$$A^{\mathbb{R}}: \quad \mathbb{W}_1^{\mathbb{R}} \rightarrow \mathbb{W}_2^{\mathbb{R}},$$

где $A^{\mathbb{R}}$ – овеществление отображения A . Допустим, что $\dim_{\mathbb{C}} \mathbb{W}_1 = m$ и $\dim_{\mathbb{C}} \mathbb{W}_2 = n$. Пусть $\mathbf{e}_{\alpha}$, $\alpha = 1, \dots, m$, и $\mathbf{e}_i$, $i = 1, \dots, n$ – базисы соответствующих пространств. Тогда линейное отображение (1.54) задается комплексной $m \times n$ матрицей A с элементами A_{α}^i . Выделим из этой матрицы вещественную и мнимую части:

$$A = B + iC.$$

Тогда $2m \times 2n$ матрица овеществленного отображения в базисах $\{\mathbf{e}_{\alpha}, i\mathbf{e}_{\alpha}\}$ и $\{\mathbf{e}_i, i\mathbf{e}_i\}$ будет иметь вид

$$A = B + iC \quad \mapsto \quad A^{\mathbb{R}} = \begin{pmatrix} B & C \\ -C & B \end{pmatrix}. \quad (1.55)$$

Пусть задано два линейных отображения комплексных векторных пространств:

$$\mathbb{W}_1 \xrightarrow{A_1} \mathbb{W}_2 \xrightarrow{A_2} \mathbb{W}_3.$$

Если во всех пространствах заданы базисы, то отображения A_1 и A_2 взаимно однозначно определяются комплексными матрицами соответствующих размерностей, которые мы разложим на вещественную и мнимую части. Тогда легко проверить формулу для произведения матриц самих отображений и их овеществлений:

$$\begin{aligned} A_1 A_2 &= B_1 B_2 - C_1 C_2 + i(B_1 C_2 + C_1 B_2) \quad \mapsto \\ &\mapsto \begin{pmatrix} B_1 & C_1 \\ -C_1 & B_1 \end{pmatrix} \begin{pmatrix} B_2 & C_2 \\ -C_2 & B_2 \end{pmatrix} = \begin{pmatrix} B_1 B_2 - C_1 C_2 & B_1 C_2 + C_1 B_2 \\ -B_1 C_2 - C_1 B_2 & B_1 B_2 - C_1 C_2 \end{pmatrix}, \end{aligned}$$

т.е. произведение комплексных матриц при овеществлении переходит в произведение вещественных. Нетрудно проверить следующие свойства матриц овеществленного отображения:

$$\begin{aligned} \text{id}^{\mathbb{R}} &= \text{id}, \\ (A_1 A_2)^{\mathbb{R}} &= A_1^{\mathbb{R}} A_2^{\mathbb{R}}, \\ (aA_1 + bA_2)^{\mathbb{R}} &= aA_1^{\mathbb{R}} + bA_2^{\mathbb{R}}, \quad a, b \in \mathbb{R}. \end{aligned}$$

Вторая формула была доказана выше. В последнем случае оба отображения A_1 и A_2 действуют из $\mathbb{W}_1$ в $\mathbb{W}_2$.

Предложение 1.3.17. Пусть $A: \mathbb{W} \rightarrow \mathbb{W}$ – линейный оператор на конечномерном комплексном векторном пространстве. Тогда $\det A^{\mathbb{R}} = |\det A|^2$.

Доказательство. Пусть отображение A в некотором базисе представимо квадратной матрицей $B + iC$. Тогда, применяя элементарные преобразования к блочной структуре сначала к строкам, а потом к столбцам, получим:

$$\begin{aligned} \det A^{\mathbb{R}} &= \det \begin{pmatrix} B & C \\ -C & B \end{pmatrix} = \det \begin{pmatrix} B - iC & C + iB \\ -C & B \end{pmatrix} = \\ &= \det \begin{pmatrix} B - iC & 0 \\ -C & B + iC \end{pmatrix} = \det(B - iC) \det(B + iC) = |\det A|^2. \quad \square \end{aligned}$$

Таким образом, определитель о вещественного линейного оператора в конечномерном комплексном векторном пространстве всегда положителен.

Пусть $\mathbb{W}$ – комплексное линейное пространство и $\mathbb{W}^{\mathbb{R}}$ – его о веществление. Для того, чтобы восстановить умножение на комплексные числа в вещественном пространстве $\mathbb{W}^{\mathbb{R}}$, достаточно знать оператор $J : \mathbb{W}^{\mathbb{R}} \rightarrow \mathbb{W}^{\mathbb{R}}$, который соответствует умножению на мнимую единицу i : $J(\mathbf{x}) = i\mathbf{x}$, где в левой части $\mathbf{x} \in \mathbb{W}^{\mathbb{R}}$, а в правой части $\mathbf{x} \in \mathbb{W}$ (взаимно однозначное соответствие векторов задается формулой (1.53)). Очевидно, что этот оператор линеен над $\mathbb{R}$ и удовлетворяет условию $J^2 = -\mathbb{1}$. Если оператор J известен, то для любого комплексного числа справедливо равенство

$$(a + ib)\mathbf{x} = a\mathbf{x} + bJ(\mathbf{x}), \quad a, b \in \mathbb{R}.$$

Это наблюдение приводит к следующему важному понятию.

Определение. Пусть $\mathbb{V}$ – вещественное векторное пространство. *Комплексной структурой* на $\mathbb{V}$ называется линейный оператор $J : \mathbb{V} \rightarrow \mathbb{V}$ (эндоморфизм), удовлетворяющий условию $J^2 = -\text{id}$. $\square$

Из определения следует, что оператор комплексной структуры всегда невырожден.

Данное определение оправдывается следующим утверждением.

Предложение 1.3.18. Пусть $(\mathbb{V}, J)$ – вещественное векторное пространство четной размерности, $\dim_{\mathbb{R}} \mathbb{V} = 2n$, с комплексной структурой. Введем на $\mathbb{V}$ операцию умножения на комплексные числа:

$$(a + ib)\mathbf{x} := a\mathbf{x} + bJ(\mathbf{x}).$$

Тогда $\mathbb{V}$ станет комплексным векторным пространством $\tilde{\mathbb{V}}$, для которого $\tilde{\mathbb{V}}^{\mathbb{R}} = \mathbb{V}$.

Доказательство. Прямая проверка. $\square$

Ясно, что $\dim_{\mathbb{C}} \tilde{\mathbb{V}} = n$, поскольку векторы $\mathbf{e}_{\alpha}$ и $J(\mathbf{e}_{\alpha})$ пропорциональны над полем комплексных чисел, а также $\dim_{\mathbb{R}} \tilde{\mathbb{V}}^{\mathbb{R}} = 2n$.

Следствие. Если $(\mathbb{V}, J)$ – вещественное векторное пространство с комплексной структурой, то его размерность четна, $\dim_{\mathbb{R}} \mathbb{V} = 2n$, и матрица комплексной структуры в подходящем базисе имеет вид

$$J = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}, \quad (1.56)$$

где $\mathbb{1}$ – единичная $n \times n$ матрица.

Доказательство. Из предложения 1.3.18 следует, что $\dim_{\mathbb{R}} \mathbb{V} = 2 \dim_{\mathbb{C}} \tilde{\mathbb{V}}$. Далее, выберем базис e_{α} , $\alpha = 1, \dots, n$, в комплексном векторном пространстве $\tilde{\mathbb{V}}$. Оператор умножения на i в этом базисе имеет вид $i\mathbb{1}$. Тогда из представления (1.55) следует, что матрица оператора комплексной структуры J в этом базисе имеет вид (1.56). $\square$

Определение. Комплексная структура вида (1.56) называется *канонической*. $\square$

Замечание. Каноническая комплексная структура (как матрица) с точностью до знака совпадает с канонической симплектической формой (19.1), которая будет введена позже. Однако есть существенное отличие. Комплексная структура является оператором, действующим в векторном пространстве, и поэтому несет один ковариантный и один контравариантный индексы. В общем случае (без метрики) про нее нельзя сказать, что она антисимметрична, т.к. индексы разных типов переставлять нельзя. В то же время симплектическая форма задается антисимметричным ковариантным тензором второго ранга. $\square$

Заметим, что для вещественной формы (1.55) любого комплексного линейного оператора выполнены равенства:

$$iA = Ai = -C + iB \quad \mapsto \quad (iA)^{\mathbb{R}} = \begin{pmatrix} -C & B \\ -B & -C \end{pmatrix} = JA^{\mathbb{R}} = A^{\mathbb{R}}J. \quad (1.57)$$

Предложение 1.3.19. Пусть $\mathbb{W}$ – комплексное векторное пространство и $B : \mathbb{W}^{\mathbb{R}} \rightarrow \mathbb{W}^{\mathbb{R}}$ – вещественное линейное отображение его о веществления. Тогда комплексное линейное отображение $A : \mathbb{W} \rightarrow \mathbb{W}$ такое, что $A^{\mathbb{R}} = B$, существует если и только если оператор B коммутирует с естественной комплексной структурой на $\mathbb{W}^{\mathbb{R}}$.

Доказательство. Пусть оператор A существует. Тогда оператор B коммутирует с естественной комплексной структурой J на $\mathbb{W}^{\mathbb{R}}$, ибо

$$B(J(\mathbf{x})) = A(i\mathbf{x}) = iA(\mathbf{x}) = J(B(\mathbf{x})), \quad \forall \mathbf{x} \in \mathbb{W}, \mathbb{W}^{\mathbb{R}},$$

где $\mathbf{x}$ понимается как вектор либо в $\mathbb{W}$, либо в $\mathbb{W}^{\mathbb{R}}$ в соответствии с разложением (1.53). Это условие является также достаточным, потому что из него вытекает линейность A над $\mathbb{C}$:

$$\begin{aligned} A((a + ib)\mathbf{x}) &= aA(\mathbf{x}) + bA(i\mathbf{x}) = aB(\mathbf{x}) + bB(J(\mathbf{x})) = \\ &= aB(\mathbf{x}) + bJ(B(\mathbf{x})) = (a + bJ)B(\mathbf{x}) = (a + ib)A(\mathbf{x}). \end{aligned} \quad \square$$

Линейный оператор $A : \mathbb{W} \rightarrow \mathbb{W}$ в комплексном векторном пространстве размерности $\dim_{\mathbb{C}} \mathbb{W} = n$ с некоторым базисом – это произвольная матрица из группы $A \in \mathrm{GL}(n, \mathbb{C})$. Каждой комплексной матрице взаимно однозначно ставится в соответствие $2n \times 2n$ вещественная матрица вида (1.55) с сохранением правила умножения матриц. Поэтому у группы $\mathrm{GL}(n, \mathbb{C})$ есть вещественное представление в группе матриц $\mathrm{GL}(2n, \mathbb{R})$. Оно состоит из вещественных $2n \times 2n$ матриц, которые коммутируют с канонической комплексной структурой (1.56) (следствие предложения 1.3.19).

Пусть $\mathbb{V}$ – четномерное вещественное линейное пространство и $B : \mathbb{V} \rightarrow \mathbb{V}$ – вещественный линейный оператор. Поставим вопрос, когда на $\mathbb{V}$ существует такая комплексная структура J , что B является о веществлением некоторого комплексного линейного отображения $A : \tilde{\mathbb{V}} \rightarrow \tilde{\mathbb{V}}$, где $\tilde{\mathbb{V}}$ – комплексное векторное пространство, построенное с помощью J (предложение 1.3.18). Частичный ответ на этот вопрос дает

Предложение 1.3.20. Пусть $\mathbb{V}$, $\dim_{\mathbb{R}} \mathbb{V} = 2$, – вещественное линейное пространство и $B : \mathbb{V} \rightarrow \mathbb{V}$ – вещественный линейный оператор. Тогда, если B не имеет собственных векторов в $\mathbb{V}$, то существует такая комплексная структура J на $\mathbb{V}$, что $B = A^{\mathbb{R}}$, где $A : \tilde{\mathbb{V}} \rightarrow \tilde{\mathbb{V}}$ – некоторый комплексный линейный оператор, который действует в комплексном векторном пространстве $\tilde{\mathbb{V}}$, построенном из $\mathbb{V}$ с помощью J в предложении 1.3.18.

Доказательство. По условию предложения оператор B имеет два комплексно сопряженных собственных значения $\lambda \pm i\mu$, где $\lambda, \mu \in \mathbb{R}$ и $\mu \neq 0$. Положим

$$J := \frac{1}{\mu}(B - \lambda \mathbb{1}).$$

По теореме Гамильтона–Кэли 1.3.20 характеристический многочлен от матрицы оператора равен нулю:

$$B^2 - 2\lambda B + (\lambda^2 + \mu^2)\mathbb{1} = 0.$$

Отсюда следует равенство

$$J^2 = \frac{1}{\mu^2}(B^2 - 2\lambda B + \lambda^2 \mathbb{1}) = -\mathbb{1}.$$

Кроме того, J коммутирует с B . □

Дадим эквивалентное определение комплексификации вещественного векторного пространства $\mathbb{V}$, которое возникает путем расширения поля скаляров с $\mathbb{R}$ до $\mathbb{C}$.

Определение. Пусть $\mathbb{V}$ – вещественное векторное пространство. Введем на прямой сумме $\mathbb{V} \oplus \mathbb{V}$ комплексную структуру по формуле:

$$J(\mathbf{x}_1, \mathbf{x}_2) := (-\mathbf{x}_2, \mathbf{x}_1), \quad \forall (\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{V} \oplus \mathbb{V}.$$

Ясно, что $J^2 = -\mathbb{1}$. Назовем *комплексификацией* $\mathbb{V}^{\mathbb{C}}$ пространства $\mathbb{V}$ комплексное пространство $\widetilde{\mathbb{V} \oplus \mathbb{V}}$, построенное по этой структуре в предложении 1.3.18. □

Отождествим векторное пространство $\mathbb{V}$ с подмножеством векторов вида $(\mathbf{x}, 0) \in \mathbb{V} \oplus \mathbb{V}$. Тогда

$$i(\mathbf{x}, 0) = J(\mathbf{x}, 0) = (0, \mathbf{x}),$$

и любой вектор из $\mathbb{V}^{\mathbb{C}}$ можно записать в виде

$$(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1, 0) + (0, \mathbf{x}_2) = (\mathbf{x}_1, 0) + i(\mathbf{x}_2, 0) = \mathbf{x}_1 + i\mathbf{x}_2 = (x_1^{\alpha} + ix_2^{\alpha})(\mathbf{e}_{\alpha}, 0).$$

Иными словами, $\mathbb{V}^{\mathbb{C}} = \mathbb{V} \oplus_{\mathbb{R}} i\mathbb{V}$, где прямая сумма рассматривается над $\mathbb{R}$, но не над $\mathbb{C}$.

Произвольный базис вещественного векторного пространства $\mathbb{V}$ можно рассматривать в качестве базиса в его комплексификации $\mathbb{V}^{\mathbb{C}}$. Поэтому $\dim_{\mathbb{R}} \mathbb{V} = \dim_{\mathbb{C}} \mathbb{V}^{\mathbb{C}}$. Другими словами, мы выбираем некоторый базис в вещественном векторном пространстве $\mathbb{V}$ и меняем вещественные компоненты векторов на комплексные.

Вспомнив понятие тензорного произведения, можно записать

$$\mathbb{V}^{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} \mathbb{V},$$

где комплексная плоскость $\mathbb{C}$ рассматривается, как двумерное вещественное векторное пространство. Оба способа обозначения комплексификации вещественного векторного пространства $\mathbb{V}$ являются стандартными.

Пусть $A : \mathbb{V}_1 \rightarrow \mathbb{V}_2$ – отображение вещественных векторных пространств. Если в пространствах $\mathbb{V}_1$ и $\mathbb{V}_2$ заданы базисы, то их же можно рассматривать в качестве базисов в комплексных пространствах $\mathbb{V}_1^{\mathbb{C}}$ и $\mathbb{V}_2^{\mathbb{C}}$. Тогда матрица отображения A в исходной паре базисов совпадает с матрицей комплексифицированного отображения $A^{\mathbb{C}}$. В частности, комплексные собственные значения отображений A и $A^{\mathbb{C}}$ и их жордановы формы совпадают. Справедливы также формулы: $\det A^{\mathbb{C}} = \det A$ и $\operatorname{tr} A^{\mathbb{C}} = \operatorname{tr} A$.

Теперь проследим, что происходит с векторными пространствами при последовательных комплексификации и овеществлении, которые проводятся в различных порядках.

Пусть $\mathbb{V}$ – вещественное векторное пространство. Тогда существует естественный изоморфизм

$$(\mathbb{V}^{\mathbb{C}})^{\mathbb{R}} \simeq \mathbb{V} \oplus \mathbb{V}.$$

Действительно, из определения следует, что $\mathbb{V}^{\mathbb{C}}$ совпадает с $\mathbb{V} \oplus \mathbb{V}$ как вещественное векторное пространство. Аналогично, $(A^{\mathbb{C}})^{\mathbb{R}} \mapsto A \oplus A$ в смысле этого отождествления для любого вещественного линейного отображения $A : \mathbb{V} \rightarrow \mathbb{V}$.

Композиция преобразований в обратном порядке приводит к менее очевидному ответу. Для формулировки результата введем

Определение. *Сопряженным комплексным векторным пространством $\bar{\mathbb{V}}$ к комплексному векторному пространству $\mathbb{V}$ называется множество $\mathbb{V}$ с той же структурой аддитивной группы, но с новым умножением на скаляры из $\mathbb{C}$, которое обозначим через $a * \mathbf{x}$:*

$$a * \mathbf{x} := \bar{a}\mathbf{x}, \quad \forall a \in \mathbb{C}, \quad \forall \mathbf{x} \in \mathbb{V},$$

где черта обозначает комплексное сопряжение.

Если $(\mathbb{V}, J)$ – вещественное векторное пространство с комплексной структурой, то оператор $-J$ называется *сопряженной комплексной структурой*. $\square$

Легко проверить, что все аксиомы векторного пространства выполнены.

Если $\tilde{\mathbb{V}}$ – комплексное векторное пространство из предложения 1.3.18, соответствующее паре $(\mathbb{V}, J)$, то $\tilde{\tilde{\mathbb{V}}}$ – комплексное векторное пространство, отвечающее паре $(\mathbb{V}, -J)$.

Пусть теперь $\mathbb{W}$ – комплексное векторное пространство. Построим комплексно линейный изоморфизм

$$A : (\mathbb{W}^{\mathbb{R}})^{\mathbb{C}} \simeq \mathbb{W} \oplus \bar{\mathbb{W}}.$$

С этой целью заметим, что на $(\mathbb{W}^{\mathbb{R}})^{\mathbb{C}}$ имеются два вещественно линейных оператора: оператор канонической комплексной структуры $J(\mathbf{x}_1, \mathbf{x}_2) = (-\mathbf{x}_2, \mathbf{x}_1)$ и оператор умножения на i , отвечающий исходной комплексной структуре $\mathbb{V}$: $i(\mathbf{x}_1, \mathbf{x}_2) = (i\mathbf{x}_1, i\mathbf{x}_2)$. Так как J коммутирует с i , он комплексно линеен в этой структуре. Поскольку $J^2 = -\operatorname{id}$, его собственные значения равны $\pm i$. Введем обозначения для двух подпространств, отвечающих этим собственным значениям:

$$\mathbb{W}^{1,0} := \{(\mathbf{x}_1, \mathbf{x}_2) \in (\mathbb{W}^{\mathbb{R}})^{\mathbb{C}} : J(\mathbf{x}_1, \mathbf{x}_2) = i(\mathbf{x}_1, \mathbf{x}_2)\},$$

$$\mathbb{W}^{0,1} := \{(\mathbf{x}_1, \mathbf{x}_2) \in (\mathbb{W}^{\mathbb{R}})^{\mathbb{C}} : J(\mathbf{x}_1, \mathbf{x}_2) = -i(\mathbf{x}_1, \mathbf{x}_2)\}.$$

Оба множества $\mathbb{W}^{1,0}$ и $\mathbb{W}^{0,1}$ являются комплексными подпространствами в $(\mathbb{W}^{\mathbb{R}})^{\mathbb{C}}$. Эти подпространства замкнуты относительно сложения и умножения на веществен-

ные числа, а замкнутость относительно умножения на J следует из того, что операторы J и i коммутируют. Покажем, что $\mathbb{W} = \mathbb{W}^{1,0} \oplus \mathbb{W}^{0,1}$, а также, что $\mathbb{W}^{1,0}$ естественно изоморфно $\mathbb{W}$, тогда как $\mathbb{W}^{0,1}$ естественно изоморфно $\overline{\mathbb{W}}$.

Из определений следует, что подпространство $\mathbb{W}^{1,0}$ состоит из векторов вида $(\mathbf{x}, -i\mathbf{x})$, а $\mathbb{W}^{0,1}$ – из векторов вида $(\mathbf{y}, i\mathbf{y})$. Для заданных $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{W}$ уравнение

$$(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}, -i\mathbf{x}) + (\mathbf{y}, i\mathbf{y})$$

на $\mathbf{x}, \mathbf{y}$ имеет единственное решение

$$\mathbf{x} = \frac{\mathbf{x}_1 + i\mathbf{x}_2}{2}, \quad \mathbf{y} = \frac{\mathbf{x}_1 - i\mathbf{x}_2}{2}.$$

Следовательно, $\mathbb{W} = \mathbb{W}^{1,0} \oplus \mathbb{W}^{0,1}$. Отображения

$$\begin{aligned} \mathbb{W} \ni \mathbf{x} &\mapsto (\mathbf{x}, -i\mathbf{x}) && \in \mathbb{W}^{1,0}, \\ \overline{\mathbb{W}} \ni \mathbf{x} &\mapsto (\mathbf{x}, i\mathbf{x}) && \in \mathbb{W}^{0,1} \end{aligned}$$

являются вещественно линейными изоморфизмами. Кроме того, они перестановочны с действием i на $\mathbb{W}, \overline{\mathbb{W}}$ и действием J на $\mathbb{W}^{1,0}, \mathbb{W}^{0,1}$ в силу определений.

Определение. Линейное отображение $A : \mathbb{W}_1 \rightarrow \mathbb{W}_2$ двух комплексных векторных пространств называется *антилинейным* или *полулинейным*, если отображение $A : \mathbb{W}_1 \rightarrow \overline{\mathbb{W}_2}$ линейно. Иными словами A – гомоморфизм аддитивных групп и

$$A(a\mathbf{x}) = \bar{a}A(\mathbf{x}), \quad \forall a \in \mathbb{C}, \quad \mathbf{x} \in \mathbb{W}_1. \quad \square$$

Пример 1.3.67. Оператор обращения времени в нерелятивистской квантовой механике является антилинейным оператором. $\square$

В заключение раздела укажем одно из приложений понятия комплексификации.

Предложение 1.3.21. Пусть $A : \mathbb{V} \rightarrow \mathbb{V}$ – линейный оператор в вещественном векторном пространстве. Тогда оператор A имеет инвариантное подпространство размерности 1 или 2.

Доказательство. Если оператор A имеет вещественное собственное значение, то подпространство, натянутое на соответствующий собственный вектор, инвариантно. В противном случае все собственные значения комплексны. Выберем одно из них $\lambda + i\mu$. Оно будет также собственным значением оператора $A^{\mathbb{C}}$, действующим в комплексифицированном пространстве $\mathbb{V}^{\mathbb{C}}$. Согласно определениям

$$A^{\mathbb{C}}(\mathbf{x}_1 + i\mathbf{x}_2) = A(\mathbf{x}_1) + iA(\mathbf{x}_2) = (\lambda + i\mu)(\mathbf{x}_1 + i\mathbf{x}_2) = (\lambda\mathbf{x}_1 - \mu\mathbf{x}_2) + i(\mu\mathbf{x}_1 + \lambda\mathbf{x}_2).$$

Следовательно, $A(\mathbf{x}_1) = \lambda\mathbf{x}_1 - \mu\mathbf{x}_2$ и $A(\mathbf{x}_2) = \mu\mathbf{x}_1 + \lambda\mathbf{x}_2$. Это означает, что линейная оболочка векторов $\mathbf{x}_1$ и $\mathbf{x}_2$ инвариантна. $\square$

1.3.4 $\mathbb{R}^n$ как аффинное пространство

Прямая в евклидовом пространстве обладает двумя важными и независимыми свойствами. Во-первых, отрезок, соединяющий две точки евклидова пространства, имеет наименьшую длину среди всех кривых, соединяющих эти точки (метрическое свойство.) Во-вторых, касательный вектор к прямой остается касательным при параллельном переносе вдоль прямой. В настоящем разделе мы придадим смысл последнему утверждению. Это важно, т.к. параллельный перенос векторов в дифференциальной геометрии является нетривиальным обобщением параллельного переноса векторов в аффинном пространстве.

В предыдущем разделе было установлено, что точки евклидова пространства $\mathbb{R}^n$ находятся во взаимно однозначном соответствии с множеством векторов $\mathbb{V}$ евклидова пространства, имеющих начало в начале системы координат. Компоненты этих векторов совпадают с координатами точек – концов векторов – и поэтому их можно складывать с координатами произвольных точек из $\mathbb{R}^n$, которые сами уже не рассматриваются, как элементы векторного пространства. Введение этой операции приводит к понятию аффинного пространства. Чтобы подчеркнуть разницу между точками и векторами дадим сначала общее определение.

Определение. *Аффинным пространством* над полем вещественных чисел называется тройка $(\mathbb{A}, \mathbb{V}, +)$, состоящая из множества $\mathbb{A}$, элементы которого называются точками, ассоциированного векторного пространства $\mathbb{V}$ и отображения

$$\mathbb{A} \times \mathbb{V} \ni P, x \mapsto P + x \in \mathbb{A}, \quad (1.58)$$

удовлетворяющего следующим аксиомам:

- 1) $(P + x) + y = P + (x + y) \quad \forall P \in \mathbb{A}, \quad \forall x, y \in \mathbb{V},$
- 2) $P + 0 = P, \quad \forall P \in \mathbb{A},$
- 3) для двух произвольных точек $P, Q \in \mathbb{A}$ уравнение $Q = P + x$ на вектор x имеет единственное решение. $\square$

В рассматриваемом случае множество точек представляет собой множество точек евклидова пространства $\mathbb{A} = \mathbb{R}^n$. Векторное пространство $\mathbb{V}$ также совпадает с евклидовым пространством $\mathbb{V} = \mathbb{R}^n$, которое рассматривается как векторное пространство. При этом отображение (1.58) называется *сдвигом* аффинного пространства на вектор x . Сдвиг на противоположный вектор $-x$ обозначается знаком минус: $P \mapsto P - x$.

Замечание. Отметим, что в определении аффинного пространства не фигурирует явно структура умножения на скаляры. Вообще говоря, для точек аффинного пространства не определено понятие суммы $P + Q$ и умножения на числа aP . $\square$

Тот единственный вектор $x \in \mathbb{V}$, для которого $Q = P + x$, удобно обозначать $Q - P$. Эта операция вычитания

$$\mathbb{A} \times \mathbb{A} \ni Q, P \mapsto Q - P \in \mathbb{V}$$

обладает следующими свойствами:

- 1) $(R - Q) + (Q - P) = R - P;$
- 2) $P - P = 0;$
- 3) $(P + x) - (Q + y) = (P - Q) + (x - y).$

Свойство 1), справедливое для трех произвольных точек аффинного пространства, называется *соотношением Шаля* и его можно записать в эквивалентной форме

$$(R - Q) + (Q - P) + (P - R) = 0.$$

Интуитивно аффинное пространство $(\mathbb{A}, \mathbb{V}, +)$ можно представлять себе как векторное пространство $\mathbb{V}$ с забытым началом координат $\mathbf{0}$, в котором введена операция сдвига на вектор.

В аффинном пространстве можно ввести аффинные координаты, состоящие из фиксированной точки $P_0 \in \mathbb{A}$ (начала координат) и базиса $\mathbf{e}_\alpha$ соответствующего векторного пространства. Это является следствием третьего свойства в определении аффинного пространства. Произвольная точка $P \in \mathbb{A}$ имеет единственное представление

$$P = P_0 + x_P^\alpha \mathbf{e}_\alpha, \quad (1.59)$$

где числа x_P^α называются координатами точки P в данной системе аффинных координат. Если предположить, что точка P_0 совпадает с началом координат в евклидовом векторном пространстве $\mathbb{R}^n$, то произвольная точка P определяется только координатами соответствующего вектора $P = \{x_P^\alpha\}$. Таким образом, мы имеем взаимно однозначное соответствие между точками аффинного пространства и векторами из $\mathbb{V}$. При этом сдвиг точки P (1.58) на вектор $\mathbf{x} = \{x^\alpha\}$ записывается в виде

$$x_P^\alpha \mapsto x_P^\alpha + x^\alpha.$$

После отождествления точек аффинного и векторного пространства можно считать, что в общем случае точка P имеет координаты

$$P = \{x_0^\alpha + x_P^\alpha\},$$

где x_0^α – координаты точки P_0 в разложении (1.59). Поскольку количество аффинных координат точек аффинного пространства равно n , то *размерностью* аффинного пространства называется размерность ассоциированного векторного пространства, $\dim \mathbb{A} := \dim \mathbb{V}$.

Пусть аффинное пространство имеет конечную размерность n . Если координаты всех точек аффинного пространства преобразовать по правилу

$$x^\alpha \mapsto x'^\alpha = x^\beta A_\beta^\alpha + a^\alpha, \quad (1.60)$$

где $A \in \mathrm{GL}(n, \mathbb{R})$ – произвольная невырожденная матрица и $\mathbf{a} = a^\alpha \mathbf{e}_\alpha \in \mathbb{V}$ – произвольный фиксированный вектор, то получим взаимно однозначное отображение аффинного пространства на себя. Последовательное выполнение двух преобразований (1.60) с параметрами $(A, \mathbf{a})$ и $(B, \mathbf{b})$ имеет вид

$$x''^\alpha = x^\gamma A_\gamma^\beta B_\beta^\alpha + a^\beta B_\beta^\alpha + b^\alpha.$$

Определение. Совокупность преобразований вида (1.60) с законом композиции

$$(B, \mathbf{b}) \circ (A, \mathbf{a}) = (AB, \mathbf{a}B + \mathbf{b})$$

образует группу Ли *аффинных преобразований* $\mathbb{A}(n)$ аффинного пространства $\mathbb{A}$. Преобразования (1.60) при $A_\beta^\alpha = \delta_\beta^\alpha$ называются *трансляциями* на вектор $\mathbf{a} \in \mathbb{V}$. $\square$

Множество трансляций образует инвариантную (нормальную) подгруппу аффинной группы, а сама аффинная группа представляет собой полупрямое произведение, которое будем обозначать символом $\ltimes$, группы невырожденных матриц $\mathrm{GL}(n, \mathbb{R})$ на подгруппу трансляций, которая естественным образом отождествляется с векторным пространством $\mathbb{V}$. Таким образом группа аффинных преобразований имеет вид

$$\mathbb{A}(n) = \mathrm{GL}(n, \mathbb{R}) \ltimes \mathbb{V}.$$

Подгруппа $\mathbb{A}(n)$, сохраняющая начало координат, совпадает с $\mathbb{GL}(n, \mathbb{R})$.

Аффинная группа является неабелевой и имеет размерность $\dim \mathbb{A}(n) = n^2 + n$. Она действует в аффинном пространстве транзитивно и эффективно (см. раздел 9.1). Аффинные преобразования называются также *движениями аффинного пространства*. Если $\det A > 0$ или $\det A < 0$, то движения аффинного пространства (1.60) называются, соответственно, *собственными* или *несобственными*.

Замечание. В векторном пространстве действует только группа $\mathbb{GL}(n, \mathbb{R})$, т.к. в нем отсутствует понятие сдвига (трансляции). $\square$

Определение. Подгруппа сдвигов действует на $\mathbb{R}^n$ свободно. Рассмотрим подгруппу $\mathbb{G} \subset \mathbb{V} = \mathbb{R}^n$, где $\mathbb{V}$ рассматривается, как группа трансляций, состоящую из всех сдвигов

$$\sum_{i=1}^m n^i \mathbf{a}_i, \quad n^i \in \mathbb{Z}, \quad (1.61)$$

на линейно независимые векторы $\mathbf{a}_i$, $i = 1, \dots, m \leq n$ с целыми коэффициентами. Эта подгруппа действует собственно разрывно (см. раздел 9.1). Поэтому, отождествляя точки евклидова пространства, связанные трансляциями (1.61), получим важный класс многообразий, называемых *цилиндрами*. Они являются фактор пространствами $\mathbb{R}^n / \mathbb{G}$. При $m = n$ цилиндр компактен и называется *тором* $\mathbb{T}^n$. $\square$

Пример 1.3.68. В частном случае, сдвиги евклидовой плоскости $\mathbb{R}^2$ с декартовыми координатами x, y на постоянный вектор $\mathbf{a} = \{a, 0\}$, $a \neq 0$, вдоль оси абсцисс:

$$x \mapsto x + a, \quad y \mapsto y,$$

приводит к простейшему двумерному цилиндру, хорошо известному из курса элементарной геометрии. $\square$

В одномерном случае, $n = 1$, тор совпадает с окружностью $\mathbb{T}^1 = \mathbb{S}^1$. В общем случае при $1 \leq m < n$ векторное пространство разлагается в прямую сумму $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$, где $\mathbb{W}$ – подпространство, натянутое на векторы $\mathbf{a}_i$. Отсюда следует

$$\frac{\mathbb{R}^n}{\mathbb{G}} = \frac{\mathbb{R}^{n-m} \times \mathbb{R}^m}{\mathbb{G}} = \mathbb{R}^{n-m} \times \frac{\mathbb{R}^m}{\mathbb{G}}.$$

То есть цилиндры представляют собой прямое произведение $(n - m)$ -мерного аффинного пространства и m -мерного тора.

Вернемся к аффинным преобразованиям (1.60).

Теорема 1.3.21. *Всякое аффинное преобразование n -мерного евклидова пространства $\mathbb{A} = \mathbb{R}^n$ может быть представлено в виде композиции трех отображений:*

- 1) *п растяжений с положительными коэффициентами вдоль n попарно ортогональных осей, проходящих через некоторую точку $P \in \mathbb{A}$;*
- 2) *движения, оставляющего неподвижной точку P ;*
- 3) *сдвига.*

Доказательство. Первые два преобразования вытекают из полярного разложения матриц, теорема 37.1.5. Остаются еще сдвиги. Детали доказательства приведены, например, в [12], часть 3, §2, теорема 8. $\square$

Определение. В аффинном пространстве *прямой линией*, проходящей через точку $P \in \mathbb{A}$ в направлении вектора $x \in \mathbb{V}$, называется множество точек вида

$$P + ax, \quad a \in \mathbb{R}. \quad (1.62)$$

Две прямые линии называются *параллельными*, если они задаются уравнениями

$$P + ax \quad \text{и} \quad Q + ax, \quad a \in \mathbb{R}, \quad P \neq Q.$$

Подмножество $\mathbb{B} \subset \mathbb{A}$ называется *аффинным подпространством*, если множество векторов $(Q - P)$, $\forall Q, P \in \mathbb{B}$ образует подпространство $\mathbb{U} \subset \mathbb{V}$ в ассоциированном векторном пространстве. $\square$

Легко проверить, что две прямые линии параллельны тогда и только тогда, когда они не имеют общих точек.

Ясно, что каждое аффинное подпространство $\mathbb{B} \subset \mathbb{A}$ состоит из точек вида

$$\mathbb{B} := \{P + x : x \in \mathbb{U}\},$$

где P – произвольная точка из $\mathbb{B}$.

Пример 1.3.69. Прямая линия является одномерным аффинным подпространством в $\mathbb{A}$. $\square$

При аффинном преобразовании прямые линии переходят в прямые, причем пересекающиеся линии переходят в пересекающиеся, а параллельные – в параллельные.

Определение. Пусть $P_1, \dots, P_k$ – произвольные точки аффинного пространства, и числа $a_1, \dots, a_k$ удовлетворяют условию $\sum_{i=1}^k a_i = 1$. Определим сумму

$$P = \sum_{i=1}^k a_i P_i := P_0 + \sum_{i=1}^k a_i (P_i - P_0), \quad (1.63)$$

где P_0 – произвольная точка из $\mathbb{A}$. Поскольку $(P - P_0) \in \mathbb{V}$ и выражение (1.63) не зависит от выбора P_0 , то оно корректно определяет точку аффинного пространства $P \in \mathbb{A}$. Эта точка называется *барицентрической комбинацией* точек $P_1 \dots P_k$ с коэффициентами $a_1 \dots a_k$. $\square$

Пример 1.3.70. В евклидовом аффинном пространстве барицентрическая сумма

$$\sum_{i=1}^k \frac{m_i}{M} P_i, \quad M := \sum_{i=1}^k m_i$$

представляет собой положение центра масс системы масс m_i , расположенных в точках P_i . $\square$

Из определения (1.63) следует, что система

$$\{P_0, (P_1 - P_0), \dots, (P_n - P_0)\},$$

состоящая из точки P_0 и множества векторов $(P_i - P_0)$, образует систему аффинных координат тогда и только тогда, когда любая точка $P \in \mathbb{A}$ представима в виде барицентрической суммы

$$P = \sum_{i=1}^n x_i P_i, \quad \sum_{i=1}^n x_i = 1.$$

Если это условие выполнено, то множество точек $\{P_0, \dots, P_n\}$ называется *барицентрической* системой координат в аффинном пространстве, а числа $x_i \in \mathbb{R}$ – барицентрическими координатами точки $P \in \mathbb{A}$.

Определение. Пусть точки P_i , $i = 1, \dots, n$ имеют аффинные координаты

$$P_i = \{0, \dots, \underbrace{0}_{i-1}, 1, 0, \dots, 0\}.$$

Эти точки вместе с началом системы координат образуют барицентрическую систему координат в аффинном пространстве. Рассмотрим пересечение множества точек с барицентрическими координатами, не превосходящими единицы, $0 \leq x_i \leq 1$, с положительным 2^n -тантом. Это множество называется *стандартным $(n-1)$ -мерным симплексом*. В общем случае множество точек

$$\left\{ \sum_{i=1}^n x_i P_i : \sum_{i=1}^n x_i = 1, \quad 0 \leq x_i \leq 1 \right\} \quad (1.64)$$

называется *замкнутым симплексом* с вершинами в точках P_i в аффинном пространстве $\mathbb{A}$. Симплекс называется *вырожденным*, если векторы $(P_2 - P_1), \dots, (P_n - P_1)$ линейно зависимы. $\square$

Пример 1.3.71. Одномерный симплекс – это отрезок прямой, двумерный – равносторонний треугольник, трехмерный – тетраэдр (см. рис.1.3). $\square$

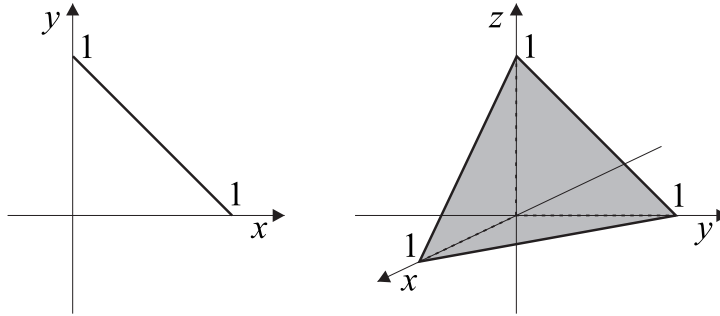


Рис. 1.3: Одномерный и двумерный симплексы.

Определим теперь параллельный перенос векторов в аффинном пространстве.

Определение. Рассмотрим произвольную кривую $\gamma = x(t)$, соединяющую две произвольно выбранные точки P и Q в аффинном пространстве. С каждой точкой аффинного пространства и, следовательно, с каждой точкой кривой γ ассоциировано векторное пространство $\mathbb{V}$. Будем говорить, что вектор $x \in \mathbb{V}$ *параллельно переносится* вдоль кривой γ , если его компоненты относительно некоторого фиксированного базиса в $\mathbb{V}$ не меняются при переходе от точки к точке. $\square$

Замечание. В данном определении важно различать точки аффинного и векторного пространств. $\square$

Очевидно, что параллельный перенос вектора из точки P в точку Q не зависит от кривой, соединяющей эти точки.

В настоящем и предыдущем разделах векторы обозначались жирным шрифтом. В дальнейшем мы упростим обозначения. Если у символа индекс присутствует, то это значит, что рассматривается соответствующая компонента вектора. Например, x^α обозначает компоненту вектора с индексом α . В то же время, если индекс опущен или используются фигурные скобки, то подразумевается весь вектор $x = \{x^\alpha\} = (x^1, \dots, x^n)$, и мы, как правило, не будем использовать жирный шрифт.

1.4 Отображения

Определение. Рассмотрим два множества $\mathbb{M}_1$ и $\mathbb{M}_2$ и их *отображение*

$$f : \mathbb{M}_1 \rightarrow \mathbb{M}_2 \quad \text{или} \quad \mathbb{M}_1 \xrightarrow{f} \mathbb{M}_2, \quad (1.65)$$

которое ставит каждому элементу $x \in \mathbb{M}_1$ единственный элемент $f(x) \in \mathbb{M}_2$. Отображение f называется также *функцией*, а переменная $x \in \mathbb{M}_1$ — *аргументом*. Вообще говоря, функция может отображать различные элементы из $\mathbb{M}_1$ в один и тот же элемент множества $\mathbb{M}_2$. Ситуация, когда один элемент из $\mathbb{M}_1$ отображается сразу в несколько элементов множества $\mathbb{M}_2$, не допускается. Пространство $\mathbb{M}_2$ в (1.65) называется *пространством-мишенью*. *Тождественное* отображение

$$\text{id}(\mathbb{M}) : \mathbb{M} \rightarrow \mathbb{M}$$

определяется равенством $\text{id}(x) = x$ для всех $x \in \mathbb{M}$. Отображение f называется *взаимно однозначным* или *инъективным*, если различные точки $x_1 \neq x_2$ из $\mathbb{M}_1$ отображаются в различные точки $f(x_1) \neq f(x_2)$ из $\mathbb{M}_2$. Отображение f является отображением $\mathbb{M}_1$ на $\mathbb{M}_2$ или *сюръективным*, если $f(\mathbb{M}_1) = \mathbb{M}_2$. Если отображение f инъективно и сюръективно одновременно, то оно называется *биективным отображением* или *биекцией*. Каждое отображение порождает некоторое подмножество в прямом произведении $\mathbb{M}_1 \times \mathbb{M}_2$,

$$\text{Gr}_f := \{x, f(x) \in \mathbb{M}_1 \times \mathbb{M}_2\},$$

которое называется *графиком* отображения f . □

Пример 1.4.1. Рассмотрим дифференцируемую функцию одной переменной $f(x) \in C^1(\mathbb{R})$. Если эта функция монотонно возрастает, $f' > 0$, или монотонно убывает, $f' < 0$, на интервале $x \in (p, q)$, то она задает биективное отображение интервалов $f : (p, q) \rightarrow (f(p), f(q))$. □

Пример 1.4.2. Преобразование координат с невырожденным якобианом является биекцией двух областей евклидова пространства (теорема 1.5.1). □

Сопоставление аргументу $x \in \mathbb{M}_1$ в (1.65) значения функции $f(x) \in \mathbb{M}_2$ будем обозначать ограниченной стрелкой:

$$f : \mathbb{M}_1 \ni x \mapsto f(x) \in \mathbb{M}_2.$$

Определение. Множество $f(\mathbb{M}_1) = \mathbb{U} \subset \mathbb{M}_2$, которое может и не совпадать со всем $\mathbb{M}_2$, называется *образом* множества $\mathbb{M}_1$ или *областью значений* функции f . Множество $\mathbb{M}_1$ при этом называется *прообразом* $\mathbb{U}$ или *областью определения* f и обозначается $\mathbb{M}_1 = f^{-1}(\mathbb{U})$. Вообще, прообразом произвольного подмножества $\mathbb{V} \subset \mathbb{U} \subset \mathbb{M}_2$ называется множество тех точек $\mathbb{W} \subset \mathbb{M}_1$, для которых $f(\mathbb{W}) = \mathbb{V}$. При этом символ f^{-1} является отображением

$$f^{-1} : \quad \mathbb{U} \ni \quad f(x) \mapsto x \quad \in \mathbb{M}_1$$

тогда и только тогда, когда отображение f является взаимно однозначным. В этом случае оно единственно и называется *обратным отображением*. $\square$

Для биективного отображения $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ всегда существует обратное отображение

$$f^{-1} : \quad \mathbb{M}_2 \rightarrow \mathbb{M}_1.$$

При этом отображение f^{-1} также биективно и выполнено равенство

$$(f^{-1})^{-1} = f.$$

Если отображение f не является взаимно однозначным, то прообразом одной точки $f^{-1}(x)$, где $x \in \mathbb{U}$, может быть несколько элементов из $\mathbb{M}_1$ или их бесконечное число. В этом случае символ f^{-1} не является отображением.

Пример 1.4.3. Рассмотрим три отображения:

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad g : \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{0\}, \quad h : \mathbb{R}_+ \rightarrow \mathbb{R}_+,$$

которые определены одним и тем же правилом $x \mapsto x^2$. Эти отображения различны: f не сюръективно, не инъективно; g сюръективно, но не инъективно; h биективно. Этот пример показывает, что задание области определения и области значений является существенной частью определения отображения (функции). $\square$

Определение. Пусть f – функция (1.65) и $\mathbb{U}_1$ – некоторое подмножество $\mathbb{M}_1$, тогда отображение

$$f|_{\mathbb{U}_1} : \quad \mathbb{U}_1 \rightarrow \mathbb{M}_2,$$

называется *сужением* функции (или отображения) на $\mathbb{U}_1 \subset \mathbb{M}_1$. Функция f является *продолжением* функции

$$g : \quad \mathbb{U}_1 \rightarrow \mathbb{M}_2,$$

заданной на некотором подмножестве $\mathbb{U}_1 \subset \mathbb{M}_1$, если ее сужение на $\mathbb{U}_1$ совпадает с g , $f|_{\mathbb{U}_1} = g$. $\square$

Ясно, что сужение функции единственно, а продолжение – нет.

Определение. *Композицией*, (*произведением* или *суперпозицией*) двух отображений $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ и $g : \mathbb{M}_2 \rightarrow \mathbb{M}_3$ называется отображение

$$g \circ f : \quad \mathbb{M}_1 \rightarrow \mathbb{M}_3,$$

которое состоит в последовательном применении этих отображений:

$$(g \circ f)(x) = g(f(x)).$$

$\square$

Замечание. Произведение определено не для всех отображений, а только для тех, у которых множество $\mathbb{M}_2$ в предыдущих обозначениях общее. Если рассматриваются отображения некоторого множества в себя, то для них композиция всегда определена. $\square$

Композиция отображений есть не что иное, как сложная функция.

Если задано отображение $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$, то

$$f \circ \text{id}(\mathbb{M}_1) = f \quad \text{и} \quad \text{id}(\mathbb{M}_2)f = f.$$

Для биективного отображения выполнены равенства:

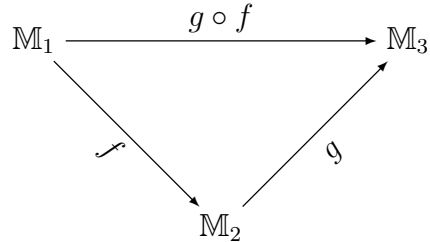
$$f^{-1} \circ f = \text{id}(\mathbb{M}_1) \quad \text{и} \quad f \circ f^{-1} = \text{id}(\mathbb{M}_2).$$

Если заданы два биективных отображения f и g и определена их композиция, т.е. область значений f совпадает с областью определения g , то обратное отображение для композиции существует и задается формулой

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

В дальнейшем вместо $g \circ f$ мы часто будем писать просто gf .

Произведение отображений играет огромную роль в математике, и его изображают в виде диаграммы



которая называется *коммутативной*. Это значит, что результат перехода от $\mathbb{M}_1$ к $\mathbb{M}_3$ не зависит от того, по какому пути мы движемся: то ли вдоль стрелки $g \circ f$, то ли последовательно вдоль стрелок f и g . Изображение произведений отображений в виде коммутативных диаграмм наглядно и полезно. Это будет продемонстрировано в следующем утверждении.

Предложение 1.4.1. Пусть определено произведение трех отображений:

$$f : \mathbb{M}_1 \rightarrow \mathbb{M}_2, \quad g : \mathbb{M}_2 \rightarrow \mathbb{M}_3, \quad h : \mathbb{M}_3 \rightarrow \mathbb{M}_4.$$

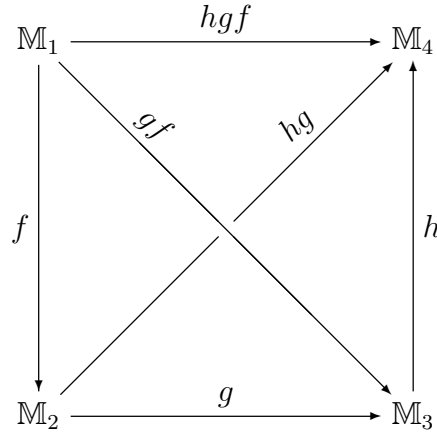
Тогда произведение этих отображений ассоциативно:

$$h(gf) = (hg)f$$

и его обозначают $hgf : \mathbb{M}_1 \rightarrow \mathbb{M}_4$.

Доказательство. Все необходимые рассуждения содержатся в следующей диаграм-

ме:



Для доказательства необходимо проследить куда отобразится произвольный элемент $x \in M_1$ под действием композиций $h(gf)$ и $(hg)f$. Формально доказательство выглядит следующим образом:

$$(h(gf))x = h((gf)x) = h(g(fx)) = (hg)(hx) = ((hg)f)x. \quad \square$$

Отображения конечных множеств обладают простыми свойствами.

Предложение 1.4.2. *Если отображение $f : M \rightarrow M$ конечного множества M в себя инъективно или сюръективно, то оно биективно.*

Доказательство. Пусть отображение f инъективно. Покажем, что оно является также сюръективным. Для произвольного элемента множества $x \in M$ определим последовательность

$$f^0(x) := x, \quad \text{и} \quad f^k(x) := \underbrace{f(\dots f(x) \dots)}_k = f(f^{k-1}(x)), \quad k = 1, 2, \dots$$

Из конечности множества M вытекает, что в этой последовательности есть повторения. Допустим, скажем, что $f^m(x) = f^n(x)$ при некоторых $m > n$. Если $n > 0$, то из равенства $f(f^{m-1}(x)) = f(f^{n-1}(x))$ и инъективности f следует $f^{m-1}(x) = f^{n-1}(x)$. Сократив отображение f n раз, приходим к равенству

$$f^{m-n}(x) = x.$$

В этом случае для любого $x \in M$ найдется такое $x' = f^{m-n-1}(x)$, что $f(x') = x$, т.е. отображение f сюръективно.

Второе утверждение предложения докажем от противного. Пусть отображение f сюръективно, но не инъективно. Тогда существует такой элемент $x \in M$, что его прообраз состоит не менее, чем из двух элементов. Пусть множество M содержит N элементов. Тогда f должно отобразить оставшиеся элементы, которых не более, чем $N - 2$, на $N - 1$ элемент. Для сюръективного отображения это невозможно. $\square$

В приведенном доказательстве существенна конечность множества M .

Пример 1.4.4. *Отображение следования, определенное для натуральных чисел сдвигом на единицу,*

$$\sigma : \mathbb{N} \ni n \mapsto n + 1 \in \mathbb{N},$$

является инъективным, но не сюръективным, поскольку единица не принадлежит образу $\sigma(\mathbb{N})$. $\square$

Определение. Мы говорим, что множества $\mathbb{M}_1$ и $\mathbb{M}_2$ имеют одинаковую *мощность*, если существует биективное отображение $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$. Множества, имеющие ту же мощность, что и множество натуральных чисел $\mathbb{N}$ называется *счетным*. $\square$

1.4.1 Отображения топологических пространств

При отображении топологических пространств важную роль играют отображения, согласованные с их топологией.

Определение. Отображение (1.65) двух топологических пространств называется *непрерывным* в точке $x \in \mathbb{M}_1$, если для каждой окрестности $\mathbb{U}_2$ точки $f(x) \in \mathbb{M}_2$ существует такая окрестность $\mathbb{U}_1$ точки x , что ее образ $f(\mathbb{U}_1)$ содержится в $\mathbb{U}_2$. Отображение называется непрерывным в области, если оно непрерывно в каждой точке этой области. $\square$

Замечание. В общем случае окрестность $\mathbb{U}_2 \subset \mathbb{M}$ может содержать точки, не лежащие в образе $f(\mathbb{M}_1)$ и образ $f(\mathbb{U}_1)$ совсем не обязан совпадать с $\mathbb{U}_2$. $\square$

Определение. Говорят, что два топологических пространства $\mathbb{M}_1$ и $\mathbb{M}_2$ *гомеоморфны*, если существуют взаимно обратные непрерывные отображения $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ и $g : \mathbb{M}_2 \rightarrow \mathbb{M}_1$ такие, что $g \circ f = \text{id}(\mathbb{M}_1)$ и $f \circ g = \text{id}(\mathbb{M}_2)$. Тогда отображения f и g называются *гомеоморфизмами*. $\square$

Ясно, что любой гомеоморфизм является биекцией. Обратное утверждение неверно.

Пример 1.4.5. Рассмотрим функцию

$$f(x) = \begin{cases} x, & x \leq -1 \text{ или } x \geq 1, \\ -x, & x \in (-1, 1), \end{cases}$$

показанную на рис.1.4. Эта функция задает отображение $\mathbb{R} \rightarrow \mathbb{R}$, которое биективно, но не задает гомеоморфизм. $\square$

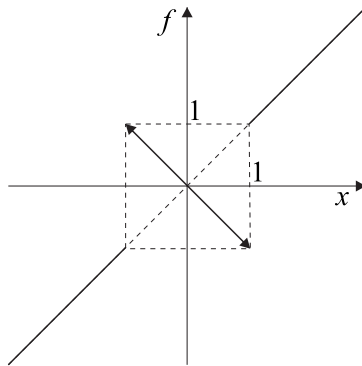


Рис. 1.4: Биективное отображение прямых, которое не является гомеоморфизмом.

Справедливы две фундаментальные теоремы, которые приведем без доказательства.

Теорема 1.4.1 (Инвариантность размерности). *Открытое подмножество в $\overline{\mathbb{R}}_+^m$ не может быть гомеоморфно никакому открытому подмножеству в $\overline{\mathbb{R}}_+^n$, если $m \neq n$.*

В этой теореме замкнутое полупространство $\overline{\mathbb{R}}_+^n$ можно заменить на все евклидово пространство $\mathbb{R}^n$.

Теорема 1.4.2 (Инвариантность края). *Пусть U и V – два открытых подмножества $\overline{\mathbb{R}}_+^n$ и f – гомеоморфизм U на V , тогда $f(U \cap \mathbb{R}^{n-1}) = V \cap \mathbb{R}^{n-1}$, где $\mathbb{R}^{n-1}$ – гиперплоскость (край), определяемая условием $x^n = 0$. Другими словами, при гомеоморфизмах точки края отображаются в край.*

Приведем пять критериев (необходимых и достаточных условий) непрерывности отображения, которые будут использоваться в дальнейшем.

Теорема 1.4.3. *Если отображение $f : M_1 \rightarrow M_2$ сюръективно, то оно непрерывно тогда и только тогда, когда выполнено одно из следующих условий:*

- 1) *образ любого замкнутого подмножества M_2 замкнут в M_1 ;*
- 2) *образ любого элемента базы топологии в M_2 открыт в M_1 ;*
- 3) $\forall U_1 \subset M_1$ справедливо включение $f(\overline{U_1}) \subset \overline{f(U_1)}$;
- 4) $\forall U_2 \subset M_2$ справедливо включение $f^{-1}(\overline{U_2}) \subset \overline{f^{-1}(U_2)}$;
- 5) $\forall U_2 \subset M_2$ справедливо включение $f^{-1}(\text{int } U_2) \subset \text{int } f^{-1}(U_2)$.

Доказательство. См., например, [20], теорема 1.4.1. □

Определение и критерии непрерывности отображений говорят о топологических свойствах прообраза, исходя из свойств образа. При этом нельзя сделать однозначного утверждения относительно свойств образа при заданных свойствах прообраза. Действительно, из критерия 3) следует, что образ замкнутого множества, вообще говоря, не является замкнутым. Критерий 5), записанный в виде

$$\text{int } U_2 \subset f(\text{int } f^{-1}(U_2)),$$

говорит о том, что образ открытого множества может не быть открытым.

Определение. Непрерывное отображение (1.65) называется *открытым* (замкнутым) если образ любого открытого (замкнутого) подмножества M_1 открыт (замкнут) в M_2 . □

Пример 1.4.6. Гомеоморфизм двух топологических пространств является открытым и замкнутым одновременно, т.е. открытое множество отображается в открытое, а замкнутое – в замкнутое. □

Пример 1.4.7. Приведем пример замкнутого, но не открытого непрерывного сюръективного отображения $f : \mathbb{R} \rightarrow \mathbb{R}$ (см. рис.1.5), которое задано следующим образом

$$f(x) = \begin{cases} x + 1, & x \leq -2, \\ -1, & -2 \leq x \leq -1, \\ x, & -1 \leq x \leq 1, \\ 1, & 1 \leq x \leq 2, \\ x - 1, & 2 \leq x. \end{cases}$$

При таком отображении открытый интервал (a, b) , где $a \in (-2, -1)$ и $b \in (1, 2)$ отображается в замкнутый отрезок $[-1, 1]$. □

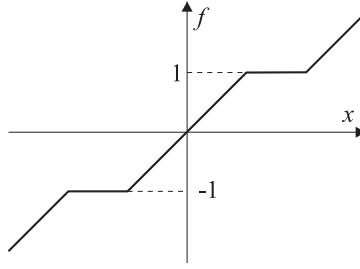


Рис. 1.5: Пример замкнутого, но не открытого непрерывного сюръективного отображения.

Определение. Рассмотрим отображение $f : M_1 \rightarrow M_2$ топологического пространства M_1 в некоторое множество M_2 . Если отображение f является сюръективным, т.е. отображением на, то на множестве M_2 можно определить *топологию идентификации* или *фактортопологию*. А именно, назовем подмножество $U \subset M_2$ открытым, если его прообраз $f^{-1}(U) \subset M_1$ открыт в M_1 . Это самая тонкая топология на M_2 , в которой отображение f является непрерывным. $\square$

Пример 1.4.8. Пусть M – топологическое пространство, на котором задано некоторое отношение эквивалентности $x_1 \sim x_2$. Пусть $M/\sim$ – множество классов эквивалентности, и $f : M \rightarrow M/\sim$ – каноническая проекция элемента $x \in M$ на его класс эквивалентности. Тогда отображение f сюръективно и на фактор пространстве $M/\sim$ можно ввести топологию идентификации. $\square$

На самом деле, рассмотренный пример является общим. Если задано сюръективное отображение $f : M_1 \rightarrow M_2$, то на множестве M_1 всегда можно определить отношение эквивалентности: будем считать точки $x_1, x_2 \in M_1$ эквивалентными, если $f(x_1) = f(x_2)$. Другими словами, в множестве M_1 мы идентифицируем те точки, которые отображаются в одну точку пространства-мишени M_2 . Этим объясняется название “топология идентификации”.

Определение. Если некоторое свойство P топологического пространства сохраняется при гомеоморфизме, то оно называется *топологическим инвариантом* или *свойством*. Другими словами, топологическое пространство M обладает свойством P в том и только в том случае, если этим свойством обладает любое другое гомеоморфное ему топологическое пространство. $\square$

Так как гомеоморфизм устанавливает взаимно однозначное соответствие не только между точками, но и между открытыми множествами, то любое свойство, определенное в терминах открытых множеств, является топологическим инвариантом.

Пример 1.4.9. Число связных компонент топологического пространства является топологическим инвариантом. $\square$

Если отображение (1.65) непрерывно, но не является гомеоморфизмом, то число связных компонент может уменьшаться, но не увеличиваться.

Пример 1.4.10. Если пространство $M_1 = I_1 \cup I_2$ является объединением двух непересекающихся открытых интервалов I_1 и I_2 , то его можно непрерывно отобразить на произвольный открытый интервал I_3 . При этом $I_1 \rightarrow I_3$ и $I_2 \rightarrow I_3$, т.е. каждый из

интервалов $\mathbb{I}_1$ и $\mathbb{I}_2$ отображается на весь интервал $\mathbb{I}_3$. Обратное утверждение неверно. Один интервал невозможно непрерывно отобразить на два. Действительно, если существует непрерывное отображение

$$g : \mathbb{I}_3 \rightarrow \mathbb{I}_1 \cup \mathbb{I}_2,$$

то в силу непрерывности g прообразы $\mathbb{I}_1$ и $\mathbb{I}_2$ являются открытыми и замкнутыми непустыми подмножествами в $\mathbb{I}_3$ одновременно. Поскольку интервал $\mathbb{I}_3$ связан, то $\mathbb{I}_3 = g^{-1}(\mathbb{I}_1) = g^{-1}(\mathbb{I}_2)$. Это означает, что каждой точке интервала $\mathbb{I}_3$ соответствует две точки в объединении $\mathbb{I}_1 \cup \mathbb{I}_2$, что противоречит определению отображения (1.65). $\square$

Определение. *Компактификацией* топологического пространства $\mathbb{M}$ называется пара $(\mathbb{U}, f)$, где $\mathbb{U}$ – компактное топологическое пространство, а f – гомеоморфизм $\mathbb{M}$ на всюду плотное подмножество в $\mathbb{U}$. $\square$

Пример 1.4.11. Отображение евклидовой плоскости на сферу, $f : \mathbb{R}^2 \rightarrow \mathbb{S}^2$, которое задается стереографической проекцией (см. раздел 29.1) является компактификацией. При этом все бесконечно удаленные точки евклидовой плоскости отображаются на одну точку сферы – северный полюс. $\square$

Пример 1.4.12. Добавим к евклидову пространству $\mathbb{R}^n$ бесконечно удаленную точку ∞ и объявим ее окрестностями множества вида $(\mathbb{R}^n \setminus \mathbb{B}) \cup \infty$ для всех ограниченных замкнутых подмножеств $\mathbb{B} \subset \mathbb{R}^n$. Тогда объединение $\mathbb{R}^n \cup \infty$ будет компактным пространством, гомеоморфным сфере $\mathbb{S}^n$. В этом примере пара $((\mathbb{R}^n \cup \infty), f)$, где отображение $f : \mathbb{R}^n \rightarrow \mathbb{R}^n \cup \infty$ имеет вид

$$f : \mathbb{R}^n \ni x \mapsto (x, \infty) \in \mathbb{R}^n \cup \infty,$$

является компактификацией евклидова пространства. $\square$

Одно и то же топологическое пространство может иметь много различных компактификаций.

Пример 1.4.13. Выше мы показали, что сфера $\mathbb{S}^n$ является компактификацией евклидова пространства $\mathbb{R}^n$. Другой компактификацией $\mathbb{R}^n$ является проективное пространство $\mathbb{P}\mathbb{R}^n$. В этом случае к евклидову пространству добавляется не одна точка, а сфера $\mathbb{S}^{n-1}$. $\square$

Отображения хаусдорфовых топологических пространств. Поскольку многообразия, по определению, являются хаусдорфовыми топологическими пространствами, то отображения хаусдорфовых пространств будут играть в дальнейшем важную роль.

Предложение 1.4.3. Пусть $\mathbb{M}_1$ и $\mathbb{M}_2$ – хаусдорфовы топологические пространства. Функция $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ непрерывна тогда и только тогда, когда для любой сходящейся направленности $S = \{x_\sigma : \sigma \in \Sigma\}$ в $\mathbb{M}_1$ такой, что $x = \lim_{\sigma \in \Sigma} x_\sigma$, направленность $\{f(x_\sigma) : \sigma \in \Sigma\}$ сходится к точке $f(x)$ в $\mathbb{M}_2$.

Доказательство. См., например, [6], утверждение 2.7. $\square$

Доказательство следующих двух теорем и следствий содержится, например, в [20], теоремы 3.1.10 и 3.1.11.

Теорема 1.4.4. Если компактное пространство M_1 непрерывно отображается на пространство M_2 , то M_2 компактно.

Следствие. Если $f : M_1 \rightarrow M_2$ – непрерывное отображение компактного пространства M_1 в хаусдорфово пространство M_2 , то $\overline{f(U)} = f(\overline{U})$ для каждого подмножества $U \subset M_1$.

Следствие. Каждое непрерывное отображение компактного пространства в хаусдорфово пространство замкнуто.

Теорема 1.4.5. Каждое непрерывное взаимно однозначное отображение компактного пространства на хаусдорфово пространство является гомеоморфизмом.

Сформулируем три теоремы, доказательство которых можно найти, например, в [19], теоремы 8.7, 8.9, 8.10 и 8.11.

Теорема 1.4.6. Компактное подмножество U хаусдорфова топологического пространства M хаусдорфово и замкнуто.

Теорема 1.4.7. Пусть M_1 и M_2 – топологические пространства. Тогда M_1 и M_2 хаусдорфовы в том и только в том случае, если произведение $M_1 \times M_2$ хаусдорфово.

Хотя подпространства и произведения хаусдорфовых пространств хаусдорфовы, факторпространство хаусдорфова пространства, вообще говоря, нехаусдорфово.

Пример 1.4.14. Рассмотрим вещественную прямую $\mathbb{R}$ и открытый интервал $(0, 1)$. Введем отношение эквивалентности $\sim$ на $\mathbb{R}$, при котором $x \sim x'$ тогда и только тогда, когда $x = x'$ или $x, x' \in (0, 1)$. Грубо говоря, фактор пространство $X/\sim$ – это вещественная прямая $\mathbb{R}$ со стянутым в точку интервалом $(0, 1)$. Если снабдить $\mathbb{R}/\sim$ фактортопологией относительно естественной проекции $\pi : \mathbb{R} \ni x \mapsto \tilde{x} \in \mathbb{R}/\sim$, то прообраз точки $\tilde{x}_0$, где $x_0 \in (0, 1)$ есть интервал $(0, 1)$, который открыт в $\mathbb{R}$. Следовательно, точка $\tilde{x}_0$ открыта в $\mathbb{R}/\sim$. Таким образом фактор пространство $M/\sim$ представляет собой объединение двух лучей $(-\infty, 0]$, $[1, \infty)$ и точки $\tilde{x}_0$, которая открыта. Любые окрестности точек 0 и 1 содержат точку $\tilde{x}_0$ и, потому, пересекаются. Следовательно, факторпространство $\mathbb{R}/\sim$ нехаусдорфово. Подчеркнем, что в рассматриваемом примере фактор пространство $\mathbb{R}/\sim$ снабжено фактор топологией (топологией идентификации), которая отличается от топологии, индуцируемой вложением в $\mathbb{R}^2$. $\square$

В рассмотренном примере факторпространство $\mathbb{R}/\sim$ не является многообразием, хотя проекция π непрерывна. Для того, чтобы обеспечить хаусдорфовость факторпространства N хаусдорфова пространства M по некоторому отношению эквивалентности необходимо наложить дополнительные условия. Достаточное условие дает следующая

Теорема 1.4.8. Пусть N – факторпространство топологического пространства M , определенное при помощи сюръективного отображения $f : M \rightarrow N$. Если пространство M компактно и хаусдорфово, а отображение f замкнуто, то N компактно и хаусдорфово.

Следствие. Если M – компактное хаусдорфово пространство, на котором действует конечная группа преобразований G , то M/G – компактное хаусдорфово пространство. $\square$

Пример 1.4.15. Проективное пространство $\mathbb{RP}^n$ получается из сферы $\mathbb{S}^n$ отождествлением диаметрально противоположных точек, $\mathbb{RP}^n = \mathbb{S}^n/\mathbb{Z}_2$. Это есть компактное хаусдорфово пространство. $\square$

Чтобы сформулировать другое следствие теоремы 1.4.8, рассмотрим пространство $\mathbb{M}$ вместе с его подмножеством $\mathbb{U} \subset \mathbb{M}$. Обозначим через $\mathbb{M}/\mathbb{U}$ фактор пространство $\mathbb{M}/\sim$, где $\sim$ – отношение эквивалентности на $\mathbb{M}$, при котором $x \sim x'$ тогда и только тогда, когда $x = x'$ или $x, x' \in \mathbb{U}$.

Следствие. Если $\mathbb{M}$ – компактное хаусдорфово пространство и $\mathbb{U}$ – его замкнутое подмножество, то $\mathbb{M}/\mathbb{U}$ – компактное хаусдорфово пространство. $\square$

Замечание. В примере 1.4.14 подмножество $(0, 1) \subset \mathbb{R}$ открыто, и фактор пространство $\mathbb{R}/\sim$ не является хаусдорфовым. $\square$

1.5 Преобразования координат

Важным примером отображений являются преобразования координат, которые рассмотрены в настоящем разделе. Евклидово пространство $\mathbb{R}^n$ было определено в разделе 1.3.2 как топологическое произведение n прямых. При этом каждая точка $x \in \mathbb{R}^n$ задается набором n вещественных чисел:

$$x = (x^1, \dots, x^n) = \{x^\alpha\}, \quad \alpha = 1, \dots, n,$$

– декартовых координат, которые покрывают все евклидово пространство. Рассмотрим n дифференцируемых функций (преобразование координат)

$$x^{\alpha'} = x^{\alpha'}(x) \in \mathcal{C}^1(\mathbb{U}), \quad \alpha' = 1, \dots, n, \quad (1.66)$$

от декартовых координат x^α , заданных в некоторой области $\mathbb{U} \subset \mathbb{R}^n$. В общем случае эти функции отображают область $\mathbb{U} \subset \mathbb{R}^n$ на некоторое множество точек $\mathbb{U}' \subset \mathbb{R}^n$ – образ отображения. В наших обозначениях запись $x^{\alpha'}$ эквивалентна записи $x'^{\alpha'}$. При этом штрихованные индексы пробегают те же значения, что и нештрихованные $\alpha, \alpha' = 1, \dots, n$. Такая запись удобнее, т.к. позволяет во многих случаях опускать букву x . В сокращенной записи мы опускаем индексы:

$$\mathbb{U} \ni \quad x \mapsto x'(x) \quad \in \mathbb{U}'. \quad (1.67)$$

Определение. Матрица, составленная из частных производных,

$$J_\alpha^{\alpha'} = \frac{\partial x^{\alpha'}}{\partial x^\alpha} = \partial_\alpha x^{\alpha'} \quad (1.68)$$

называется *матрицей Якоби*. Здесь индекс α считается первым, а индекс α' – вторым. Определитель матрицы Якоби называется *якобианом*,

$$J = \frac{\partial(x^{\alpha'})}{\partial(x^\alpha)} := \det \left(\frac{\partial x^{\alpha'}}{\partial x^\alpha} \right). \quad \square \quad (1.69)$$

Из курса математического анализа известно следующее утверждение.

Теорема 1.5.1. Пусть функции (1.66) задают дифференцируемое отображение класса C^k открытого множества $U \subset \mathbb{R}^n$ в $\mathbb{R}^n$. Если якобиан этого отображения не обращается в нуль в точке $x_0 \in U$, то существуют такие окрестности $U_0 \subset U$ и $U'_0 \subset \mathbb{R}^n$ точек x_0 и $x'_0 = x'_0(x_0)$, что функции $x^{\alpha'}(x)$ задают взаимно однозначное отображение окрестности U_0 на U'_0 (биекцию), и обратное ему отображение дифференцируемо и того же класса гладкости C^k в области U'_0 . В частности, возможно $k = \infty$.

Доказательство. См., например, [24], §41.8, теорема 7. $\square$

Следствие. Пусть $x'(x)$ – дифференцируемое отображение открытого множества $U \subset \mathbb{R}^n$ в $\mathbb{R}^n$. Если якобиан этого отображения отличен от нуля на всем U , то образ этого множества U' также является открытым множеством. В частности, окрестность точки x отображается в окрестность точки x' . $\square$

Пусть задана последовательность биекций:

$$U \xrightarrow{x'(x)} U' \xrightarrow{x''(x')} U'' \rightarrow \dots$$

Мы говорим, что точка $x \in U \subset \mathbb{R}^n$ евклидова пространства задается n вещественными числами

$$x = \{x^\alpha\} = \{x^{\alpha'}\} = \{x^{\alpha''}\} = \dots,$$

которые являются координатами одной и той же точки в различных системах координат. Теперь мы можем расширить понятие евклидова пространства, допустив произвольные замены координат. Тогда мы считаем, что евклидово пространство $\mathbb{R}^n$ покрывается некоторыми областями (картами)

$$\mathbb{R}^n = \bigcup_i U_i,$$

на каждой из которых выбрана своя система координат. Рассмотрим две пересекающиеся карты U_i и U_j в $\mathbb{R}^n$. В частном случае одна или обе области могут совпадать со всем евклидовым пространством. Обозначим координаты точек, принадлежащих U_i и U_j , соответственно, через x^α и $x^{\alpha'}$. Тогда в области пересечения карт $U_i \cap U_j$ координаты точек будут связаны некоторым преобразованием вида (1.66). По этой причине функции (1.66) называются *функциями перехода* к новой системе координат или *функциями склейки*. Заметим, что их одновременно можно рассматривать и как функции, осуществляющие склейку двух областей и как замену координат на пересечении $U_i \cap U_j$.

Замечание. В дифференциальной геометрии такие объекты как векторные поля, метрика, связность и другие геометрические структуры вводятся инвариантным образом, независимо от выбранной системы координат. В приложениях выбор той или иной системы координат диктуется симметрией задачи, и эти координаты, как правило, не покрывают все евклидово пространство. Например, сферические и цилиндрические координаты определены всюду в $\mathbb{R}^3$ за исключением оси z . $\square$

Функции (1.66) осуществляют замену или *преобразование координат*. Это преобразование в зависимости от обстоятельств можно рассматривать как *пассивное* или *активное*. При пассивном рассмотрении считается, что одной и той же точке области соответствуют различные координаты в различных системах отсчета. Именно так мы

и рассматривали преобразование координат до сих пор. Во втором случае считается, что заданная точка евклидова пространства с координатами x^α в результате некоторого преобразования или деформирования области занимает новое положение с координатами $x^{\alpha'}$ относительно старой координатной системы. Такая точка зрения принята в теории групп преобразований.

Пример 1.5.1. Вращение евклидова пространства относительно начала в данной декартовой системе координат обычно рассматривают как активное. С математической точки зрения оба подхода равноправны, т.к. определяются одним и тем же набором функций (1.66), а деление преобразований на активные и пассивные зависит от физической интерпретации и традиций. $\square$

Обычно от функций перехода (1.66) требуется, чтобы они и их обратные $x^\alpha = x^\alpha(x')$ были достаточно гладкими. В частности, непрерывность функций перехода означает гомеоморфность отображения (1.66). В дальнейшем, если не оговорено противное, мы рассматриваем преобразования координат класса $\mathcal{C}^\infty$. В этом случае гомеоморфизм (1.67) называется *диффеоморфизмом*.

Отличие от нуля якобиана (1.69) означает, что функции (1.66), определяющие преобразования координат, функционально независимы.

Определение. Набор функций $u^i(x)$, $i = 1, \dots, N$, в некоторой области $\mathbb{U}$ называется *функционально зависимым*, если на любом компакте из $\mathbb{U}$ существует функция от N переменных $F(u)$, определенная в области значений u^i , которая непрерывна вместе со всеми частными производными первого порядка, не равна тождественно нулю ни в какой подобласти и для которой выполнено соотношение $F(u(x)) = 0$. $\square$

Дифференцируя уравнение $F(u(x)) = 0$ по x^α , получим равенство

$$\frac{\partial F}{\partial u^i} \frac{\partial u^i}{\partial x^\alpha} = 0.$$

Поскольку среди функций $\partial F / \partial u^i$ по крайней мере одна отлична от нуля, то для функциональной зависимости необходимо, чтобы ранг матрицы $\partial u^i / \partial x^\alpha$ был меньше N . Это условие является и достаточным. В частности, при $N > n$ функции всегда будут функционально зависимы. Тем самым можно дать эквивалентное

Определение. N достаточно гладких функций, для которых ранг матрицы $\partial u^i / \partial x^\alpha$ равен N в каждой точке области $\mathbb{U} \subset \mathbb{R}^n$ называются *функционально независимыми* в этой области. $\square$

Если ранг матрицы $\partial_\alpha u^i$ меньше N , то функции $u^i(x)$ функционально зависимы, и уравнение $F(u) = 0$ согласно теореме о неявной функции локально разрешимо, и одну из функций можно выразить через остальные в некоторой области.

Таким образом, n достаточно гладких функционально независимых функций (1.66) задают преобразование координат в некоторой области $\mathbb{U} \subset \mathbb{R}^n$.

Преобразование координат (1.66) называется *невырожденным* в некоторой точке или области, если якобиан преобразования отличен от нуля в этой точке или области, соответственно. Если преобразование координат невырождено в некоторой области $\mathbb{U}$, то согласно теореме 1.5.1 в достаточно малой окрестности произвольной точки $x \in \mathbb{U}$ это преобразование является взаимно однозначным. Для невырожденных преобразований координат локально существует обратное преобразование

$$x^{\alpha'} \mapsto x^\alpha = x^\alpha(x'),$$

при этом якобиан обратного преобразования равен

$$J^{-1} = \det \left(\frac{\partial x^\alpha}{\partial x^{\alpha'}} \right).$$

Если якобиан преобразования обращается в нуль в некоторой точке, то в этой точке преобразование координат является вырожденным, и обратного преобразования не существует.

Совершим два преобразования координат $x \mapsto x' \mapsto x''$. Тогда координаты точки x можно выразить, как в системе координат x' , так и в x'' . При этом, по определению, справедливы равенства $x^\alpha(x'') = x^\alpha(x'(x''))$. Отсюда следует, что каждая координатная функция x^α при преобразовании координат $x' \rightarrow x''$ ведет себя, как скалярное поле (см. раздел 2.5). Верно также и обратное утверждение. Если в некоторой области евклидова пространства $U \subset \mathbb{R}^n$ задано n функций (скалярных полей) $f^a(x)$, $a = 1, \dots, n$, таких, что $\det(\partial f^a / \partial x^\alpha) \neq 0$, то их всегда можно выбрать в качестве новой системы координат на U .

Обозначим множество невырожденных преобразований координат в $\mathbb{R}^n$ через $\mathbb{G}(\mathbb{R}^n)$. При этом будем рассматривать преобразования координат как активные, т.е. отображающие области в одном фиксированном евклидовом пространстве. Для заданного преобразования координат $f = x'(x)$ будем писать $f \in \mathbb{G}(\mathbb{R}^n)$. Эти преобразования согласованы с топологией в $\mathbb{R}^n$.

Определение. Множество преобразований $\mathbb{G}(\mathbb{R}^n)$ топологического пространства $\mathbb{R}^n$ называется *псевдогруппой преобразований*, если выполнены следующие условия:

- 1) каждое преобразование $f \in \mathbb{G}(\mathbb{R}^n)$ является гомеоморфизмом открытого множества U_i (область определения f) на другое открытое множество $f(U_i) = U_j$ (область значений f);
- 2) если $f \in \mathbb{G}(\mathbb{R}^n)$, то сужение f на произвольное открытое подмножество области определения f также принадлежит $\mathbb{G}(\mathbb{R}^n)$;
- 3) пусть множество U есть объединение произвольного числа открытых множеств, $U = \bigcup_i U_i$, тогда гомеоморфизм f множества U принадлежит $\mathbb{G}(\mathbb{R}^n)$, если сужение f на U_i принадлежит $\mathbb{G}(\mathbb{R}^n)$ для каждого i ;
- 4) тождественное преобразование каждого открытого множества $U \subset \mathbb{R}^n$ принадлежит $\mathbb{G}(\mathbb{R}^n)$;
- 5) для каждого $f \in \mathbb{G}(\mathbb{R}^n)$ существует обратное преобразование $f^{-1} \in \mathbb{G}(\mathbb{R}^n)$;
- 6) если f_1 есть гомеоморфизм $U_1 \rightarrow V_1$, а f_2 – гомеоморфизм $U_2 \rightarrow V_2$, и пересечение $V_1 \cap U_2$ не пусто, то гомеоморфизм $f_2 f_1$ множества $f_1^{-1}(V_1 \cap U_2)$ на множество $f_2(V_1 \cap U_2)$ также принадлежит $\mathbb{G}(\mathbb{R}^n)$. $\square$

Замечание. Первые три свойства являются топологическими, а остальные – групповыми, если под композицией преобразований понимать их последовательное выполнение в тех точках, где они определены. В общем случае преобразования координат группы не образуют, т.к. при композиции двух преобразований необходимо следить за их областью определений и значений. Например, если область значений преобразования f_1 не имеет общих точек с областью определения f_2 , то композиция $f_2 \circ f_1$ не определена. Следовательно, преобразования координат группы не образуют, т.к. закон композиции в группе должен быть определен для всех элементов группы. Тем не менее часто употребляется термин группа преобразований координат. Для того, чтобы преобразования координат действительно образовали группу необходимо предположить, что области определений и значений $U \subset \mathbb{R}^n$ для всех

$f \in \mathbb{G}(\mathbb{R}^n)$ совпадают и отказаться от условия 2). Эта группа бесконечномерна, т.к. параметризуется n функциями от n переменных. $\square$

Отображение f областей евклидова пространства называется *диффеоморфизмом*. Если все диффеоморфизмы $f \in \mathbb{G}(\mathbb{R}^n)$ одного класса гладкости $\mathcal{C}^k$, то будем писать $f \in \text{diff}^k(\mathbb{R}^n)$. Для гладкой псевдогруппы преобразований координат индекс обычно опускают $\text{diff}(\mathbb{R}^n) := \text{diff}^\infty(\mathbb{R}^n)$. Ее часто называют просто *группой диффеоморфизмов* евклидова пространства $\mathbb{R}^n$. Очевидно, что $\text{diff}^l(\mathbb{R}^n)$ – подпсевдогруппа $\text{diff}^k(\mathbb{R}^n)$ при $l > k$.

Множество преобразований координат с положительным определителем $\text{diff}_0^k(\mathbb{R}^n)$ также является подпсевдогруппой: $\text{diff}_0^k(\mathbb{R}^n) \subset \text{diff}^k(\mathbb{R}^n)$. Все невырожденные преобразования координат можно разделить на два класса: с положительным, $\text{diff}_0^k(\mathbb{R}^n)$, и отрицательным, $\text{diff}^k(\mathbb{R}^n) \setminus \text{diff}_0^k(\mathbb{R}^n)$, якобианом. Говорят, что координатный базис касательного пространства (см. раздел 2.6) имеет одинаковую *ориентацию* в двух системах координат, если якобиан соответствующего преобразования положителен. Если якобиан преобразования отрицателен, то при переходе между системами координат базис касательного пространства, по определению, меняет ориентацию.

1.6 Группа двумерных вращений $\mathbb{O}(2)$

Рассмотрим евклидову плоскость $\mathbb{R}^2$ с декартовыми координатами $\{x, y\} = \{x^\alpha\}$, $\alpha = 1, 2$, и евклидовой метрикой $g_{\alpha\beta} = \delta_{\alpha\beta}$.

Определение. Невырожденные линейные однородные преобразования координат,

$$x^\alpha \mapsto x'^\alpha = x^\beta S_\beta^\alpha, \quad \det S_\beta^\alpha \neq 0,$$

относительно которых евклидова метрика остается инвариантной,

$$\delta_{\alpha\beta} = S_\alpha^\gamma S_\beta^\epsilon \delta_{\gamma\epsilon}, \quad (1.70)$$

образуют группу, которая называется *группой двумерных вращений* и обозначается $\mathbb{O}(2)$. Квадратные матрицы, удовлетворяющие условию (1.70), которое можно переписать в матричном виде

$$SS^T = S^T S = 1,$$

где S^T обозначает транспонированную матрицу, называются *ортогональными*. $\square$

Группа $\mathbb{O}(2)$ является группой ортогональных 2×2 матриц с обычным правилом умножения матриц. Вычисляя определитель левой и правой части уравнения (1.70) получим, что определитель матриц S по модулю равен единице,

$$\det S = \pm 1.$$

Предложение 1.6.1. Любое решение уравнения (1.70) принадлежит одному из двух классов:

$$S_+(\omega) := \begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix}, \quad \det S_+ = 1, \quad (1.71)$$

$$S_-(\omega) := \begin{pmatrix} \cos \omega & \sin \omega \\ \sin \omega & -\cos \omega \end{pmatrix}, \quad \det S_- = -1, \quad (1.72)$$

при некотором угле вращения $\omega \in [0, 2\pi]$. При этом точки $\omega = 0$ и $\omega = 2\pi$ отождествляются. Следовательно, параметр ω лежит на окружности, $\omega \in \mathbb{S}^1$.

Доказательство. Уравнение (1.70) – это система трех квадратных уравнений:

$$\begin{aligned}(S_\alpha^1)^2 + (S_\alpha^2)^2 &= 1, & \alpha &= 1, 2, \\ S_1^1 S_2^1 + S_1^2 S_2^2 &= 0.\end{aligned}$$

Перебор всех возможностей показывает, что с точностью до переопределения угла поворота формулы (1.71), (1.72) дают общее решение. $\square$

Для краткости обозначим множество матриц вида (1.71), (1.72) теми же, но ажурными буквами:

$$\begin{aligned}\mathbb{S}_+ &:= \{S_+(\omega) : \omega \in \mathbb{S}^1\}, \\ \mathbb{S}_- &:= \{S_-(\omega) : \omega \in \mathbb{S}^1\}.\end{aligned}$$

Тогда $\mathbb{O}(2) = \mathbb{S}_+ \cup \mathbb{S}_-$, и справедлива следующая таблица умножения:

$$\mathbb{S}_+ \mathbb{S}_+ = \mathbb{S}_+, \quad \mathbb{S}_+ \mathbb{S}_- = \mathbb{S}_-, \quad \mathbb{S}_- \mathbb{S}_+ = \mathbb{S}_-, \quad \mathbb{S}_- \mathbb{S}_- = \mathbb{S}_+.$$

Обратные матрицы для $\mathbb{S}_+$ имеют вид

$$S_+^{-1}(\omega) = S_+(-\omega) \tag{1.73}$$

и соответствует вращению на тот же угол ω , но в противоположном направлении. При этом единичная матрица $\mathbb{1} = S_+(0) \in \mathbb{S}_+$.

Таким образом, группа двумерных вращений является однопараметрической компактной группой Ли, состоящей из двух компонент. Матрицы $\mathbb{S}_+$ с единичным определителем образуют нормальную подгруппу группы вращений, которая называется группой *собственных* вращений и обозначается $\mathbb{SO}(2) = \mathbb{S}_+$. Эти матрицы образуют связную компоненту единицы. В этом нетрудно убедиться, заменив параметр ω на $t\omega$ в (1.71). Тогда мы получим непрерывное семейство ортогональных матриц, связывающее заданную матрицу при $t = 1$ с единичной матрицей при $t = 0$.

Отметим, что матрица $-\mathbb{1}$ также принадлежит связной компоненте единицы, $-\mathbb{1} = S_+(\pi/2)$.

Матрицы несобственных вращений $\mathbb{S}_-$ группы не образуют, т.к. не содержат, например, единицы. Их можно представить в виде произведения двух матриц

$$S_- = S_+ P, \tag{1.74}$$

где

$$P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = S_-(0), \quad \det P = -1, \tag{1.75}$$

– *оператор четности*, который отражает ось ординат,

$$P : (x, y) \mapsto (x, -y).$$

При этом меняется ориентация декартовых координат.

Таким образом, матрицы с положительным и отрицательным определителем образуют две связные компоненты группы вращений $\mathbb{O}(2)$. Связная компонента единицы образует группу собственных вращений $\mathbb{SO}(2)$, а оператор четности определяет диффеоморфизм между компонентами.

Легко проверить, что группа собственных вращений $\mathbb{SO}(2)$ является абелевой. В то же время полная группа вращений $\mathbb{O}(2)$ неабелева. В этом просто убедиться, вычислив коммутатор

$$[S_-(\omega_1), S_-(\omega_2)] = 2 \begin{pmatrix} 0 & \sin(\omega_1 - \omega_2) \\ -\sin(\omega_1 - \omega_2) & 0 \end{pmatrix},$$

который не лежит ни в $\mathbb{S}_+$, ни в $\mathbb{S}_-$.

Элементы группы $\mathbb{SO}(2)$ можно представить в виде

$$S_{+\alpha}{}^\beta = (e^{\omega\varepsilon})_\alpha{}^\beta = \delta_\alpha^\beta \cos \omega + \varepsilon_\alpha{}^\beta \sin \omega, \quad (1.76)$$

где экспонента определена с помощью ряда Тейлора, и $\varepsilon_{\alpha\beta}$ — полностью антисимметричный тензор второго ранга (см. приложение 37.4).

Матрица (1.76) задает экспоненциальное отображение элементов алгебры Ли $\omega \in \mathfrak{so}(2) \approx \mathbb{R}$ на группу вращений $S \in \mathbb{SO}(2)$. Экспоненциальное отображение определено для всех элементов алгебры и в данном случае сюръективно. При этом элементы алгебры ω и $\omega + 2\pi t$, $t \in \mathbb{Z}$, отображаются в один и тот же элемент группы. Это отображение является накрытием (см. главу 11).

Подгруппа $\mathbb{SO}(2) \subset \mathbb{O}(2)$ является нормальной подгруппой. Следовательно, фактор пространство $\mathbb{O}(2)/\mathbb{SO}(2)$ является группой, состоящей из двух элементов

$$\mathbb{Z}_2 \simeq \frac{\mathbb{O}(2)}{\mathbb{SO}(2)},$$

где $\mathbb{Z}_2 := \{1, -1\}$ — группа из двух элементов по умножению.

Матрицы (1.71), (1.72) можно рассматривать как представление группы вращений $\mathbb{O}(2)$ в двумерном векторном пространстве. Оно неприводимо и называется *фундаментальным* или *векторным* представлением.

Бесконечно малое вращение на угол $\delta\omega \ll 1$ в линейном приближении имеет вид

$$S_{+\alpha}{}^\beta \approx \delta_\alpha^\beta + \delta\omega L_\alpha{}^\beta, \quad (1.77)$$

где $L_\alpha{}^\beta = \varepsilon_\alpha{}^\beta$ — генератор (вектор базиса алгебры Ли) группы $\mathbb{SO}(2)$ в фундаментальном представлении.

Вращения евклидовой плоскости показаны на рис. 1.6, а. Орбитами действия и связной компоненты единицы группы $\mathbb{SO}(2)$, и полной группы $\mathbb{O}(2)$ являются концентрические окружности,

$$x^2 + y^2 = r^2, \quad r \in \mathbb{R}_+, \quad (1.78)$$

с центром в начале координат. Само начало координат является неподвижной точкой относительно действия группы вращений.

Определим *полярные координаты* на евклидовой плоскости

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad (1.79)$$

где радиус точки r и полярный угол φ определены в интервалах: $r \in (0, \infty)$, $\varphi \in [0, 2\pi)$. Обратное преобразование координат имеет вид

$$r = \sqrt{x^2 + y^2}, \quad \varphi = \arctg \left(\frac{y}{x} \right). \quad (1.80)$$

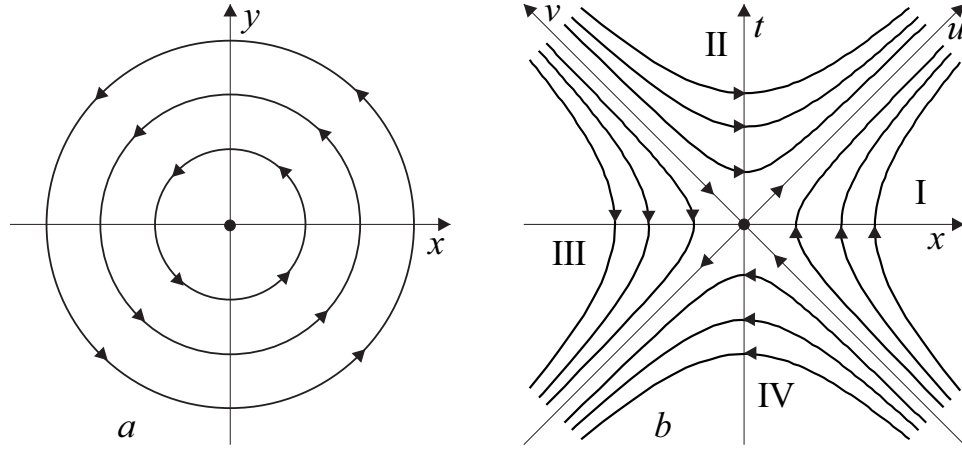


Рис. 1.6: Траектории Киллинга = орбиты действия групп $\mathbb{SO}(2)$ и $\mathbb{SO}_0(1, 1)$ на евклидовой плоскости (a) и плоскости Минковского (b).

Замечание. Эти координаты особенно удобны в задачах, обладающих вращательной симметрией, поскольку при вращении меняется только полярный угол, а радиус точки остается прежним. Например, для собственных вращений

$$S_+(\omega) : (r, \varphi) \mapsto (r, \varphi + \omega). \quad \square \quad (1.81)$$

Евклидов интервал в полярных координатах имеет нетривиальный вид

$$ds^2 = dr^2 + r^2 d\varphi^2. \quad (1.82)$$

Недостатком полярных координат является то, что они вырождены в начале координат, поскольку для этой точки полярный угол (1.80) не определен. Соответственно, якобиан преобразования (1.79),

$$J = \frac{\partial(x, y)}{\partial(r, \varphi)} = r,$$

равен нулю в начале координат. Во всех остальных точках преобразование координат является взаимно однозначным и сохраняет ориентацию, поскольку якобиан положителен.

Если исключить (выколоть) сингулярную точку, которой является начало координат, то евклидова плоскость в полярных координатах представляет собой топологическое произведение окружности $\mathbb{S}^1$ на полубесконечный интервал $\mathbb{R}_+ = (0, \infty)$,

$$\mathbb{R}^2 \setminus \{0\} = \mathbb{S}^1 \times \mathbb{R}_+.$$

Рассмотрим два ненулевых вектора в начале координат, $X_1 = (x_1, y_1) = (r_1, \varphi_1)$ и $X_2 = (x_2, y_2) = (r_2, \varphi_2)$. Их длины равны $\|X_1\| = r_1$ и $\|X_2\| = r_2$. Тогда *угол между векторами* φ_{12} , по определению, равен

$$\cos \varphi_{12} := \frac{(X_1, X_2)}{\|X_1\| \|X_2\|}, \quad (1.83)$$

где круглые скобки обозначают скалярное произведение векторов $(X_1, X_2) := x_1 x_2 + y_1 y_2$. Это соотношение определяет угол между векторами с точностью до знака, т.к. $\cos$ является четной функцией. Чтобы устранить возникшую неоднозначность, будем

считать угол φ_{12} между векторами X_1 и X_2 положительным, если вращение от X_1 к X_2 происходит против часовой стрелки. По определению, вращение от оси абсцисс к оси ординат по меньшему углу называется *вращением против часовой стрелки*. При вращении по часовой стрелке угол считается отрицательным. Тогда

$$\varphi_{12} := \varphi_2 - \varphi_1. \quad (1.84)$$

Отсюда следует, что углы между векторами складываются. То есть, если углы между векторами X_1, X_2 и X_2, X_3 равны, соответственно, φ_{12} и φ_{23} , то угол между векторами X_1 и X_3 равен сумме углов

$$\varphi_{13} := \varphi_{12} + \varphi_{23}. \quad (1.85)$$

Это правило сложения углов эквивалентно абелевости группы собственных вращений $\mathbb{SO}(2)$.

Рассмотрим бесконечно малые вращения евклидовой плоскости на угол $|\delta\omega| \ll 1$. Разлагая матрицу (1.71) в ряд, в линейном приближении получим следующие приращения координат:

$$\begin{aligned} \delta x &= -y\delta\omega, \\ \delta y &= x\delta\omega. \end{aligned} \quad (1.86)$$

Забегая вперед, скажем, что, если евклидова плоскость рассматривается, как многообразие, то генератором вращений является векторное поле Киллинга (см. главу 15)

$$K = -y\partial_x + x\partial_y.$$

Изменение формы функции $f(x, y)$ (см. раздел 2.13) в линейном приближении можно выразить через вектор Киллинга:

$$\delta f(x, y) := -\delta x^\alpha \partial_\alpha f = -\delta\omega Kf|_{(x,y)},$$

где векторное поле действует на функцию как дифференцирование. В полярных координатах, генератор вращений (векторное поле Киллинга) есть дифференцирование по полярному углу

$$K = -y \left(\frac{\partial r}{\partial x} \partial_r + \frac{\partial \varphi}{\partial x} \partial_\varphi \right) + x \left(\frac{\partial r}{\partial y} \partial_r + \frac{\partial \varphi}{\partial y} \partial_\varphi \right) = \partial_\varphi.$$

При этом интегральные кривые для генератора вращений (траектории Киллинга) совпадают с орбитами действия группы вращений и являются концентрическими окружностями (1.78) с центром в начале координат, как показано на рис. 1.6, а.

В комплексных координатах на евклидовой плоскости,

$$z := x + iy = r e^{i\varphi}, \quad \bar{z} := x - iy = r e^{-i\varphi}, \quad (1.87)$$

где черта обозначает комплексное сопряжение, собственные и несобственные вращения на угол ω выглядят, соответственно

$$\begin{aligned} S_+(\omega) : \quad z &\mapsto z' = z e^{i\omega}, & S_-(\omega) : \quad z &\mapsto z' = \bar{z} e^{-i\omega}, \\ \bar{z} &\mapsto z' = \bar{z} e^{-i\omega}, & \bar{z} &\mapsto z' = z e^{i\omega}. \end{aligned} \quad (1.88)$$

То есть собственное вращение комплексной плоскости – это умножение на комплексное число $e^{i\omega}$, равное по модулю единице, а несобственное вращение – это умножение

на комплексно сопряженное число $e^{-i\omega}$. Отсюда следует, что связная компонента единицы $\mathbb{SO}(2)$ изоморфна группе $\mathbb{U}(1)$ комплексных чисел, равных по модулю единице, с обычным правилом умножения.

Приведем для справки ряд формул. Интервал на евклидовой плоскости в комплексных координатах (1.87) имеет вид

$$ds^2 = dzd\bar{z}.$$

Операторы дифференцирования по декартовым и комплексным координатам связаны между собой соотношениями:

$$\begin{aligned} \partial_z &= \frac{1}{2}(\partial_x - i\partial_y), & \partial_{\bar{z}} &= \frac{1}{2}(\partial_x + i\partial_y), \\ \partial_x &= \partial_z + \partial_{\bar{z}}, & \partial_y &= i(\partial_z - \partial_{\bar{z}}). \end{aligned} \quad (1.89)$$

Отсюда следует выражение для лапласиана

$$\Delta = \partial_{xx}^2 + \partial_{yy}^2 = 4\partial_{z\bar{z}}^2. \quad (1.90)$$

Из формул преобразования комплексных координат (1.89) сразу следует, что оператор Лапласа инвариантен относительно полной группы вращений $\mathbb{O}(2)$.

1.7 Группа двумерных преобразований Лоренца $\mathbb{O}(1, 1)$

Определение. *Плоскостью Минковского $\mathbb{R}^{1,1}$ называется евклидова плоскость с декартовыми координатами $\{x^\alpha\} = \{t, x\}$, $\alpha = 0, 1$, в которых задана двумерная метрика Лоренца*

$$\eta = \{\eta_{\alpha\beta}\} := \text{diag}(+-) := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.91)$$

В пространстве Минковского декартова система координат, в которой задана метрика (1.91), называется *инерциальной* системой отсчета. Метрике Лоренца соответствует интервал

$$ds^2 = dx^\alpha dx^\beta \eta_{\alpha\beta} = dt^2 - dx^2 = dudv, \quad (1.92)$$

где введены *координаты светового конуса* (см. рис. 1.6, b):

$$u = t + x, \quad v = t - x, \quad (1.93)$$

которые являются аналогом комплексных координат (1.87) на евклидовой плоскости. Координатные оси t и x называют, соответственно, *временем* и *пространством*. $\square$

Якобиан преобразования координат (1.93) равен

$$J = \frac{\partial(u, v)}{\partial(x, t)} = 2.$$

Это преобразование сохраняет ориентацию, если координаты упорядочены следующим образом: $x, t \mapsto u, v$. В координатах светового конуса линии, параллельные осям координат, $u = \text{const}$ и $v = \text{const}$, имеют нулевую длину, т.е. светоподобны.

Замечание. На плоскости Минковского $\mathbb{R}^{1,1}$ время и пространство играют равноправную роль, поскольку пространство одномерно, и переобозначение $t \leftrightarrow x$ приводит только к изменению общего знака метрики. $\square$

Подчеркнем, что на плоскости Минковского $\mathbb{R}^{1,1}$ задано две метрики: евклидова метрика и метрика Минковского (1.91). Первая необходима для задания топологии и определения дифференцируемых в обычном смысле функций на $\mathbb{R}^{1,1}$.

Предложение 1.7.1. На плоскости Минковского любая изотропная кривая $\{t(s), x(s)\}$, т.е. линия, касательный вектор к которой имеет нулевую длину:

$$\left(\frac{dt}{ds}\right)^2 - \left(\frac{dx}{ds}\right)^2 = 0, \quad (1.94)$$

является прямой линией вида $t = \pm x + \text{const}$.

Доказательство. Из уравнения (1.94) следует, что $dt = \pm dx$. Общим решением данного уравнения являются прямые линии $t = \pm x + \text{const}$. $\square$

Определение. Уравнение

$$(t - t_0)^2 - (x - x_0)^2 = 0$$

определяет на плоскости Минковского $\mathbb{R}^{1,1}$ две перпендикулярные прямые, пересекающиеся в точке (t_0, x_0) . Они называются *световым конусом* в данной точке, рис. 1.7. Световой конус состоит из двух связанных компонент: светового конуса *прошлого*, $t \leq t_0$, и *будущего*, $t \geq t_0$. Световые конусы прошлого и будущего имеют общую вершину в точке (t_0, x_0) . $\square$

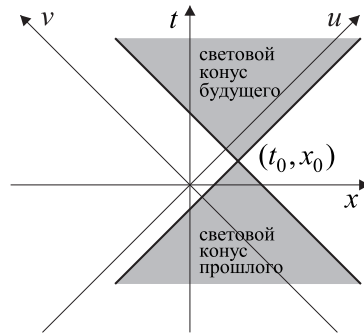


Рис. 1.7: Световой конус на плоскости Минковского.

Любая времениподобная кривая $\{x^\alpha(s)\}$,

$$\dot{x}^\alpha \dot{x}^\beta \eta_{\alpha\beta} > 0, \quad \dot{x}^\alpha := \frac{dx^\alpha}{ds},$$

проходящая через точку (t_0, x_0) , будет целиком лежать внутри светового конуса (затемненная область на рисунке).

Определение. Невырожденные линейные однородные преобразования координат

$$x^\alpha \mapsto x'^\alpha = x^\beta S_\beta^\alpha, \quad \det S_\beta^\alpha \neq 0, \quad (1.95)$$

относительно которых метрика Минковского (1.91) остается инвариантной,

$$\eta_{\alpha\beta} = S_\alpha^\gamma S_\beta^\epsilon \eta_{\gamma\epsilon}, \quad (1.96)$$

образуют группу, которая называется *группой двумерных преобразований Лоренца* и обозначается $\mathbb{O}(1, 1)$. $\square$

Из уравнения (1.96) следует, что определитель матриц S , как и в случае вращений евклидовой плоскости, по модулю равен единице, $\det S = \pm 1$.

Предложение 1.7.2. Любое решение уравнения (1.96) принадлежит одному из четырех классов:

$$S_{+\uparrow}(\omega) := \begin{pmatrix} \text{ch } \omega & \text{sh } \omega \\ \text{sh } \omega & \text{ch } \omega \end{pmatrix}, \quad \det S_{+\uparrow} = 1, \quad (1.97)$$

$$S_{-\uparrow}(\omega) := \begin{pmatrix} \text{ch } \omega & -\text{sh } \omega \\ \text{sh } \omega & -\text{ch } \omega \end{pmatrix}, \quad \det S_{-\uparrow} = -1, \quad (1.98)$$

$$S_{+\downarrow}(\omega) := \begin{pmatrix} -\text{ch } \omega & -\text{sh } \omega \\ -\text{sh } \omega & -\text{ch } \omega \end{pmatrix}, \quad \det S_{+\downarrow} = 1, \quad (1.99)$$

$$S_{-\downarrow}(\omega) := \begin{pmatrix} -\text{ch } \omega & \text{sh } \omega \\ -\text{sh } \omega & \text{ch } \omega \end{pmatrix}, \quad \det S_{-\downarrow} = -1, \quad (1.100)$$

при некотором $\omega \in \mathbb{R}$. Параметр ω называется гиперболическим углом вращения.

Доказательство. Уравнение (1.91) – это система трех квадратных уравнений:

$$\begin{aligned} (S_1^1)^2 - (S_1^2)^2 &= 1, \\ (S_2^1)^2 - (S_2^2)^2 &= -1, \\ S_1^1 S_2^1 - S_1^2 S_2^2 &= 0. \end{aligned}$$

Перебор всех возможных решений показывает, что с точностью до переопределения угла поворота формулы (1.97)–(1.100) дают общее решение. $\square$

Индекс $\pm$ у матрицы преобразований Лоренца соответствует знаку определителя, а смысл стрелки будет ясен из дальнейшего рассмотрения.

Из доказанного утверждения следует, что двумерная группа Лоренца $\mathbb{O}(1, 1)$ является однопараметрической некомпактной группой Ли и состоит из четырех несвязных компонент.

Заметим, что

$$S_{-\uparrow}^T(\omega) = S_{-\uparrow}(-\omega) \quad \text{и} \quad S_{-\downarrow}^T(\omega) = S_{-\downarrow}(-\omega).$$

Как и в случае группы вращений евклидовой плоскости $\mathbb{O}(2)$, обозначим множества матриц (1.97)–(1.100) теми же, но ажурными буквами. Например,

$$\mathbb{S}_{+\uparrow} := \{S_{+\uparrow}(\omega) : \omega \in \mathbb{R}\}.$$

Тогда выполнена следующая таблица умножения

	$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{+\downarrow}$
$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{+\downarrow}$
$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{+\downarrow}$	$\mathbb{S}_{-\downarrow}$
$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{+\downarrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{+\uparrow}$
$\mathbb{S}_{+\downarrow}$	$\mathbb{S}_{+\downarrow}$	$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{+\uparrow}$

(1.101)

Обратные матрицы из $\mathbb{S}_{+\uparrow}$ имеют вид

$$S_{+\uparrow}^{-1}(\omega) = S_{+\uparrow}(-\omega),$$

т.е. получаются заменой $\omega \mapsto -\omega$.

Единичная матрица $\mathbb{1} = S_{+\uparrow}(0)$ и матрица $-\mathbb{1} = S_{+\downarrow}(0)$ принадлежат разным компонентам группы.

Определение. Введем оператор *обращения времени и пространственного отражения (четности)*

$$\begin{aligned} T &:= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = S_{-\downarrow}(0), & \det T = -1, \\ P &:= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = S_{-\uparrow}(0), & \det P = -1. \end{aligned} \quad (1.102)$$

Действие этих операторов мы записываем в виде $x^\alpha \mapsto x^\beta T_\beta^\alpha$ и $x^\alpha \mapsto x^\beta P_\beta^\alpha$. Они отражают, соответственно, временную и пространственную координаты:

$$\begin{aligned} T: & \quad (t, x) \mapsto (-t, x), \\ P: & \quad (t, x) \mapsto (t, -x). \end{aligned}$$

Введем также обозначение для их композиции

$$R := PT = TP = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = S_{+\uparrow}(0), \quad \det R = 1, \quad (1.103)$$

которая соответствует полному отражению координат $R: (t, x) \mapsto (-t, -x)$. $\square$

Тогда матрицы преобразований Лоренца (1.98)–(1.100) можно представить в виде

$$\begin{aligned} S_{-\uparrow} &= S_{+\uparrow}P, \\ S_{-\downarrow} &= S_{+\uparrow}T, \\ S_{+\downarrow} &= S_{+\uparrow}R. \end{aligned} \quad (1.104)$$

Преобразования (1.97) и (1.98) не меняют направления времени и называются *ортохронными*, что отмечено направленной вверх стрелкой. Эти преобразования образуют компоненты $S_{+\uparrow}$ и $S_{-\uparrow}$, для которых $S_0^0 > 0$. Матрицы из компонент $S_{+\downarrow}$ и $S_{-\downarrow}$ не являются ортохронными, и для них $S_0^0 < 0$.

Из таблицы умножения (1.101) следует, что только одна из четырех связных компонент $\mathbb{S}_{+\uparrow}$ (связная компонента единицы) образует группу. Эта группа обозначается

$$\mathbb{SO}_0(1, 1) \approx \mathbb{S}_{+\uparrow}, \quad \det S = 1, \quad S_0^0 > 0, \quad (1.105)$$

и называется *собственной ортохронной* группой Лоренца. Кроме этого группа Лоренца $\mathbb{O}(1, 1)$ имеет еще три нетривиальные подгруппы, состоящие из двух компонент, которые включают отражение времени T , преобразование четности P и полного отражения R . Введем для этих подгрупп следующие обозначения:

$$\begin{aligned}\mathbb{O}_{\downarrow}(1, 1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{-\downarrow}, & \det S &= \pm 1, \\ \mathbb{O}_0(1, 1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{-\uparrow}, & \det S &= \pm 1, \quad S_0^0 > 0, \\ \mathbb{SO}(1, 1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{+\downarrow}, & \det S &= 1.\end{aligned}\tag{1.106}$$

Матрицы подгрупп $\mathbb{SO}(1, 1)$ и $\mathbb{SO}_0(1, 1)$ имеют положительный определитель. Легко проверить, что они являются абелевыми. Подгруппы $\mathbb{O}_0(1, 1)$ и $\mathbb{O}_{\downarrow}(1, 1)$, а также полная группа Лоренца $\mathbb{O}(1, 1)$ являются неабелевыми.

Замечание. Знание подгрупп полной группы Лоренца $\mathbb{O}(1, 1)$ важно для описания представлений, в том числе спинорных. $\square$

Определение. Если при пространственных отражениях скаляр, вектор и вообще произвольный тензор меняет знак, то их принято называть *псевдоскаляр*, *псевдовектор* и *псевдотензор*. $\square$

Все подгруппы в полной группе Лоренца $\mathbb{O}(1, 1)$ являются нормальными. Поэтому определена факторгруппа

$$\mathbb{K}_4 := \frac{\mathbb{O}(1, 1)}{\mathbb{SO}_0(1, 1)},$$

состоящая из четырех элементов. Из выражений (1.104) следует, что в качестве представителей смежных классов можно выбрать четыре матрицы $\mathbb{1}, P, T, R$. Эти матрицы, как легко проверить, имеют следующую таблицу умножения

	$\mathbb{1}$	P	T	R
$\mathbb{1}$	$\mathbb{1}$	P	T	R
P	P	$\mathbb{1}$	R	T
T	T	R	$\mathbb{1}$	P
R	R	T	P	$\mathbb{1}$

(1.107)

которая с точностью до обозначений совпадает с (1.101).

Определение. Группа состоящая из четырех элементов, удовлетворяющих таблице умножения (1.107), называется *4-группой Клейна*. $\square$

Поскольку $P^2 = 1$, $T^2 = 1$ и $PT = TP$, то группа Клейна изоморфна прямому произведению групп: $\mathbb{K}_4 \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$. Действительно, если $\mathbb{Z}_2 := \{1, -1\}$, то достаточно положить $P := (1, -1)$ и $T := (-1, 1)$.

Замечание. Группа Клейна не является циклической. $\square$

Элементы группы $\mathbb{SO}_0(1, 1)$ можно представить в виде

$$S_{+\uparrow\alpha}{}^\beta = (e^{-\omega\varepsilon})_\alpha{}^\beta = \delta_\alpha^\beta \operatorname{ch} \omega - \varepsilon_\alpha{}^\beta \operatorname{sh} \omega,\tag{1.108}$$

где $\varepsilon_{\alpha\beta} = \varepsilon_\alpha{}^\gamma \eta_{\gamma\beta}$, $\varepsilon_{01} = 1$, — полностью антисимметричный тензор второго ранга в пространстве Минковского (см. приложение 37.4).

В координатах светового конуса (1.93) преобразования из группы Лоренца имеют вид:

$$\begin{aligned} S_{+\uparrow} : \quad u &\mapsto u' = u e^{\omega}, & S_{+\downarrow} : \quad u &\mapsto u' = -u e^{\omega}, \\ &v \mapsto v' = v e^{-\omega}, & &v \mapsto v' = -v e^{-\omega}, \\ S_{-\uparrow} : \quad u &\mapsto u' = v e^{-\omega}, & S_{-\downarrow} : \quad u &\mapsto u' = -v e^{-\omega}, \\ &v \mapsto v' = u e^{\omega}, & &v \mapsto v' = -u e^{\omega}. \end{aligned} \quad (1.109)$$

Матрицы (1.97)–(1.100) можно рассматривать как представление группы Лоренца $\mathbb{O}(1, 1)$ в двумерном векторном пространстве. Оно неприводимо и называется *фундаментальным* или *векторным* представлением. Относительно группы собственных ортохронных преобразований $\mathbb{SO}_0(1, 1)$ и группы собственных преобразований $\mathbb{SO}(1, 1)$ это представление является приводимым и распадается на два неприводимых одномерных представления (1.109).

Замечание. Все неприводимые представления собственных групп Лоренца $\mathbb{SO}_0(1, 1)$ и $\mathbb{SO}(1, 1)$ являются одномерными, поскольку группы абелевы. Это является причиной того, что выбор конусных координат часто приводит к существенному упрощению вычислений. $\square$

Бесконечно малые собственные ортохронные лоренцевы вращения на угол $\delta\omega \ll 1$ имеют вид

$$S_{+\uparrow\alpha}{}^{\beta} = \delta_{\alpha}{}^{\beta} + \delta\omega L_{\alpha}{}^{\beta}, \quad (1.110)$$

где $L_{\alpha}{}^{\beta} = -\varepsilon_{\alpha}{}^{\beta}$ – генераторы группы Лоренца в фундаментальном представлении.

Гиперболические вращения плоскости Минковского при действии собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, 1)$ показаны на рис. 1.6, *b*. При этом плоскость разделяется на четыре квадранта I–IV. Собственные ортохронные преобразования Лоренца преобразуют точки внутри каждого квадранта. Орбитами этих точек являются гиперболы

$$t^2 - x^2 = \pm r^2, \quad r \in \mathbb{R}_+. \quad (1.111)$$

Здесь знак плюс соответствует II и IV, а минус – I и III квадрантам. Кроме того, орбитами группы $\mathbb{SO}_0(1, 1)$ являются четыре луча $t = \pm x$, выходящих из начала координат $(0, 0)$, которое является неподвижной точкой. В каждом из квадрантов можно ввести гиперболическую полярную систему координат:

$$\begin{aligned} \text{I} : \quad t &= r \operatorname{sh} \varphi, & \text{III} : \quad t &= -r \operatorname{sh} \varphi, \\ &x = r \operatorname{ch} \varphi, & &x = -r \operatorname{ch} \varphi, \\ \text{II} : \quad t &= r \operatorname{ch} \varphi, & \text{IV} : \quad t &= -r \operatorname{ch} \varphi, \\ &x = r \operatorname{sh} \varphi, & &x = -r \operatorname{sh} \varphi, \end{aligned} \quad (1.112)$$

где $r \in (0, \infty)$, $\varphi \in (-\infty, \infty)$ в каждом квадранте. При этом “началом” координат, $r = 0$, являются две перпендикулярные прямые

$$r = 0 \leftrightarrow t = \pm x,$$

т.е. световой конус в начале координат. На рис. 1.6, *b* стрелками показано направление возрастания угла φ внутри квадрантов и возрастание гиперболического угла вращения ω на световом конусе.

Обратные преобразования координат, например, во II квадранте имеют вид

$$r = \sqrt{t^2 - x^2}, \quad \varphi = \operatorname{arch} \left(\frac{x}{t} \right).$$

Знаки для преобразования координат в различных квадрантах (1.112) подобраны таким образом, чтобы при лоренцевом преобразовании (1.97) полярный угол менялся по закону

$$\varphi \mapsto \varphi' = \varphi + \omega, \quad (1.113)$$

при этом радиус точки остается прежним. Лоренцев интервал в гиперболических полярных координатах отличается знаком в разных квадрантах:

$$\begin{aligned} \text{I, III : } ds^2 &= -dr^2 + r^2 d\varphi^2, \\ \text{II, IV : } ds^2 &= dr^2 - r^2 d\varphi^2. \end{aligned} \quad (1.114)$$

Отсюда следует, что радиус точки играет роль пространственной координаты в I и III квадранте и времениподобной координаты во II и IV квадранте. Угол, наоборот, времениподобен в I и III квадранте и пространственноподобен во II и IV квадранте. Якобиан преобразования (1.112) имеет различные знаки в разных квадрантах:

$$\text{I, III : } J = \frac{\partial(x, t)}{\partial(r, \varphi)} = r, \quad \text{II, IV : } J = \frac{\partial(x, t)}{\partial(r, \varphi)} = -r$$

и вырожден на линиях $r = 0$. Переход к гиперболическим полярным координатам сохраняет ориентацию базиса в I и III квадрантах и меняет ориентацию во II и IV квадрантах.

С топологической точки зрения внутренность каждого квадранта является прямым произведением полубесконечного интервала и прямой

$$\text{int}(\text{I}) = \mathbb{R}_+ \times \mathbb{R}.$$

Нетрудно выразить координаты светового конуса через полярные координаты. Например, во втором квадранте

$$u = r e^\varphi, \quad v = r e^{-\varphi},$$

что является аналогом полярной записи комплексных чисел (1.87).

При пространственном отражении, обращении времени и полном отражении квадранты отображаются друг на друга следующим образом:

$$\begin{array}{llll} P : & \text{I} \leftrightarrow \text{III}, & \text{II} \rightarrow \text{II}, & \text{IV} \rightarrow \text{IV}, \\ T : & \text{I} \rightarrow \text{I}, & \text{II} \leftrightarrow \text{IV}, & \text{III} \rightarrow \text{III}, \\ R : & \text{I} \leftrightarrow \text{III}, & \text{II} \leftrightarrow \text{IV}, & \end{array}$$

Определим угол φ_{12} между векторами X_1 и X_2 из одного квадранта следующим образом:

$$\text{ch } \varphi_{12} := \frac{|(X_1, X_2)|}{|X_1||X_2|}. \quad (1.115)$$

Этот угол определен с точностью до знака. Будем считать угол φ_{12} между векторами $X_1 = \{r_1, \varphi_1\}$ и $X_2 = \{r_2, \varphi_2\}$ положительным, если поворот от X_1 к X_2 происходит в сторону увеличения полярного угла. Тогда

$$\varphi_{12} = \varphi_2 - \varphi_1. \quad (1.116)$$

Отсюда следует, что углы между векторами складываются (1.85), что соответствует абелевости собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, 1)$. Подчеркнем, что понятие угла вводится только для векторов из одного квадранта.

Бесконечно малые гиперболические вращения плоскости Минковского $\mathbb{R}^{1,1}$ на угол $\delta\omega \ll 1$ можно получить, разлагая матрицу (1.97) в ряд Тейлора. В линейном приближении получим следующие приращения координат:

$$\delta t = x\delta\omega, \quad \delta x = t\delta\omega. \quad (1.117)$$

Отсюда следует, что генератором лоренцевых преобразований является векторное поле Киллинга

$$K = x\partial_t + t\partial_x.$$

В полярных координатах, т.е. внутри каждого квадранта, генератор вращений есть дифференцирование по полярному углу:

$$K = \partial_\varphi.$$

При этом интегральные кривые для генератора вращений (траектории Киллинга) совпадают с орбитами действия группы и представляют собой гиперболы (1.111), как показано на рис. 1.6, *b*.

В физике собственные ортохронные преобразования Лоренца (1.108) обычно записывают для размерных координат $x^0 = ct$, $x^1 = x$, где c – скорость света. Они имеют вид

$$t' = \frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}}, \quad x' = \frac{x - Vt}{\sqrt{1 - V^2/c^2}}, \quad (1.118)$$

где V – скорость движения штрихованной системы координат вдоль оси x . При этом угол гиперболического поворота ω выражается через скорость относительного движения V с помощью соотношения

$$\text{th } \omega = -\frac{V}{c}. \quad (1.119)$$

Поскольку $-1 < \text{th } \omega < 1$, то преобразования Лоренца определены для скоростей, меньших скорости света $V < c$. Поэтому из постулата ковариантности законов Природы относительно преобразований Лоренца вытекает, что скорость света является предельной скоростью для частиц и волн. Это предположение является постулатом специальной теории относительности и находится в прекрасном согласии с известными экспериментальными данными.

Преобразования Лоренца (1.118) имеют следующую физическую интерпретацию. Дополним двумерное пространство Минковского до четырехмерного, введя дополнительные декартовы координаты y, z . Предположим, что дополнительные координаты не меняются

$$y' = y, \quad z' = z.$$

Тогда координаты t, x, y, z и t', x', y', z' являются координатами одного и того же события в двух инерциальных системах отсчета с параллельными осями координат (см. рис.1.8). При этом штрихованная система координат движется равномерно и прямолинейно со скоростью V относительно нештрихованной системы вдоль оси x . Системы координат выбраны таким образом, что их начала совпадают:

$$t = 0, x = 0 \quad \Leftrightarrow \quad t' = 0, x' = 0.$$

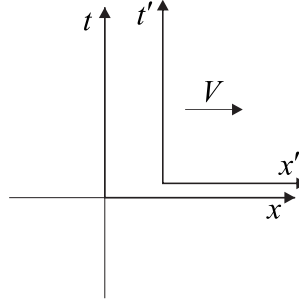


Рис. 1.8: Физическая интерпретация преобразований Лоренца. Штрихованная инерциальная система отсчета движется равномерно и прямолинейно относительно нештрихованной со скоростью V .

Преобразования Лоренца (1.118) при $c \rightarrow \infty$ переходят в преобразования Галилея

$$x' = x - Vt, \quad t' = t. \quad (1.120)$$

Эти преобразования хорошо знакомы из курса классической механики, где время является абсолютным и не зависит от выбора системы координат.

Вернемся к двумерному пространству Минковского. Поскольку в приложениях часто используются конусные координаты (1.93), в заключение настоящего раздела приведем для справок некоторые формулы:

$$t = \frac{1}{2}(u + v), \quad x = \frac{1}{2}(u - v), \quad (1.121)$$

$$\begin{aligned} \partial_u &= \frac{1}{2}(\partial_t + \partial_x), & \partial_v &= \frac{1}{2}(\partial_t - \partial_x), \\ \partial_t &= \partial_u + \partial_v, & \partial_x &= \partial_u - \partial_v. \end{aligned} \quad (1.122)$$

Оператор Даламбера на плоскости Минковского имеет вид

$$\square := \partial_{tt}^2 - \partial_{xx}^2 = 4\partial_{uv}^2. \quad (1.123)$$

Он инвариантен относительно действия полной группы Лоренца $\mathbb{O}(1, 1)$.

1.8 Трехмерное евклидово пространство

Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовыми координатами $\{x^i\} = \{x, y, z\}$, $i = 1, 2, 3$.

Замечание. В настоящем разделе для обозначения индексов используются буквы из середины латинского алфавита $i, j, k, \dots$. Такое изменение обозначений связано с тем, что в дальнейшем координаты в пространстве Минковского $\mathbb{R}^{1,3}$ часто будут обозначаться через $\{x^a\} = \{x^0, x^i\}$, где явно выделено время $x^0 = t$, $a = 0, 1, 2, 3$ и $i = 1, 2, 3$. В обозначениях мы используем следующее правило. Множество чисел $\{1, 2, 3\}$ является подмножеством чисел $\{0, 1, 2, 3\}$, так же как и буквы $\{i, j, \dots\}$ являются подмножеством всего алфавита $a, b, \dots$. $\square$

Дифференциально геометрическая евклидова метрика в декартовой системе координат задается единичной матрицей

$$ds^2 = \delta_{ij} dx^i dx^j = dx^2 + dy^2 + dz^2. \quad (1.124)$$

Определение. Группа линейных неоднородных преобразований декартовых координат, оставляющая метрику (1.124) инвариантной, называется *неоднородной группой вращений* и обозначается $\mathbb{I}\mathbb{O}(3)$. В общем случае преобразование из неоднородной группы вращений имеет вид

$$x^i \mapsto x'^i = x^j S_j^i + a^i,$$

где $S_j^i \in \mathbb{O}(3)$ – вещественная ортогональная матрица, определяемая условием

$$S^T S = S S^T = 1, \quad (1.125)$$

и $a^i \in \mathbb{R}^3$ – произвольный вектор. Матрицы $\mathbb{O}(3)$ образуют группу, которая называется *группой трехмерных вращений*. Абелева подгруппа в $\mathbb{I}\mathbb{O}(3)$, параметризуемая вектором a , называется *группой трансляций*. $\square$

Поскольку группа вращений сохраняет метрику (1.124), то при поворотах и трансляциях евклидова пространства сохраняются длины и углы между векторами. Конечно, трансляции можно было бы ввести и ранее в двумерном случае.

Можно показать, что любое преобразование декартовых координат, оставляющее метрику (1.124) инвариантной, является линейным. Поэтому группа $\mathbb{I}\mathbb{O}(3)$ является максимальной группой преобразований симметрии евклидова пространства $\mathbb{R}^3$ с заданной метрикой. Группа неоднородных вращений, как будет показано в дальнейшем, является группой Ли.

Группа $\mathbb{I}\mathbb{O}(3)$ представляет собой полупрямое произведение группы вращений $\mathbb{O}(3)$ вокруг начала координат на абелеву группу трансляций. Группа трансляций действует в $\mathbb{R}^3$ свободно и транзитивно. Как многообразие она диффеоморфна евклидову пространству $\mathbb{R}^3$. Группа вращений $\mathbb{O}(3)$ действует в $\mathbb{R}^3$ эффективно. При этом начало координат является неподвижной точкой. Многообразие группы вращений будет описано немного позже.

Из условия ортогональности (1.125) следует, что определитель ортогональной матрицы равен по модулю единице

$$\det S = \pm 1.$$

Группа трехмерных вращений $\mathbb{O}(3)$ является неабелевой и состоит из двух компонент. Связная компонента единицы обозначается $\mathbb{SO}(3)$ и называется группой *собственных* вращений. К ней относятся ортогональные матрицы с единичным определителем. Вторая связная компонента состоит из ортогональных матриц с отрицательным определителем и сама по себе группы не образует.

Определение. *Оператором пространственных отражений* называется матрица

$$P := \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{или} \quad P_j^i = -\delta_j^i, \quad (1.126)$$

при действии которой все координаты меняют знак $x^i \mapsto x^j P_j^i = -x^i$. $\square$

Произвольную ортогональную матрицу с отрицательным определителем можно однозначно представить в виде произведения собственно ортогональной матрицы и оператора пространственных отражений. Это значит, что множество ортогональных

матриц с отрицательным определителем представляет собой смежный класс матрицы отражений по подгруппе $\mathbb{SO}(3)$.

Связная компонента единицы $\mathbb{SO}(3)$ является нормальной подгруппой в группе вращений $\mathbb{O}(3)$. Поэтому определена фактор группа

$$\mathbb{Z}_2 = \frac{\mathbb{O}(3)}{\mathbb{SO}(3)},$$

состоящая из двух элементов $\mathbb{Z}_2 = \{1, P\}$. Отображение

$$\mathbb{O}(3) \ni S \mapsto \det S \in \mathbb{Z}_2$$

есть гомоморфизм группы вращений на мультипликативную группу $\mathbb{Z}_2$. Как многообразие полная группа вращений гомеоморфна прямому произведению: $\mathbb{O}(3) \approx \mathbb{Z}_2 \times \mathbb{SO}(3)$.

Группа $\mathbb{SO}(3)$ проста (т.е. не содержит инвариантных абелевых подгрупп), и, значит, нетривиальные гомоморфизмы на другие группы Ли невозможны (см. раздел 8.8).

При вращениях произвольная точка $x \in \mathbb{R}^3$, не совпадающая с началом координат, пробегает все точки сферы $\mathbb{S}^2$ с центром в начале координат и содержащей точку $x \in \mathbb{S}^2$. При этом сфера является орбитой точки $x \in \mathbb{R}^3$ при действии групп $\mathbb{SO}(3)$ и $\mathbb{O}(3)$. Отметим, что две точки сферы $x_1, x_2 \in \mathbb{S}^2$ не определяют вращение однозначно (см. рис. 1.9). Действительно, пусть точки x_1 и x_2 лежат в плоскости x, z и симметричны относительно оси z . Тогда точку x_2 можно получить из точки x_1 вращением вокруг оси z на угол $\pm\pi$. Этого же можно добиться вращением вокруг оси y на меньший угол, причем траекторией точки x_1 будет часть большой окружности, соединяющей точки x_1 и x_2 . В общем случае осью вращения может служить любой единичный вектор, лежащий в плоскости, перпендикулярной отрезку, соединяющему точки x_1 и x_2 и проходящей через начало координат. В рассматриваемом случае ось вращения должна лежать в плоскости y, z . Если на сфере задать перемещение не двух, а всех точек одновременно, тогда вращение будет определено однозначно.

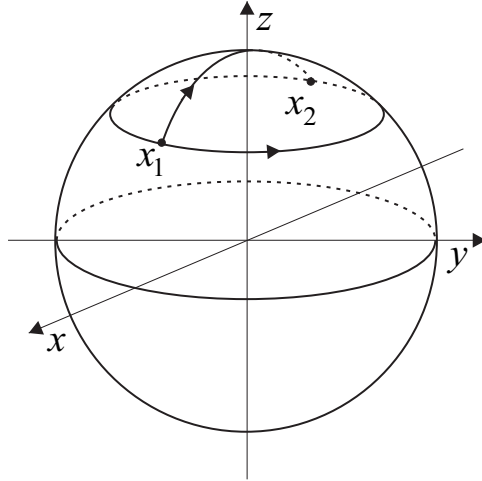


Рис. 1.9: Отображение двух точек на сфере $x_1 \mapsto x_2$ не определяет трехмерное вращение однозначно.

Генераторы группы вращений $\mathbb{SO}(3)$ могут быть представлены в евклидовом пространстве в виде дифференциальных операторов (векторных полей Киллинга)

$$L_{ij} = x_i \partial_j - x_j \partial_i. \quad (1.127)$$

Прямое вычисление коммутатора приводит к следующей алгебре Ли

$$[L_{ij}, L_{kl}] = -\delta_{ik}L_{jl} + \delta_{il}L_{jk} + \delta_{jk}L_{il} - \delta_{jl}L_{ik}. \quad (1.128)$$

Отсюда следует, что группа трехмерных вращений $\mathbb{SO}(3)$ является неабелевой группой Ли. Эта группа трехмерна, т.к. ввиду антисимметрии по индексам, существуют только три независимых вектора: L_{12} , L_{13} и L_{23} . Алгебру Ли (1.128) можно переписать в более компактном виде для эквивалентного набора генераторов

$$J_i := \frac{1}{2}\varepsilon_{ijk}L^{jk}, \quad L_{ij} = \varepsilon_{ijk}J^k, \quad (1.129)$$

где ε_{ijk} — полностью антисимметричный тензор третьего ранга, $\varepsilon_{123} = 1$, (см. Приложение 37.4), а подъем и опускание индексов производится с помощью евклидовой метрики. Тогда алгебра Ли группы $\mathbb{SO}(3)$ принимает вид

$$[J_i, J_j] = -\varepsilon_{ijk}J^k. \quad (1.130)$$

Алгебры Ли групп Ли $\mathbb{SO}(3)$ и $\mathbb{O}(3)$, естественно, совпадают.

При проведении расчетов с группой $\mathbb{O}(3)$ удобно использовать явную параметризацию элементов группы элементами ее алгебры. Элемент алгебры $\mathfrak{so}(3)$ может быть представлен в виде произвольной антисимметричной 3×3 -матрицы

$$(\omega\varepsilon)_i^j = (\omega^k\varepsilon_k)_i^j = \omega^k\varepsilon_{ki}^j \in \mathfrak{so}(3), \quad (1.131)$$

где $\omega = \{\omega^k\} \in \mathbb{R}^3$ — вектор вращения. Здесь первый индекс k нумерует базис алгебры, а индексы i, j рассматриваются, как матричные. Для сравнения с (1.130) заметим, что полностью антисимметричный тензор третьего ранга можно рассматривать в качестве базиса алгебры $\mathfrak{so}(3)$:

$$(J_k)_i^j := \varepsilon_{ki}^j.$$

Элемент алгебры параметризуется трехмерным вектором $\omega \in \mathbb{R}^3$, и, значит, группа $\mathbb{O}(3)$ является трехмерной группой Ли.

Предложение 1.8.1. *Каждому элементу алгебры Ли (вектору вращения) $\omega = \{\omega^i\} \in \mathfrak{so}(3)$ соответствует элемент группы Ли $S = \{S_i^j\} \in \mathbb{SO}(3)$ (связной компоненты единицы):*

$$S_i^j(\omega) = (e^{(\omega\varepsilon)})_i^j = \delta_i^j \cos \omega + \frac{(\omega\varepsilon)_i^j}{\omega} \sin \omega + \frac{\omega_i \omega^j}{\omega^2} (1 - \cos \omega) \in \mathbb{SO}(3), \quad (1.132)$$

где $\omega := \sqrt{\omega^i \omega_i}$ — модуль вектора вращения ω .

Доказательство. Последнее равенство в (1.132) доказывается путем разложения обеих частей равенства в ряды, которые равномерно сходятся для всех $\omega < \infty$. С помощью прямых вычислений нетрудно убедиться, что S_i^j является действительно ортогональной матрицей. $\square$

Формула (1.132) задает экспоненциальное отображение элемента алгебры $\omega \in \mathfrak{so}(3) \approx \mathbb{R}^3$ в элемент группы $S \in \mathbb{SO}(3)$. Можно также доказать, что любая ортогональная матрица с единичным определителем имеет вид (1.132) для некоторого

вектора ω . Поэтому данное экспоненциальное отображение сюръективно. В некоторой окрестности единицы группы оно взаимно однозначно. Однако в целом отображение (1.132) не является взаимно однозначным, т.к. различные элементы алгебры отображаются в один элемент группы. Как мы увидим, в каждый элемент группы отображается бесконечное счетное число элементов алгебры.

В отличие от элемента алгебры, элемент группы имеет как симметричную, так и антисимметричную части.

Элемент группы $\mathbb{SO}(3)$ инвариантен относительно сдвига вектора вращения $\omega^i \mapsto \omega^i + 2\pi\omega^i/\omega$:

$$S_i^j \left(\omega^k + 2\pi \frac{\omega^k}{\omega} \right) = S_i^j(\omega^k),$$

и это единственная инвариантность. При сдвиге вектора ω меняется только его длина $\omega \mapsto \omega + 2\pi$, а направление остается неизменным. В инвариантности матрицы вращений нетрудно убедиться, если заметить, что отношение ω^i/ω , определяющее направление вектора ω , не меняется при произвольном сдвиге:

$$\frac{\omega^i}{\omega} \mapsto \frac{\omega^i + c\omega^i/\omega}{\omega + c} = \frac{\omega^i}{\omega}, \quad c \in \mathbb{R}.$$

Тем самым элемент группы вращений параметризуется точками евклидова пространства $\omega \in \mathbb{R}^3$ с единственным отношением эквивалентности

$$\omega^i \sim \omega^i + 2\pi \frac{\omega^i}{\omega}. \quad (1.133)$$

При этом неопределенность в нуле раскрывается по радиальным направлениям $\omega^i = \epsilon k^i$, $\epsilon \rightarrow 0$. То есть начало координат отождествляется со всеми сферами радиуса $2\pi m$, $m = 1, 2, \dots$

Вектор вращения ω параметризует группу $\mathbb{SO}(3)$ следующим образом. Направление вектора ω определяет ось вращения, а модуль вектора ω равен углу поворота. Таким образом, каждый элемент группы $\mathbb{SO}(3)$ отождествляется с точкой трехмерного шара $\mathbb{B}_\pi^3 \subset \mathbb{R}^3$ радиуса π с центром в начале координат. При этом различными внутренним точкам шара соответствуют различные вращения, а диаметрально противоположные точки граничной сферы $\mathbb{S}_\pi^2$ необходимо отождествить, т.к. повороты вокруг фиксированной оси на углы π и $-\pi$ приводят к одному и тому же результату.

Таким образом, мы построили экспоненциальное отображение (1.132) алгебры Ли в группу Ли:

$$\mathfrak{so}(3) \ni \quad \omega \mapsto S \quad \in \mathbb{SO}(3).$$

Это отображение сюръективно, но не является взаимно однозначным. Оно также не является накрытием. Действительно, прообразом каждого нетривиального вращения вокруг некоторой оси является счетное число точек, лежащих на этой оси на расстоянии 2π друг от друга. В то же время прообразом единицы группы $e \in \mathbb{SO}(3)$ является начало координат в $\mathbb{R}^3 \approx \mathfrak{so}(3)$ и все сферы $\mathbb{S}_{2\pi m}^2$ радиусов $2\pi m$, $m = 1, 2, \dots$

Полная группа трехмерных вращений $\mathbb{O}(3)$ состоит из двух связных компонент: ортогональных матриц S_+ с положительным и S_- отрицательным определителем. Элементы полной группы вращений $\mathbb{O}(3)$ параметризуются элементом алгебры (1.131) следующим образом:

$$S_{\pm i}^j = \pm S_i^j \in \mathbb{O}(3). \quad (1.134)$$

То есть матрицы S_- получаются из матриц связной компоненты единицы умножением на оператор пространственного отражения $P := -\mathbb{1}$. При этом каждому элементу

алгебры $\mathfrak{so}(3)$ соответствуют два элемента группы $S_{\pm}(\omega) \in \mathbb{O}(3)$ – по одному из каждой компоненты.

Обратные матрицы собственных вращений имеют вид

$$S^{-1}_i{}^j(\omega) = S_i{}^j(-\omega) = \delta_i^j \cos \omega - \frac{(\omega \varepsilon)_i{}^j}{\omega} \sin \omega + \frac{\omega_i \omega^j}{\omega^2} (1 - \cos \omega) \in \mathbb{SO}(3), \quad (1.135)$$

т.е. соответствуют противоположному вектору вращения $-\omega \in \mathbb{R}^3$. Другими словами, обратный элемент группы $\mathbb{SO}(3)$ представляет собой поворот евклидова пространства вокруг той же оси, но на противоположный угол.

Свернем матрицы вращений $S_{\pm}$ с вектором $\{\omega_j\}$

$$S_{\pm i}{}^j \omega_j = \pm \omega_i.$$

Это значит, что вектор ω является собственным вектором матриц вращений с собственными значениями ± 1 . Другими словами, вращения S_+ и S_- оставляют, соответственно, ось вращения без изменения и меняют ее направление на противоположное.

Допустим, что элемент алгебры $\mathfrak{so}(3)$ (вектор поворота) дифференцируемо зависит от точки некоторого многообразия $\mathbb{M}$, т.е.

$$\mathbb{M} \ni x \mapsto \omega(x) \in \mathfrak{so}(3).$$

Каждому вектору поворота по формуле (1.132) ставится в соответствие матрица вращений $S \in \mathbb{SO}(3)$. Введем обозначение

$$l_{\alpha i}{}^j = (\partial_{\alpha} S_{\pm}^{-1} S_{\pm})_i{}^j = (\partial_{\alpha} S^{-1} S)_i{}^j, \quad (1.136)$$

где $\partial_{\alpha} := \partial/\partial x^{\alpha}$ – частная производная в локальной системе координат. Последнее равенство следует из (1.134). С помощью прямых вычислений можно убедиться в справедливости следующей формулы

$$\begin{aligned} l_{\alpha i}{}^j = & -\frac{(\partial_{\alpha} \omega \varepsilon)_i{}^j}{\omega} \sin \omega - \frac{\partial_{\alpha} \omega (\omega \varepsilon)_i{}^j}{\omega} \left(1 - \frac{\sin \omega}{\omega}\right) + \\ & + \frac{\partial_{\alpha} \omega_i \omega^j - \omega_i \partial_{\alpha} \omega^j}{\omega^2} (1 - \cos \omega) \in \mathfrak{so}(3). \end{aligned} \quad (1.137)$$

Эта матрица антисимметрична по своим индексам, и, значит, является элементом алгебры $\mathfrak{so}(3)$. Она представляет собой тривиальную $\mathbb{SO}(3)$ связность, для которой тензор кривизны тождественно равен нулю (чистая калибровка (13.71)). Если в качестве многообразия $\mathbb{M}$ выбрать многообразие самой группы вращений $\mathbb{SO}(3)$, то 1-форма $dx^{\alpha} l_{\alpha i}{}^j$ будет являться канонической 1-формой на $\mathbb{SO}(3)$ (см. раздел 8.2).

Построим гомоморфизм группы двумерных унитарных матриц $\mathbb{SU}(2)$ на группу собственных трехмерных вращений $\mathbb{SO}(3)$. Алгебра Ли $\mathfrak{su}(2)$ состоит из антиэрмитовых матриц $A = -A^{\dagger}$ с нулевым следом $\text{tr } A = 0$. Элемент алгебры $\mathfrak{su}(2)$ имеет вид

$$r^k \frac{i}{2} (\sigma_k)_{\text{A}}{}^{\text{B}} \in \mathfrak{su}(2), \quad \mathbf{r} = \{r^k\} \in \mathbb{R}^3, \quad (1.138)$$

где σ_k – матрицы Паули (см. приложение 37.2), и $\text{A}, \text{B} = 1, 2$ – матричные индексы. Множитель $i/2$ выбран с тем, чтобы коммутаторы векторов базиса $\frac{i}{2} \sigma_k$ имели вид (1.130) как и для алгебры $\mathfrak{so}(3)$. Это доказывает, что алгебры $\mathfrak{su}(2)$ и $\mathfrak{so}(3)$ изоморфны. Задать этот изоморфизм можно в явном виде. Для этого заметим, что

произвольную антиэрмитову матрицу можно взаимно однозначно представить в виде (1.138):

$$A = \frac{i}{2} \begin{pmatrix} r^3 & r^1 - ir^2 \\ r^1 + ir^2 & -r^3 \end{pmatrix}. \quad (1.139)$$

Предложение 1.8.2. *Отображение*

$$\varphi: \mathfrak{su}(2) \ni A \mapsto \varphi(A) := (\omega\epsilon) \in \mathfrak{so}(3), \quad (1.140)$$

где $\omega = \mathbf{r}$ из представления (1.139) и элемент алгебры $(\omega\epsilon)$ определен формулой (1.131), является изоморфизмом алгебр $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$.

Доказательство. Взаимная однозначность и линейность отображения φ очевидны. Сохранение коммутатора при отображении,

$$\varphi([A, B]) = [\varphi(A), \varphi(B)], \quad A, B \in \mathfrak{su}(2),$$

проверяется прямой проверкой. Это также следует из того, что коммутаторы базисных векторов в алгебрах совпадают. $\square$

Очевидно, что отображение φ является гладким. Заметим также, что

$$\det A = \frac{1}{4} ((\omega^1)^2 + (\omega^2)^2 + (\omega^3)^2).$$

Пусть $U \in \mathbb{SU}(2)$ – произвольная унитарная матрица, $UU^\dagger = U^\dagger U = 1$, с единичным определителем, $\det U = 1$. Построим новую матрицу

$$A' = UAU^\dagger.$$

Эта матрица также антиэрмитова, $A' = -A'^\dagger$, и ее след равен нулю, $\text{tr } A' = 0$. Поэтому матрицу A' также можно представить в виде

$$A' = \begin{pmatrix} r'^3 & r'^1 - ir'^2 \\ r'^1 + ir'^2 & -r'^3 \end{pmatrix}.$$

Поскольку $\det A' = \det A$, то

$$(\omega^1)^2 + (\omega^2)^2 + (\omega^3)^2 = (\omega'^1)^2 + (\omega'^2)^2 + (\omega'^3)^2$$

Следовательно, вектор вращения ω' получается из вектора ω с помощью некоторой ортогональной матрицы. Поскольку группа $\mathbb{SU}(2)$ связна, то из непрерывности следует, что каждой унитарной матрице $U \in \mathbb{SU}(2)$ однозначно ставится в соответствие некоторая ортогональная матрица $S \in \mathbb{SO}(3)$ из связной компоненты единицы. Полученное отображение гомоморфно, т.к. групповые операции, как нетрудно проверить, согласованы. Ядро гомоморфизма состоит из двух элементов $\{1, -1\}$, где 1 – двумерная единичная матрица. Это значит, что каждой ортогональной матрице $S \in \mathbb{SO}(3)$ соответствуют две унитарные матрицы U и $-U$ группы $\mathbb{SU}(2)$. Тем самым отображение $\mathbb{SU}(2) \rightarrow \mathbb{SO}(3)$ является двулистным накрытием, и это накрытие универсально, т.к. многообразие группы $\mathbb{SU}(2)$ связно и односвязно (см. раздел 11). Таким образом, справедлива

Теорема 1.8.1. *Существует гомоморфизм $\mathbb{SU}(2) \rightarrow \mathbb{SO}(3)$, при котором группа $\mathbb{SU}(2)$ является двулистным универсальным накрытием группы $\mathbb{SO}(3)$.*

Произвольную унитарную матрицу с единичным определителем $U \in \mathbb{SU}(2)$ можно параметризовать двумя комплексными числами $a, b \in \mathbb{C}$:

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad (1.141)$$

которые удовлетворяют условию $|a|^2 + |b|^2 = 1$. Это значит, что элементы группы $\mathbb{SU}(2)$ находятся во взаимно однозначном соответствии с точками трехмерной сферы $\mathbb{S}^3$, вложенной в четырехмерное евклидово пространство $\mathbb{R}^4$, т.е.

$$\mathbb{SU}(2) \approx \mathbb{S}^3. \quad (1.142)$$

Другими словами, трехмерную сферу можно оснастить групповой структурой. При гомоморфизме $\mathbb{SU}(2) \rightarrow \mathbb{SO}(3)$ два элемента U и $-U$ отображаются в один элемент группы вращений, и этим элементам соответствуют две противоположные точки на сфере $\mathbb{S}^3$. Это значит, что с топологической точки зрения группа $\mathbb{SO}(3)$ представляет собой сферу $\mathbb{S}^3$ с отождествленными диаметрально противоположными точками, т.е. проективную плоскость

$$\mathbb{SO}(3) \approx \mathbb{RP}^3 = \frac{\mathbb{S}^3}{\mathbb{Z}_2}.$$

Важно отметить, что из каждой пары элементов U и $-U$ невозможно выбрать по одному элементу группы $\mathbb{SU}(2)$ так, чтобы эти элементы образовали бы группу сами по себе. Это связано с тем, что при вращении на угол $2\pi t$, где $t \in [0, 1]$, при $t = 1$ мы имеем тождественное преобразование в группе $\mathbb{SO}(3)$, а матрицы $U \in \mathbb{SU}(2)$ меняют свой знак $U \rightarrow -U$. Другими словами, когда параметр t пробегает единичный отрезок мы имеем замкнутый путь в $\mathbb{SO}(3)$ и отрезок в группе $\mathbb{SU}(2)$. Это доказывает, в частности, что группа $\mathbb{SO}(3)$ не является подгруппой в $\mathbb{SU}(2)$.

Гомоморфизм $\mathbb{SU}(2) \rightarrow \mathbb{SO}(3)$ имеет следующую геометрическую интерпретацию. Каждое вращение трехмерного евклидова пространства $\mathbb{R}^3$ порождает вращение сферы $\mathbb{S}^2$ с центром в начале координат. Верно и обратное утверждение: каждое вращение сферы однозначно определяет вращение евклидова пространства. При стереографической проекции все точки сферы за исключением северного полюса проектируются на евклидову плоскость $\mathbb{R}^2$, на которой можно ввести комплексные координаты $z \in \mathbb{C} \approx \mathbb{R}^2$ (детали конструкции содержатся в [25]).

Теорема 1.8.2. *Каждому собственному вращению $S \in \mathbb{SO}(3)$ соответствуют два дробно линейных преобразования комплексных координат*

$$z' = \pm \frac{az + b}{-\bar{b}\bar{z} + \bar{a}}, \quad |a|^2 + |b|^2 = 1.$$

Обратно. Каждой унитарной матрице (1.141) соответствует некоторое вращение.

Доказательство. См., например, [25], §1.5. □

Замечание. Матрицу (1.141) можно рассматривать, как представление кватернионов (37.50) с единичной нормой. □

Отображение $U \mapsto S$ унитарной матрицы (1.141) в ортогональную матрицу $S \in \mathbb{SO}(3)$ можно записать в явном виде [5]

$$S = \begin{pmatrix} \frac{1}{2}(a^2 - b^2 + \bar{a}^2 - \bar{b}^2) & \frac{i}{2}(-a^2 - b^2 + \bar{a}^2 + \bar{b}^2) & -ab - \bar{a}\bar{b} \\ \frac{i}{2}(a^2 - b^2 - \bar{a}^2 + \bar{b}^2) & \frac{1}{2}(a^2 + b^2 + \bar{a}^2 + \bar{b}^2) & -i(ab - \bar{a}\bar{b}) \\ a\bar{b} + \bar{a}b & i(-a\bar{b} + \bar{a}b) & a\bar{a} - b\bar{b} \end{pmatrix}. \quad (1.143)$$

Выше было доказано, что группа $\mathrm{SU}(2)$ дважды покрывает группу вращений $\mathrm{SO}(3)$. Покажем, как это выглядит при использовании явной параметризации унитарной группы $\mathrm{SU}(2)$.

Предложение 1.8.3. Пусть элемент алгебры $\mathfrak{su}(2)$ имеет вид (1.138) с $\mathbf{r} = \boldsymbol{\omega}$. Тогда ему соответствует элемент группы $\mathrm{SU}(2)$

$$U_{\mathbf{A}}^{\mathbf{B}}(\boldsymbol{\omega}) = (\mathbf{e}^{i\boldsymbol{\omega}^k \sigma_k / 2})_{\mathbf{A}}^{\mathbf{B}} = \delta_{\mathbf{A}}^{\mathbf{B}} \cos \frac{\omega}{2} + i \frac{\omega^k \sigma_{k\mathbf{A}}^{\mathbf{B}}}{\omega} \sin \frac{\omega}{2} \in \mathrm{SU}(2), \quad (1.144)$$

где $\omega := \sqrt{\omega^i \omega_i}$ – модуль вектора вращения $\boldsymbol{\omega}$.

Доказательство. Прямая проверка путем разложения в ряды. □

Легко проверить равенство

$$U_{\mathbf{A}}^{-1\mathbf{B}}(\boldsymbol{\omega}) = U_{\mathbf{A}}^{\mathbf{B}}(-\boldsymbol{\omega}) = \delta_{\mathbf{A}}^{\mathbf{B}} \cos \frac{\omega}{2} - i \frac{\omega^k \sigma_{k\mathbf{A}}^{\mathbf{B}}}{\omega} \sin \frac{\omega}{2}.$$

Верно также и обратное утверждение: любая унитарная матрица с единичным определителем имеет вид (1.144) для некоторого вектора $\boldsymbol{\omega} \in \mathbb{R}^3$. Из вида элемента группы следует равенство

$$U_{\mathbf{A}}^{\mathbf{B}} \left(\omega^i + 4\pi \frac{\omega^i}{\omega} \right) = U_{\mathbf{A}}^{\mathbf{B}}(\omega^i).$$

То есть для группы $\mathrm{SU}(2)$ в пространстве параметров $\boldsymbol{\omega} \in \mathbb{R}^3$ (алгебре Ли) имеется отношение эквивалентности

$$\omega^i \sim \omega^i + 4\pi \frac{\omega^i}{\omega}. \quad (1.145)$$

При этом начало координат отождествляется со всеми сферами радиуса $4\pi m$, $m = 1, 2, \dots$. По сравнению с отношением эквивалентности (1.133) для группы $\mathrm{SO}(3)$ сдвиг происходит на вектор вдвое большей длины. Тем самым групповое многообразие параметризуется внутренними точками шара $\mathbb{B}_{2\pi}^3 \subset \mathbb{R}^3$ радиуса 2π . При этом все точки граничной сферы $\mathbb{S}_{2\pi}^3$ необходимо отождествить, т.к. второе слагаемое в (1.144) пропадает. Это приводит к второму отношению эквивалентности в пространстве параметров

$$\omega_1^i|_{\omega_1=2\pi} \sim \omega_2^i|_{\omega_2=2\pi}. \quad (1.146)$$

Дополнительное соотношение эквивалентности связано с отсутствием третьего слагаемого в (1.144) по сравнению с (1.132). В этом также есть отличие от группы вращений, для которой отождествляются только диаметрально противоположные точки граничной сферы, что входит в отношение эквивалентности (1.133). Можно показать, что в пространстве параметров для группы $\mathrm{SU}(2)$ других отношений эквивалентности кроме (1.145) и (1.146) не существует.

Выше было показано, что групповое многообразие унитарной группы $\mathrm{SU}(2)$ представляет собой трехмерную сферу $\mathbb{S}_{2\pi}^3$. Если начало координат в приведенной параметризации соответствует южному полюсу сферы $\mathbb{S}_{2\pi}^3$, то граничная сфера $\omega = 2\pi$ – северному. Это является следствием отношения эквивалентности (1.146).

Из гомоморфизма $\mathrm{SU}(2) \rightarrow \mathrm{SO}(3)$ следует гомоморфизм представлений этих групп. При этом каждой матрице представления группы вращений $\mathrm{SO}(3)$ соответствуют две матрицы представления группы $\mathrm{SU}(2)$, которые отличаются знаком, если представление точное. В физической литературе принято говорить, что точное представление

группы $\mathrm{SU}(2)$ является двузначным представлением группы вращений $\mathrm{SO}(3)$. Оно называется *спинорным представлением* группы $\mathrm{SO}(3)$. (См. также пример 9.4.2) Это представление двузначно, т.к. одному вращению $S \in \mathrm{SO}(3)$ соответствует два элемента унитарной группы $\pm U \in \mathrm{SU}(2)$. И эту неоднозначность нельзя устранить. Действительно, если вращению S поставить в соответствие только одну унитарную матрицу U , то после тождественного преобразования (поворота на угол $\omega = 2\pi$), матрица U изменит знак. Это следует из явного вида унитарного преобразования (1.144), т.к. аргумент тригонометрических функций равен половине угла поворота.

Поскольку группа вращений компактна, то все ее неприводимые представления унитарны и конечномерны. При этом любое другое ее представление разлагается в прямую сумму неприводимых представлений. Подробное описание всех представлений группы вращений содержится в [5, 25].

В заключение настоящего раздела скажем несколько слов о группе вращений $\mathbb{O}(n)$ n -мерного евклидова пространства $\mathbb{R}^n$. Под группой $\mathbb{O}(1) \approx \mathbb{Z}_2$ удобно понимать группу, состоящую из двух элементов $\{1, -1\}$, которые соответствуют тождественному преобразованию и отражению вещественной прямой относительно начала координат. При $n \geq 2$ группа $\mathbb{O}(n)$ сохраняет евклидову метрику и состоит из двух несвязных компонент при любом $n \geq 2$: матриц с положительным и отрицательным определителем. Матрица пространственных отражений $-\mathbb{1}$ при четных n имеет положительный определитель, принадлежит связной компоненте единицы $\mathrm{SO}(n)$ и не меняет ориентации декартовых осей. Поэтому преобразование четности в этом случае целесообразно определить, как отражение только одной, например, первой координаты

$$P = \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1} \end{pmatrix},$$

где $\mathbb{1}$ – единичная матрица размера $(n-1) \times (n-1)$.

1.9 Пространство Минковского

В настоящем разделе мы рассмотрим пространство-время Минковского произвольного числа измерений и его общие свойства. Затем остановимся на специфических свойствах пространств Минковского трех и четырех измерений, которые наиболее часто встречаются в физических приложениях.

Определение. *Пространством или пространством-временем Минковского $\mathbb{R}^{1,n-1}$ размерности n называется евклидово пространство $\mathbb{R}^n$, на котором в декартовых координатах задана лоренцева метрика*

$$\eta = \{\eta_{ab}\} = \mathrm{diag} \left(1, \underbrace{-1, \dots, -1}_{n-1} \right) \quad (1.147)$$

Декартовы системы координат в пространстве Минковского называются *инерциальными*. $\square$

Замечание. Выбор общего знака в метрике (1.147) является условным, и часто используется метрика противоположного знака, когда время входит в интервал со знаком минус, а пространственные координаты – со знаком плюс. $\square$

Таким образом в пространстве Минковского задано две метрики: евклидова и лоренцева. Евклидова метрика задается естественным образом на прямом произведении прямых и определяет топологию $\mathbb{R}^{1,n-1}$. Лоренцева метрика (1.147) не является положительно определенной, и ее нельзя использовать для определения расстояния в топологическом смысле (см. раздел 1.3.1). Поэтому не следует рассматривать пространство Минковского просто как прямое произведение прямых, на котором задана только метрика (1.147), поскольку евклидова метрика необходима для задания топологии. На самом деле это всегда подразумевается, т.к. непрерывность функций в пространстве Минковского понимается относительно естественной топологии евклидова пространства. Инерциальные координаты, в которых метрика имеет вид (1.147), обозначим $\{x^0, x^1, \dots, x^{n-1}\} = \{x^a\}$, $a = 0, 1, \dots, n-1$. Здесь первая координата $t := x^0$ называется временем. Остальные координаты x^i , $i = 1, \dots, n-1$, называются пространственноподобными и параметризуют пространственные сечения $x^0 = \text{const}$. Очевидно, что пространство Минковского есть прямое произведение

$$\mathbb{R}^{1,n-1} = \mathbb{R} \times \mathbb{R}^{n-1},$$

где первый сомножитель соответствует времени, а второй – $(n-1)$ -мерному евклидову пространству с отрицательно определенной метрикой.

Замечание. Здесь и в дальнейшем пространственные индексы будут нумероваться латинскими буквами из середины алфавита $\{i, j, k, \dots\}$. Мнемоническое правило следующее. Эти индексы образуют подмножество всего алфавита $\{a, b, c, \dots\}$, так же как множество чисел $\{1, 2, \dots, n-1\}$ является подмножеством чисел $\{0, 1, 2, \dots, n-1\}$. $\square$

Замечание. Между евклидовым пространством и пространством Минковского существует связь. Если время x^0 умножить на мнимую единицу $x^0 \mapsto ix^0$ и изменить общий знак метрики, то получим евклидову метрику. К тому же результату можно прийти, если на мнимую единицу умножить все пространственные координаты $x^i \mapsto ix^i$. Оба преобразования называются *комплексным поворотом*. При комплексном повороте оператор Даламбера переходит в оператор Лапласа, что приводит к качественному отличию решений уравнений, содержащих эти операторы. В приложениях важную роль играет связь между этими решениями. $\square$

Определение. В каждой точке пространства Минковского $y \in \mathbb{R}^{1,n-1}$ уравнение

$$\eta_{ab}(x^a - y^a)(x^b - y^b) = 0,$$

задает *световой конус* с вершиной в точке y (см. рис.1.10,а). Конус представляет собой объединение двух связных компонент: светового конуса *прошлого* $x^0 \leq y^0$ и *будущего* $x^0 \geq y^0$ с общей вершиной в точке y . $\square$

Световые конусы прошлого и будущего являются $(n-1)$ -мерными подмногообразиями в пространстве Минковского $\mathbb{R}^{1,n-1}$. В целом световой конус подмногообразия не является из-за общей вершины. Очевидно, что образующие световых конусов имеют нулевую длину.

Определение. Дифференцируемая кривая $\{x^a(s)\}$ в пространстве Минковского называется *временноподобной*, *пространственноподобной* или *светоподобной (нулевой)*, если касательный вектор, соответственно, удовлетворяет условиям:

$$\dot{x}^a \dot{x}^b \eta_{ab} > 0, \quad \dot{x}^a \dot{x}^b \eta_{ab} < 0, \quad \dot{x}^a \dot{x}^b \eta_{ab} = 0, \quad (1.148)$$

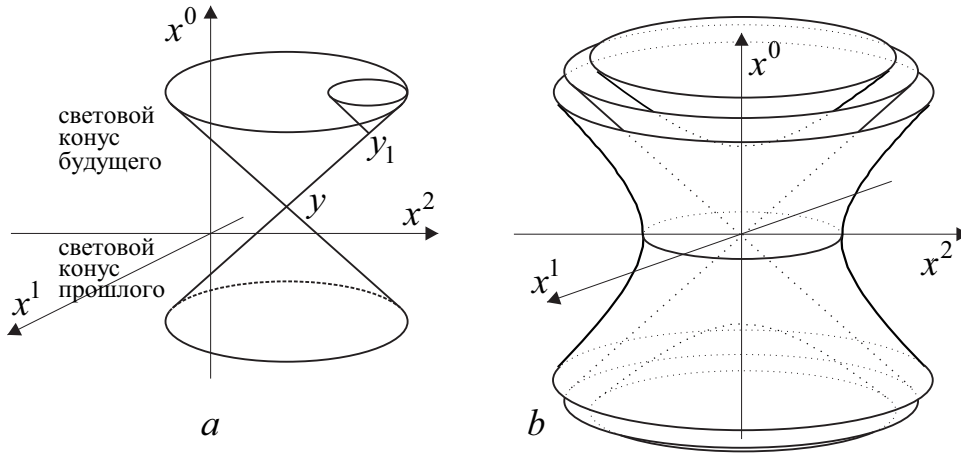


Рис. 1.10: Световой конус будущего и прошлого в трехмерном пространстве Минковского (а). Орбитами точек пространства Минковского, лежащих вне светового конуса, относительно действия собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, n-1)$ являются однополостные гиперболоиды (b).

для всех значений параметра $s \in \mathbb{R}$. Если в некоторой точке y , через которую проходит кривая, выполнены два условия $\dot{x}^a \dot{x}^b \eta_{ab} > 0$ и $\dot{x}^0 > 0$, то мы говорим, что касательный вектор к времениподобной кривой в точке y времениподобен и направлен в будущее. $\square$

Предложение 1.9.1. *Любая дифференцируемая времениподобная кривая в пространстве Минковского $\mathbb{R}^{1,n-1}$, проходящая через точку y , целиком лежит внутри светового конуса с вершиной в этой точке.*

Доказательство. Если кривая дифференцируема, времениподобна и проходит через точку y , то из дифференцируемости следует, что она будет лежать внутри светового конуса в некоторой окрестности точки y . Кроме того она не может касаться или пересекать световой конус, т.к. в такой точке она не была бы времениподобна. $\square$

Для полноты картины обобщим пространство Минковского на метрику произвольной сигнатуры.

Определение. Псевдоевклидовым пространством $\mathbb{R}^{p,q}$, $p+q=n$, называется евклидово пространство $\mathbb{R}^n$ с метрикой

$$\eta = \{\eta_{ab}\} = \text{diag}(\underbrace{+\dots+}_p, \underbrace{-\dots-}_q), \quad a, b = 1, \dots, n, \quad (1.149)$$

заданной в декартовой системе координат. $\square$

Многие свойства пространства Минковского естественным образом переносятся на псевдоевклидово пространство. При этом группа Лоренца $\mathbb{O}(1, n-1)$ заменяется на группу псевдовращений $\mathbb{O}(p, q)$, оставляющую метрику (1.149) инвариантной. Группы псевдовращений $\mathbb{O}(p, q)$ при $1 < p < n-1$ устроены сложнее группы Лоренца и в настоящей монографии рассматриваться не будут.

1.9.1 Группа Пуанкаре

Определение. Линейные неоднородные преобразования декартовых координат пространства Минковского $\mathbb{R}^{1,n-1}$, оставляющие инвариантной квадратичную форму

$$ds^2 = \eta_{ab} dx^a dx^b \quad (1.150)$$

образуют группу Ли, которая называется *группой Пуанкаре* и обозначается $\mathbb{IO}(1, n-1)$. Эта группа состоит из лоренцевых вращений с матрицами $S = \{S_b^a\}$ и сдвигов на вектор $a = \{a^a\}$:

$$x^a \mapsto x'^a = x^b S_b^a + a^a, \quad (1.151)$$

где постоянная матрица S является решением уравнения

$$\eta_{ab} = S_a^c S_b^d \eta_{cd}, \quad (1.152)$$

и $a \in \mathbb{R}^{1,n-1}$ – постоянный вектор. Множество матриц S образует группу Ли, которая называется *группой Лоренца* и обозначается $\mathbb{O}(1, n-1)$. Абелева подгруппа, параметризуемая вектором a , называется *группой трансляций*. $\square$

Замечание. В отличие от группы аффинных преобразований (1.60) матрицы S являются элементами группы Лоренца $\mathbb{O}(1, n-1)$, а не общей линейной группы $\mathbb{GL}(n, \mathbb{R})$. $\square$

Предложение 1.9.2. Любое (возможно, нелинейное) преобразование декартовых координат пространства Минковского, оставляющее квадратичную форму (1.150) инвариантной, является аффинным, т.е. имеет вид (1.151) с некоторой невырожденной матрицей S .

Доказательство. См., например, [13], теорема 2.4. $\square$

Это утверждение доказывает, что преобразования из группы Пуанкаре исчерпывают все возможные преобразования, сохраняющие метрику Лоренца (1.147). Другими словами, преобразования (1.152) являются движениями пространства Минковского общего вида.

Два последовательных преобразования S_1, a_1 и S_2, a_2 , которые отображают $x \mapsto x' \mapsto x''$ имеют вид

$$x''^a = x'^b S_{2b}^a + a_2^a = x^c (S_{1c}^b S_{2b}^a) + (a_1^b S_{2b}^a + a_2^a). \quad (1.153)$$

Группа Лоренца является подгруппой группы Пуанкаре и имеет размерность

$$\dim \mathbb{O}(1, n-1) = \frac{n(n-1)}{2}.$$

Она неабелева и некомпактна. Сдвиги образуют нормальную абелеву подгруппу группы Пуанкаре размерности n . При этом параметр сдвига a преобразуется как вектор относительно преобразований Лоренца. В целом группа Пуанкаре представляет собой полупрямое произведение группы Лоренца на подгруппу сдвигов и имеет размерность

$$\dim \mathbb{IO}(1, n-1) = \frac{n(n+1)}{2}.$$

Из определяющего уравнения (1.152) следует, что $\det S = \pm 1$. Кроме этого, 00 компоненту данного уравнения можно переписать в виде

$$(S_0^0)^2 = 1 + \sum_{i=1}^{n-1} (S_0^i)^2.$$

Заметим также, что из условия инвариантности обратной метрики Лоренца:

$$\eta^{ab} = \eta^{cd} S_c^a S_d^b,$$

которое эквивалентно (1.152), вытекает равенство

$$(S_0^0)^2 = 1 + \sum_{i=1}^{n-1} (S_i^0)^2, \quad (1.154)$$

где суммирование проходит по нижнему индексу. Отсюда вытекает, что для любой матрицы лоренцевых вращений либо $S_0^0 \geq 1$, либо $S_0^0 \leq -1$. Ясно, что никакую матрицу преобразований Лоренца с определителем $\det S = 1$ нельзя непрерывно деформировать в матрицу с определителем $\det S = -1$, т.к. множество из двух элементов $\{1, -1\}$ не является связным. Аналогично, матрицу с $S_0^0 \geq 1$ нельзя непрерывно деформировать в матрицу с $S_0^0 \leq -1$. Отсюда вытекает, что группа Лоренца состоит не менее, чем из четырех связных компонент.

Теорема 1.9.1. *Группа Лоренца $\mathcal{O}(1, n-1)$ при любом $n \geq 2$ состоит из четырех связных компонент.*

Доказательство. При $n = 2$ теорема была доказана в разделе 1.7. Доказательство для произвольного n содержится, например, в [13], теорема 2.5. $\square$

Определение. Преобразования Лоренца с положительным определителем называются *собственными*. Преобразования Лоренца, для которых $S_0^0 \geq 1$ называются *ортохронными*. $\square$

Предложение 1.9.3. *Преобразование Лоренца является ортохронным тогда и только тогда, когда оно всякий времениподобный вектор $X = \{X^0, X^1, \dots, X^{n-1}\}$ в начале координат, направленный в будущее, $X^0 > 0$, переводит во времениподобный вектор, который также направлен в будущее.*

Доказательство. Пусть $S : X^a \mapsto X'^a = X^b S_b^a$ – преобразование Лоренца. При этом времениподобный вектор переходит во времениподобный, т.к. квадрат вектора сохраняется при любом преобразовании Лоренца. Для произвольного времениподобного вектора выполнено неравенство

$$(X^0)^2 - (X^1)^2 - \dots - (X^{n-1})^2 > 0.$$

Применим неравенство Коши–Буняковского:

$$\begin{aligned} (X^1 S_1^0 + \dots + X^{n-1} S_{n-1}^0)^2 &\leq ((S_1^0)^2 + \dots + (S_{n-1}^0)^2) ((X^1)^2 + \dots + (X^{n-1})^2) \leq \\ &\leq ((S_0^0)^2 - 1) (X^0)^2 < (S_0^0)^2 (X^0)^2, \end{aligned}$$

где мы воспользовались равенством (1.154). Отсюда вытекает, что нулевая компонента преобразованного вектора

$$X'^0 := X^0 S_0^0 + X^1 S_1^0 + \dots + X^{n-1} S_{n-1}^0$$

имеет тот же знак, что и S_0^0 , если $X^0 > 0$. Отсюда же следует, что если $X^0 > 0$ и X'^0 , то $S_0^0 > 0$. $\square$

Определение. Введем оператор *обращения времени* T и *пространственного отражения* (четности) P :

$$\begin{aligned} T: & (x^0, x^i) \mapsto (-x^0, x^i), & \det T &= -1, \\ P: & (x^0, x^1, x^2, \dots, x^{n-1}) \mapsto (x^0, -x^1, x^2, \dots, x^{n-1}), & \det P &= -1. \end{aligned} \quad (1.155)$$

Введем также обозначение для их композиции

$$R := PT = TP: (x^0, x^1, x^2, \dots, x^{n-1}) \mapsto (-x^0, -x^1, x^2, \dots, x^{n-1}), \quad \det R = 1, \quad (1.156)$$

которая отражает только первые две координаты. $\square$

Замечание. Пространственные отражения при четных n можно было бы определить, как отражение всех пространственных координат, т.к. в этом случае $\det P = -1$, и ориентация осей координат меняется. Тогда оператор R соответствует полному отражению всех координат. При нечетных n отражение всех пространственных координат имеет положительный определитель и не подходит, т.к. принадлежит связной компоненте единицы группы. Отражение одной (любой) из координатных осей меняет ориентацию пространства Минковского и может быть использовано в качестве оператора четности. Такое определение подходит для пространства Минковского произвольного числа измерений. $\square$

Так же, как и в двумерном случае, полная группа Лоренца $\mathbb{O}(1, n-1)$ при $n \geq 3$ состоит из четырех связных компонент, которые получаются из связной компоненты единицы $\mathbb{S}_{+\uparrow}(1, n-1)$ обращением времени, пространственным отражением и их композицией:

$$\begin{aligned} \mathbb{S}_{-\downarrow} &= \mathbb{S}_{+\uparrow}(1, n-1)T, \\ \mathbb{S}_{-\uparrow} &= \mathbb{S}_{+\uparrow}(1, n-1)P, \\ \mathbb{S}_{+\downarrow} &= \mathbb{S}_{+\uparrow}(1, n-1)R. \end{aligned} \quad (1.157)$$

Единица полной группы Лоренца $\mathbb{O}(1, n-1)$ содержится в компоненте $\mathbb{S}_{+\uparrow}$. Эта компонента является связной группой Ли, называется *собственной ортохронной группой Лоренца* и обозначается $\mathbb{SO}_0(1, n-1)$. Кроме того, полная группа Лоренца $\mathbb{O}(1, n-1)$ содержит еще три подгруппы, состоящие из двух компонент каждая:

$$\begin{aligned} \mathbb{O}_{\downarrow}(1, n-1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{-\downarrow}, & \det S &= \pm 1, \\ \mathbb{O}_0(1, n-1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{-\uparrow}, & \det S &= \pm 1, \quad S_0^0 > 0, \\ \mathbb{SO}(1, n-1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{+\downarrow}, & \det S &= 1. \end{aligned} \quad (1.158)$$

Все четыре подгруппы являются нормальными подгруппами в $\mathbb{O}(1, n-1)$.

Очевидно, что отображения (1.157) взаимно однозначны и гладки. Поэтому как многообразия все четыре связные компоненты групп $\mathbb{O}(1, n-1)$ диффеоморфны между собой:

$$\mathbb{S}_{+\uparrow} \approx \mathbb{S}_{-\downarrow} \approx \mathbb{S}_{-\uparrow} \approx \mathbb{S}_{+\downarrow}.$$

Компоненты (1.157) являются смежными классами в группе $\mathbb{O}(1, n-1)$ по нормальной подгруппе $\mathbb{SO}_0(1, n-1)$. В качестве их представителей в фактор группе $\mathbb{O}(1, n-1)/\mathbb{SO}_0(1, n-1)$ можно выбрать матрицы T, P, R . Фактор группа $\mathbb{O}(1, n-1)/\mathbb{SO}_0(1, n-1)$ – это 4-группа Клейна $\mathbb{K}_4 = \{\mathbb{1}, T, P, R\}$, рассмотренная в разделе 1.7.

Рассмотрим автоморфизмы связной компоненты единицы группы Лоренца произвольной размерности. Пусть $S \in \mathbb{SO}_0(1, n-1)$, тогда отображение

$$S \mapsto S_0 S S_0^{-1},$$

где $S_0 \in \mathbb{SO}_0(1, n-1)$ – произвольный фиксированный элемент из связной компоненты единицы, задает внутренний автоморфизм группы собственных преобразований Лоренца. Отображения

$$S \mapsto T S T^{-1} \quad \text{и} \quad S \mapsto P S P^{-1}, \quad (1.159)$$

определяемые обращением времени и пространственным отражением, также задают автоморфизм собственных преобразований Лоренца. Однако этот автоморфизм будет внешним, т.к. преобразования T и P не принадлежат $\mathbb{SO}_0(1, n-1)$. Можно доказать, что внешние автоморфизмы (1.159) совпадают между собой с точностью до внутреннего автоморфизма². Чтобы найти явный вид внешнего автоморфизма, перепишем определение лоренцевых вращений (1.152), не делая различия между верхними и нижними индексами,

$$\eta = S \eta S^T \quad \text{или} \quad (S^T)^{-1} = \eta^{-1} S \eta.$$

Поскольку $\eta^{-1} = -T$, то отсюда следует формула для внешнего автоморфизма

$$T S T^{-1} = (S^T)^{-1}. \quad (1.160)$$

Конечно, внешние автоморфизмы связной компоненты единицы $\mathbb{SO}_0(1, n-1)$ являются внутренними для полной группы Лоренца $\mathbb{O}(1, n-1)$.

Замечание. Автоморфизм (1.160) играет большую роль при рассмотрении спинорных представлений полной группы Лоренца $\mathbb{O}(1, n-1)$, когда требуется определить оператор обращения времени и четности. $\square$

Орбитами точек пространства Минковского (см. рис.1.10, *b*), лежащих вне светового конуса, относительно действия собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, n-1)$ являются однополостные гиперboloиды, определяемые уравнением

$$(x^0)^2 - (x^1)^2 - \dots - (x^{n-1})^2 = -r^2, \quad r > 0.$$

Для точек, лежащих внутри конусов будущего и прошлого, орбитами являются, соответственно, верхняя и нижняя полы двуполостного гиперboloида

$$(x^0)^2 - (x^1)^2 - \dots - (x^{n-1})^2 = r^2, \quad r > 0.$$

Сами конусы будущего и прошлого представляют собой орбиты лежащих на них точек. Начало координат является неподвижной точкой относительно лоренцевых вращений. На каждой из орбит собственная группа Лоренца $\mathbb{SO}_0$ действует транзитивно, т.е. для любых двух точек, лежащих на одной орбите, найдется по крайней мере одно собственное ортохронное преобразование Лоренца, переводящее одну точку в другую. В целом действие группы $\mathbb{SO}_0(1, n-1)$ в пространстве Минковского $\mathbb{R}^{1, n-1}$ эффективно, но не свободно, т.к. начало координат является неподвижной точкой.

²Если в пространстве Минковского четной размерности оператор пространственного отражения определен, как отражение всех пространственных координат, то внешние автоморфизмы (1.159) просто совпадают.

Замечание. Инвариантность законов Природы относительно действия группы Пуанкаре является фундаментальным требованием к современным моделям математической физики и составляет основное содержание специальной теории относительности. Эта группа не является полупростой, т.к. содержит нормальную абелеву подгруппу (сдвиги). Следовательно, ее форма Киллинга–Картана вырождена, и это создает существенные трудности при построении физических моделей, т.к. на многообразии параметров группы Пуанкаре не существует двусторонне инвариантной невырожденной метрики. $\square$

Бесконечно малые преобразования группы Пуанкаре в пространстве Минковского $\mathbb{R}^{1,n-1}$ в линейном приближении можно записать с помощью дифференциальных операторов $L_{ab} = -L_{ba}$ и P_a :

$$\delta x^a = \left(\frac{1}{2} \delta \omega^{bc} L_{bc} + \delta a^b P_b \right) x^a, \quad (1.161)$$

где $\delta \omega^{ab} = -\delta \omega^{ba} \ll 1$ и $\delta a^a \ll 1$ – параметры преобразований Лоренца и сдвигов, соответственно. Множитель $\frac{1}{2}$ перед $\delta \omega^{bc} L_{bc}$ связан с тем, что независимые параметры входят в эту сумму дважды. Например,

$$\delta \omega^{01} L_{01} + \delta \omega^{10} L_{10} = 2 \delta \omega^{01} L_{01}.$$

Генераторы лоренцевых вращений и сдвигов представляются дифференциальными операторами (векторными полями Киллинга) на $\mathbb{R}^{1,n-1}$

$$L_{ab} = x_a \partial_b - x_b \partial_a, \quad (1.162)$$

$$P_a = \partial_a, \quad (1.163)$$

где подъем и опускание индексов осуществляется с помощью метрики Лоренца. При этом формула (1.161) переходит в равенство

$$\delta x^a = x^b \delta \omega_b^a + \delta a^a.$$

Дифференциальные операторы (1.162) и (1.163) действуют в алгебре функций $\mathcal{C}^1(\mathbb{R}^{1,n-1})$, заданных в пространстве Минковского $\mathbb{R}^{1,n-1}$. Например, изменение формы функции (см. раздел 2.13) при бесконечно малых преобразованиях в линейном приближении имеет вид

$$\delta f(x) := -\delta x^a \partial_a f(x) = -(x^b \delta \omega_b^a + \delta a^a) \partial_a f(x).$$

Матрица бесконечно малых вращений $S \in \mathbb{SO}_0(1, n-1)$ в линейном приближении имеет вид

$$S_b^a \approx \delta_b^a + \omega_b^a. \quad (1.164)$$

Отметим, что антисимметрия параметров лоренцевых вращений следует из уравнения (1.152) в линейном приближении. Каждый элемент антисимметричной матрицы ω^{ab} представляет собой угол бесконечно малого поворота в плоскости x^a, x^b . Лоренцевы вращения с параметрами ω^{0i} , $i = 1, \dots, n-1$ называют *бустом*, чтобы отличать их от чисто пространственных вращений, соответствующих параметрам ω^{ij} .

Генераторы вращений и сдвигов удовлетворяют алгебре Пуанкаре

$$[L_{ab}, L_{cd}] = -\eta_{ac} L_{bd} + \eta_{ad} L_{bc} + \eta_{bc} L_{ad} - \eta_{bd} L_{ac}, \quad (1.165)$$

$$[L_{ab}, P_c] = -\eta_{ac} P_b + \eta_{bc} P_a, \quad (1.166)$$

$$[P_a, P_b] = 0, \quad (1.167)$$

что проверяется прямой проверкой.

При действии вращений на векторы в пространстве Минковского $\mathbb{R}^{1,n-1}$ генераторы лоренцевых вращений, но не сдвиги, можно представить также в матричном виде

$$S_b^a \approx \delta_b^a + \frac{1}{2} \delta \omega^{cd} L_{cd}{}^a{}_b,$$

где

$$L_{cd}{}^a{}_b = \eta_{db} \delta_c^a - \eta_{cb} \delta_d^a. \quad (1.168)$$

Это – фундаментальное или векторное представление образующих алгебры Ли группы Лоренца. Нетрудно проверить, что матричное представление генераторов вращений также удовлетворяет алгебре Ли (1.165), которую можно записать в виде

$$[L_{ab}, L_{cd}] = f_{abcd}{}^{ef} L_{ef}, \quad (1.169)$$

где $f_{abcd}{}^{ef}$ – структурные константы группы Лоренца для коммутационных соотношений (1.165). Трансляции не имеют матричного представления в пространстве Минковского. Группа Лоренца при $n > 2$ является неабелевой и простой. Ее форма Киллинга–Картана невырождена

$$\eta_{abcd} := -f_{abef}{}^{gh} f_{cdgh}{}^{ef} = 4(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc}), \quad (1.170)$$

и поэтому может быть использована в качестве инвариантной метрики для построения инвариантов. Для матричного представления генераторов (1.168) справедливо равенство

$$\text{tr}(L_{ab}L_{cd}) = -\frac{1}{2}\eta_{abcd}. \quad (1.171)$$

Замечание. При $n = 2$ собственная ортохронная группа Лоренца $\mathbb{SO}_0(1,1)$ является абелевой, ее структурные константы равны нулю и форма Киллинга–Картана вырождена. Заметим, что правая часть (1.170) отлична от нуля и в этом случае. Формула для следа (1.171) справедлива также при $n = 2$, если под η_{abcd} понимать правую часть (1.170), а не форму Киллинга–Картана. $\square$

Группа Пуанкаре в n -мерном пространстве Минковского связана с группой псевдовращений $\mathbb{SO}(1, n)$ или $\mathbb{SO}(2, n-1)$ в $(n+1)$ -мерном пространстве следующим образом. Пусть индекс $A = \{0, 1, \dots, n-1, n\}$ нумерует координаты $(n+1)$ -мерного плоского пространства с метрикой

$$\{\eta_{AB}\} = \text{diag}(1, \underbrace{-1, \dots, -1}_{n-1}, \nu), \quad k > 0. \quad (1.172)$$

При $\nu = -1$ мы имеем пространство Минковского $\mathbb{R}^{1,n}$, а при $\nu = 1$ пространство $\mathbb{R}^{2,n-1}$. Обозначим генераторы лоренцевых вращений, затрагивающих n -тую координату, через $\tilde{P}_a := L_{an}$. Тогда алгебра псевдовращений в $(n+1)$ -мерном пространстве примет вид

$$\begin{aligned} [L_{ab}, L_{cd}] &= -\eta_{ac}L_{bd} + \eta_{ad}L_{bc} + \eta_{bc}L_{ad} - \eta_{bd}L_{ac}, \\ [L_{ab}, \tilde{P}_c] &= -\eta_{ac}\tilde{P}_b + \eta_{bc}\tilde{P}_a, \\ [\tilde{P}_a, \tilde{P}_b] &= -\nu L_{ab}. \end{aligned} \quad (1.173)$$

Эта алгебра проста и отличается от алгебры Пуанкаре тем, что “сдвиги” $\tilde{P}_a$ уже не коммутируют. Алгебра Пуанкаре получается после формального предела $\nu \rightarrow 0$, который называют *контракцией*. В этом пределе $\tilde{P} \rightarrow P$ и метрика (1.172) в $(n+1)$ -мерном пространстве вырождается.

1.9.2 Группа Галилея

Модели математической физики, которые инвариантны относительно действия группы Пуанкаре называются *релятивистскими*. Постулат о том, что физические модели должны быть инвариантны относительно группы Пуанкаре, лежит в основе *специальной теории относительности*, которая была предложена А. Эйнштейном в 1905 году [26]. До создания специальной теории относительности ньютонова механика точечных частиц рассматривалась как основная фундаментальная модель. В механике Ньютона важнейшую роль играют преобразования Галилея, которые также образуют группу симметрии пространства-времени.

Рассмотрим пространство-время Галилея $\mathbb{R}^1 \times \mathbb{R}^{n-1} = \mathbb{R}^n$, где первый сомножитель соответствует времени, $t = x^0 \in \mathbb{R}$, а второй – пространству. Пространственные декартовы координаты занумеруем буквами из середины латинского алфавита x^i , $i = 1, \dots, n-1$.

Определение. Зададим преобразования в $\mathbb{R}^1 \times \mathbb{R}^{n-1}$ следующими формулами:

$$\begin{aligned} t &\mapsto t' = t + a, \\ x &\mapsto x'^i = x^j S_j^i + v^i t + b^i, \end{aligned} \quad (1.174)$$

где $S_j^i \in \mathbb{O}(n-1)$ – матрица ортогональных вращений, действующая в пространстве. Преобразования (1.174) содержат дополнительные параметры: *скорость* $\{v^i\} \in \mathbb{R}^{n-1}$, сдвиги по времени $a \in \mathbb{R}$ и сдвиги пространства $\{b^i\} \in \mathbb{R}^{n-1}$. Эти преобразования образуют *группу Галилея* $\mathbb{G}(1, n-1)$. Декартовы системы координат, связанные преобразованием (1.174), как и в пространстве Минковского, называются *инерциальными*.

Подгруппа группы Галилея, для которой начало координат является неподвижной точкой, называется *однородной группой Галилея* и обозначается $\mathbb{HG}(1, n-1)$. Она состоит из пространственных вращений с матрицей S_j^i и преобразований, которые параметризуются вектором скорости v^i . Последние преобразования называются *галилеевыми бустами*. $\square$

Посчитаем размерность групп. Пространственные вращения параметризуются $(n-1)(n-2)/2$ параметрами. Галилеевы бусты – $(n-1)$ параметром. Кроме того, сдвиги времени и пространства задаются n параметрами. Таким образом, преобразования из полной группы Галилея задаются $n(n+1)/2$ параметрами, также как и преобразования из группы Пуанкаре. Размерность однородной группы Галилея равна размерности группы Лоренца,

$$\dim \mathbb{HG}(1, n-1) = \frac{n(n-1)}{2}.$$

Пусть в галилеевом пространстве-времени произошло два события в точках (t_1, x_1) и (t_2, x_2) . Эти события разделены во времени $\Delta t := t_2 - t_1$ и пространстве $\Delta x^i := x_2^i - x_1^i$. Тогда преобразования Галилея – это такие линейные неоднородные преобразования пространства-времени, которые оставляют инвариантными временной интервал Δt и расстояние между двумя одновременными событиями $\delta_{ij} \Delta x^i \Delta x^j$, при $t = \text{const}$.

Генераторы полной группы Галилея можно представить в виде дифференциальных операторов (векторных полей):

$$\begin{aligned} L_{ij} &= x_i \partial_j - x_j \partial_i, \\ P_i &= \partial_i, \\ B_i &= t \partial_i, \\ P_0 &= \partial_0. \end{aligned}$$

Эти генераторы удовлетворяют алгебре Галилея:

$$\begin{aligned}[L_{ij}, L_{kl}] &= -\delta_{ik}L_{jl} + \delta_{il}L_{jk} + \delta_{jk}L_{il} - \delta_{jl}L_{ik}, \\ [L_{ij}, P_k] &= -\delta_{ik}P_j + \delta_{jk}P_i, \\ [L_{ij}, B_k] &= -\delta_{ik}B_j + \delta_{jk}B_i, \\ [B_i, P_0] &= -P_i,\end{aligned}\tag{1.175}$$

где выписаны только отличные от нуля коммутаторы. Отсюда следует, что при пространственных вращениях генераторы галилеевых бустов B_i преобразуются, как ко-векторы $\{0, x_i\} \in \mathbb{R}^{1,n-1}$.

В этом представлении группа Галилея действует в алгебре функций $\mathcal{C}^1(\mathbb{R}^{1,n-1})$. При этом временная координата может быть сдвинута только на постоянный вектор. Механика Ньютона инварианта относительно преобразований Галилея. Поэтому говорят, что в механике Ньютона время имеет абсолютное значение.

Однородная группа Галилея представляет собой полупрямое произведение группы пространственных вращений $\mathbb{O}(n-1)$ на группу галилеевых бустов. Группа галилеевых бустов образует абелеву нормальную подгруппу однородной группы Галилея, которая, следовательно, не является полупростой. Отсюда следует, что в галилеевом пространстве $\mathbb{R} \times \mathbb{R}^{n-1}$ не существует метрики, инвариантной относительно преобразований Галилея.

1.9.3 Группа конформных преобразований

Рассмотрим пространство Минковского $\mathbb{R}^{1,n-1}$ с декартовыми координатами x^a , $a = 0, 1, \dots, n-1$. В настоящем разделе подъем и опускание индексов производится с помощью метрики Лоренца (1.147). Рассмотрим преобразования координат $x \mapsto x'(x)$ и ослабим требование инвариантности лоренцевой метрики (1.152), заменив его следующим условием

$$\eta_{ab}dx'^a dx'^b = \Omega^2(x)\eta_{ab}dx^a dx^b,\tag{1.176}$$

где $\Omega(x)$ – произвольная отличная от нуля функция (конформный множитель). Для определенности, будем считать, что $\Omega > 0$.

Определение. Преобразования координат пространства Минковского, при которых метрика Лоренца умножается на некоторый отличный от нуля множитель (1.176) образуют *конформную группу*. $\square$

При конформных преобразованиях длины векторов меняются, а углы между ними сохраняются.

Решения уравнения (1.176) для функций преобразования координат $x'(x)$ зависят от размерности пространства-времени.

В двумерном пространстве-времени, $n = 2$, уравнение (1.176) удобно переписать в светоподобных координатах (1.93)

$$du'dv' = \Omega^2(u, v)dudv.$$

Ясно, что конформные преобразования

$$u \mapsto u'(u), \quad v \mapsto v'(v)\tag{1.177}$$

где $u'(u)$ и $v'(v)$ – произвольные монотонные функции с отличными от нуля производными, удовлетворяют поставленному условию. Действительно,

$$du'dv' = \frac{du'}{du} \frac{dv'}{dv} dudv.$$

То есть для конформных преобразований

$$\Omega(u, v) = \frac{du' dv'}{du dv}.$$

Кроме того, двумерные преобразования координат из группы Пуанкаре также удовлетворяют равенству (1.176). Для этих преобразований $\Omega = 1$. Можно доказать и обратное утверждение. Преобразования из группы Пуанкаре и конформные преобразования (1.177) исчерпывают все возможные преобразования координат, для которых выполнено равенство (1.176).

Группа конформных преобразований (1.177) является бесконечномерной группой Ли.

Теперь рассмотрим случай $n \geq 3$.

Нетрудно доказать, что преобразования пространства Минковского (возможно, нелинейные), удовлетворяющие условию (1.176), образуют группу Ли.

Предложение 1.9.4. *Конформная группа пространства Минковского при $n \geq 3$ состоит из подгруппы преобразований из группы Пуанкаре (1.151), для которых $\Omega = 1$, специальных конформных преобразований, которые параметризуются постоянным вектором $b = \{b^a\} \in \mathbb{R}^{1, n-1}$,*

$$x'^a = \frac{x^a + b^a x^2}{1 + 2bx + b^2 x^2}, \quad b \neq 0, \quad \Omega = \frac{1}{1 + 2bx + b^2 x^2}, \quad (1.178)$$

где $bx := b^a x_a$, $b^2 := b^a b_a$, $x^2 := x^a x_a$, и дилатаций

$$x'^a = \Omega x^a, \quad \Omega = \text{const} \neq 1. \quad \square \quad (1.179)$$

Доказательство. См., например, [27], теорема §8.5. □

Дилатации называются также *гомотетией*.

В монографии [27], §8.5 описаны все (псевдо)римановы многообразия с нетривиальной конформной группой движений.

Специальные конформные преобразования (1.178) определены при

$$1 + 2bx + b^2 x^2 \neq 0.$$

Нетрудно проверить, что из формулы преобразования координат (1.178) следует равенство

$$x'^2 = x^2 \frac{1}{1 + 2bx + b^2 x^2}.$$

Если $x^2 \neq 0$, то отсюда вытекает правило

$$\frac{x'^a}{x'^2} = \frac{x^a}{x^2} + b^a, \quad x^2 \neq 0. \quad (1.180)$$

То есть специальные конформные преобразования – это сдвиг “обращенных” координат

$$x^a \mapsto \frac{x^a}{x^2},$$

для которых бесконечно удаленная точка отображается в начало координат.

Генераторы бесконечно малых специальных конформных преобразований K_a и дилатаций D можно представить в виде дифференциальных операторов (векторных полей) на $\mathbb{R}^{1,n-1}$:

$$K_a = x^2 \partial_a - 2x_a x^b \partial_b, \quad (1.181)$$

$$D = x^a \partial_a. \quad (1.182)$$

Эти векторные поля не являются полями Киллинга, т.к. при соответствующих преобразованиях метрика меняется. Прямые вычисления с учетом представления (1.162), (1.163) для генераторов группы Пуанкаре приводят к следующим коммутационным соотношениям

$$\begin{aligned} [L_{ab}, K_c] &= \eta_{ac} K_b - \eta_{bc} K_a, \\ [P_a, K_b] &= -2\eta_{ab} D - 2L_{ab}, \\ [P_a, D] &= P_a, \\ [K_a, D] &= -K_a, \\ [L_{ab}, D] &= [K_a, K_b] = [D, D] = 0. \end{aligned} \quad (1.183)$$

Эти соотношения вместе с коммутационными соотношениями (1.165)–(1.167) для группы Пуанкаре показывают, что генераторы L_{ab} , P_a , K_a и D образуют алгебру Ли, соответствующую конформной группе преобразований. Она имеет размерность

$$\frac{1}{2}(n+1)(n+2).$$

В отличие от группы Пуанкаре группа конформных преобразований проста и ее связная компонента единицы изоморфна связной компоненте единицы группы псевдовращений $\mathbb{SO}(2, n)$, действующей в пространстве $\mathbb{R}^{2,n}$, размерность которого на два превышает размерность исходного пространства Минковского $\mathbb{R}^{1,n-1}$. Изоморфизм алгебр Ли можно задать следующим образом

$$L_{AB} = \begin{pmatrix} \frac{1}{2}L_{ab} & \frac{1}{4}(P_a + K_a) & \frac{1}{4}(P_a - K_a) \\ -\frac{1}{4}(P_b + K_b) & 0 & \frac{1}{2}D \\ -\frac{1}{4}(P_b - K_b) & -\frac{1}{2}D & 0 \end{pmatrix}, \quad (1.184)$$

где индексы A, B пробегает значения $A = \{a, n, n+1\} = \{0, 1, \dots, n+1\}$, и метрика в пространстве $\mathbb{R}^{2,n}$ имеет вид

$$\eta_{AB} = \begin{pmatrix} \eta_{ab} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Поскольку конформная группа проста, то на групповом многообразии существует двусторонне инвариантная метрика. Это – существенное отличие от группы Пуанкаре.

Оператор Даламбера в n -мерном пространстве Минковского

$$\square = \partial_0^2 - \sum_{i=1}^{n-1} \partial_i^2$$

инвариантен относительно группы Пуанкаре $\mathbb{O}(1, n-1)$, причем сдвиги действуют на оператор дифференцирования тривиально $\partial_a \mapsto \partial_a$. При специальных конформных преобразованиях и дилатациях оператор Даламбера умножается на конформный множитель:

$$\square = \Omega^2 \square',$$

где $\square' := \eta^{ab} \partial'_a \partial'_b$. Это означает, что, если некоторая функция удовлетворяет уравнению Даламбера $\square f = 0$, то она будет удовлетворять уравнению Даламбера после произвольного конформного преобразования координат.

Замечание. Конформные преобразования были введены в физику в 1909 году Куннигамом и Бейтменом [28, 29], которые показали, что уравнения Максвелла ковариантны относительно конформных преобразований. Позже Дирак показал конформную инвариантность уравнения для нейтрино (безмассового уравнения Дирака) [30]. В настоящее время принято считать, что конформная инвариантность играет большую роль в асимптотических режимах, где отсутствуют размерные параметры. Подробное изложение конформной квантовой теории поля можно найти в монографии [31]. $\square$

Замечание. Конформные преобразования координат можно рассматривать также в евклидовом пространстве. Для евклидовой плоскости конформные преобразования (1.177) заменяются на конформные преобразования комплексной плоскости $z \mapsto w(z)$. При $n \geq 3$ связная компонента единицы группы конформных преобразований изоморфна группе Лоренца $\mathbb{SO}_0(1, n+1)$.

1.9.4 Трехмерное пространство Минковского

Рассмотрим трехмерное пространство Минковского $\mathbb{R}^{1,2}$ с декартовыми координатами x^a , $a = 0, 1, 2$, и метрикой Лоренца $\{\eta_{ab}\} = \text{diag}(+ - -)$. В трехмерном случае группа Лоренца $\mathbb{O}(1, 2)$ имеет всего три независимых генератора: L_{01} , L_{02} и L_{12} . Первые два генератора соответствуют лоренцевым бустам, а третий – пространственным вращениям в плоскости x^1, x^2 . Генераторы группы Лоренца удовлетворяют алгебре (1.165), которая в рассматриваемом случае существенно упрощается:

$$\begin{aligned} [L_{01}, L_{02}] &= L_{12}, \\ [L_{01}, L_{12}] &= L_{02}, \\ [L_{02}, L_{12}] &= -L_{01}, \end{aligned} \tag{1.185}$$

а все остальные коммутаторы равны нулю.

Используя полностью антисимметричный тензор третьего ранга ε^{abc} , $\varepsilon^{012} = 1$, (см. приложение 37.4) удобно перейти к дуальному базису алгебры Ли $\mathfrak{so}(1, 2)$

$$M^a := \frac{1}{2} \varepsilon^{abc} L_{bc}, \quad L_{bc} = \varepsilon_{bca} M^a,$$

где подъем и опускание индексов осуществляется с помощью метрики Лоренца. При этом генератор $M^0 = L_{12}$ соответствует вращениям в пространственной плоскости, а генераторы $M^1 = -L_{02}$ и $M^2 = L_{01}$ – лоренцевым бустам, соответственно, в плоскостях x^0, x^2 и x^0, x^1 . Нетрудно проверить, что новый базис удовлетворяет коммутационным соотношениям

$$[M^a, M^b] = \varepsilon^{abc} M_c.$$

В отличие от алгебры Ли группы трехмерных вращений (1.130), подъем и опускание индексов в рассматриваемом случае осуществляется не с помощью евклидовой метрики, а с помощью метрики Лоренца: $M_a := \eta_{ab} M^b$ ($M_0 = M^0$, $M_{1,2} = -M^{1,2}$).

Группа Лоренца $\mathbb{O}(1, 2)$ неабелева и некомпактна.

Собственная ортохронная группа Лоренца $\mathbb{SO}_0(1, 2)$ имеет двумерные неабелевы некомпактные подгруппы. В качестве такой подгруппы можно выбрать группу, генерируемую, например, базисными векторами M_1 и $K = M_0 + M_2$, которые соответствуют бусту и вращению в перпендикулярной нулевой плоскости. Эти векторы образуют базис двумерной неабелевой алгебры Ли $\mathfrak{g} \subset \mathfrak{so}(1, 2)$ и удовлетворяют следующим коммутационным соотношениям

$$\begin{aligned} [M_1, M_1] &= 0, & [K, K] &= 0, \\ [M_1, K] &= K. \end{aligned}$$

Это алгебра генерирует простейшую двумерную неабелеву группу Ли $\mathbb{G}$ (группу аффинных преобразований прямой). Ее групповое многообразие некомпактно и подробно изучено в разделе 8.7. С топологической точки зрения эта подгруппа диффеоморфна евклидовой плоскости $\mathbb{R}^2$. Любое преобразование из группы $\mathbb{SO}_0(1, 2)$ можно представить в виде композиции некоторого преобразования из данной подгруппы и некоторого пространственного вращения из подгруппы $\mathbb{SO}(2) \subset \mathbb{SO}_0(1, 2)$. Поэтому в соответствии с общей теоремой 8.11.10 как многообразие трехмерная группа Лоренца диффеоморфна прямому произведению

$$\mathbb{SO}_0(1, 2) \approx \mathbb{S}^1 \times \mathbb{R}^2,$$

т.к. $\mathbb{SO}(2) \approx \mathbb{S}^1$. Этот диффеоморфизм описан явно в примере 8.15.

Поскольку трехмерная группа Лоренца $\mathbb{SO}_0(1, 2)$ является подгруппой всех групп Лоренца более высоких размерностей $\mathbb{SO}_0(1, n-1)$, $n > 3$, то соответствующая двумерная подгруппа является подгруппой всех групп Лоренца $\mathbb{G} \subset \mathbb{SO}_0(1, n-1)$, $n \geq 3$.

Построим гомоморфизм группы вещественных 2×2 матриц с единичным определителем $\mathbb{SL}(2, \mathbb{R})$ на связную компоненту единицы трехмерной группы Лоренца $\mathbb{SO}_0(1, 2)$. Для этого рассмотрим произвольную симметричную 2×2 матрицу A (поскольку матрица вещественна, то она также эрмитова), которую параметризуем следующим образом

$$A = \begin{pmatrix} t+x & y \\ y & t-x \end{pmatrix}, \quad t, x, y \in \mathbb{R}. \quad (1.186)$$

Рассмотрим числа t, x, y как декартовы координаты в пространстве Минковского $\mathbb{R}^{1,2}$. Поскольку

$$\det A = t^2 - x^2 - y^2,$$

то определитель $\det A$ задает квадратичную форму Лоренца в $\mathbb{R}^{1,2}$.

Пусть $M \in \mathbb{SL}(2, \mathbb{R})$ – произвольная вещественная 2×2 матрица с единичным определителем. Тогда каждому преобразованию

$$A \mapsto A' = MAM^T \quad (1.187)$$

можно сопоставить некоторое преобразование Лоренца в пространстве Минковского. Действительно, при таком преобразовании матрица A' остается симметричной, и ее можно представить в виде

$$A' = \begin{pmatrix} t' + x' & y' \\ y' & t' - x' \end{pmatrix}$$

с некоторыми новыми координатами t', x', y' . Поскольку при преобразовании (1.187) определитель не меняется,

$$\det A' = \det A,$$

то каждому элементу $M \in \mathrm{SL}(2, \mathbb{R})$ однозначно ставится в соответствие элемент из группы Лоренца $\mathrm{SO}_0(1, 2)$. При этом групповые операции, как нетрудно видеть, согласованы. Обратно, по заданному преобразованию Лоренца уравнение (1.187) определяет матрицу M , имеющую единичный определитель, с точностью до знака. Поскольку группа $\mathrm{SL}(2, \mathbb{R})$ связна, то это устанавливает изоморфизм групп

$$\mathrm{SO}_0(1, 2) \simeq \frac{\mathrm{SL}(2, \mathbb{R})}{\mathbb{Z}_2}. \quad (1.188)$$

Отображение $\mathrm{SL}(2, \mathbb{R}) \rightarrow \mathrm{SO}_0(1, 2)$ является двулистным накрытием. Группа $\mathrm{SL}(2, \mathbb{R})$ как многообразие диффеоморфна прямому произведению $\mathbb{S}^1 \times \mathbb{R}^2$ (см. пример 8.15). Поэтому она является связной, но не односвязной. Следовательно, накрытие (1.188) не является универсальным. Универсальная накрывающая для групп $\mathrm{SO}_0(1, 2)$ и $\mathrm{SL}(2, \mathbb{R})$ построена в разделе 8.15.

Произвольную 2×2 матрицу можно параметризовать четырьмя числами:

$$M = \begin{pmatrix} a - d & b - c \\ b + c & a + d \end{pmatrix}, \quad a, b, c, d \in \mathbb{R}.$$

Для того, чтобы она принадлежала группе $\mathrm{SL}(2, \mathbb{R})$ на нее необходимо наложить условие

$$\det M = a^2 - d^2 - b^2 + c^2 = 1. \quad (1.189)$$

Таким образом, как многообразие, группа $\mathrm{SL}(2, \mathbb{R})$ представляет собой “гиперболоид” (1.189), вложенный в четырехмерное евклидово пространство $\mathbb{R}^4$.

Построим явную параметризацию элементов собственной ортохронной группы Лоренца $\mathrm{SO}_0(1, 2)$ элементами ее алгебры (экспоненциальное отображение). Для этого рассмотрим произвольный элемент алгебры $X \in \mathfrak{su}(2)$ (1.138)

$$X = \omega^k \frac{i}{2} \sigma_k = \frac{i}{2} \begin{pmatrix} \omega^3 & \omega^1 - i\omega^2 \\ \omega^1 + i\omega^2 & -\omega^3 \end{pmatrix}$$

где σ_k , $k = 1, 2, 3$, – матрицы Паули. Положим

$$\omega^1 = -2ix, \quad \omega^2 = -2t, \quad \omega^3 = -2iy, \quad (1.190)$$

где t, x, y – некоторые вещественные числа, тогда матрица X примет вид

$$X = \begin{pmatrix} y & x - t \\ x + t & -y \end{pmatrix}, \quad t, x, y \in \mathbb{R}.$$

Здесь мы используем те же буквы, что и в параметризации (1.186), но они имеют другой смысл. Полученная матрица вещественна, симметрична и имеет нулевой след. То есть представляет собой элемент алгебры $\mathfrak{sl}(2, \mathbb{R})$ общего вида, т.к. параметризуется тремя параметрами.

Поскольку экспонента от произвольной комплексной $n \times n$ матрицы равномерно сходится в любой ограниченной области комплексного пространства $\mathbb{C}^{n^2}$, то экспоненциальное отображение (1.144) можно переписать для группы $\mathrm{SL}(2, \mathbb{R})$, просто произведя замену координат (1.190). Поскольку

$$\sin(ia) = i \operatorname{sh} a, \quad \cos(ia) = \operatorname{ch} a, \quad a \in \mathbb{R},$$

то преобразованные матрицы выглядят по-разному в различных точках пространства Минковского $(t, x, y) \in \mathbb{R}^{1,2}$. Внутри светового конуса $t^2 > x^2 + y^2$ экспоненциальное отображение имеет вид

$$M = e^X = \begin{pmatrix} \cos s + \frac{y}{s} \sin s & \frac{x-t}{s} \sin s \\ \frac{x+t}{s} \sin s & \cos s - \frac{y}{s} \sin s \end{pmatrix} \in \mathbb{SL}(2, \mathbb{R}), \quad (1.191)$$

где $s := \sqrt{t^2 - x^2 - y^2}$. На световом конусе $t^2 = x^2 + y^2$

$$M = e^X = \begin{pmatrix} 1+y & x-t \\ x+y & 1-y \end{pmatrix} \in \mathbb{SL}(2, \mathbb{R}). \quad (1.192)$$

Вне светового конуса

$$M = e^X = \begin{pmatrix} \operatorname{ch} r + \frac{y}{r} \operatorname{sh} r & \frac{x-t}{r} \operatorname{sh} r \\ \frac{x+t}{r} \operatorname{sh} r & \operatorname{ch} r - \frac{y}{r} \operatorname{sh} r \end{pmatrix} \in \mathbb{SL}(2, \mathbb{R}), \quad (1.193)$$

где $r := \sqrt{x^2 + y^2 - t^2}$. Матрицы (1.191) и (1.193) имеют одинаковый предел при $s \rightarrow 0$ и $r \rightarrow 0$, который совпадает с матрицей (1.192).

Экспоненциальное отображение

$$\mathfrak{sl}(2, \mathbb{R}) \ni X \mapsto M \in \mathbb{SL}(2, \mathbb{R})$$

не является однозначным, и многим элементам алгебры соответствует один и тот же элемент группы. То есть групповое многообразие $\mathbb{SO}_0(1, 2)$ получается из $\mathbb{R}^3$ после отождествления некоторых точек. Например, единице группы соответствует начало координат и все двуполостные гиперболоиды

$$s = 2\pi m, \quad m = 1, 2, \dots$$

Из вида матрицы (1.191) следует, что внутри светового конуса существует отношение эквивалентности

$$x^a \sim x^a + 2\pi \frac{x^a}{s}.$$

Как и в случае группы $\mathbb{SU}(2)$ существует второе отношение эквивалентности. А именно, необходимо отождествить также точки, лежащие на гиперболоиде $s = \pi$, т.к. всем этим точкам соответствует одна и та же матрица $-\mathbb{1}$.

1.9.5 Четырехмерное пространство Минковского

Обсудим ряд специфических свойств группы Лоренца в четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$, которые часто используются в физических приложениях.

Предложение 1.9.5. В четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$ любая дифференцируемая изотропная кривая $\{x^a(s)\}$, определяемая уравнением

$$\eta_{ab} \dot{x}^a \dot{x}^b = 0, \quad (1.194)$$

представима в виде

$$\begin{aligned} x^0 &= \int^s du R, & x^2 &= \int^s du R \sin \theta \sin \varphi, \\ x^1 &= \int^s du R \sin \theta \cos \varphi, & x^3 &= \int^s du R \cos \theta, \end{aligned} \quad (1.195)$$

где каждая из координат определена с точностью до постоянной, а $R(u)$, $\theta(u)$ и $\varphi(u)$ – произвольные непрерывные функции от u . Эти линии являются прямыми тогда и только тогда, когда функции R, θ, φ постоянны.

Доказательство. Явное решение алгебраического уравнения (1.194) относительно производных и последующее интегрирование. $\square$

Отсюда следует, что, в отличие от плоскости Минковского, класс изотропных кривых состоит не только из прямых линий. Поскольку все экстремали (см. главу 16) в пространстве Минковского и только они являются прямыми, то это доказывает, что не всякая изотропная кривая является экстремалью. Исключение составляет только двумерное пространство Минковского, где все изотропные кривые являются прямыми и, следовательно, экстремальями.

Предложение 1.9.6. *Любая изотропная кривая в пространстве Минковского, проходящая через точку y , не может выйти за пределы светового конуса с вершиной в этой точке.*

Доказательство. Допустим, что в некоторой точке y_1 , лежащей на световом конусе будущего и не совпадающей с y , изотропная кривая покидает световой конус (см. рис.1.10). Тогда в некоторой окрестности данной точки условие (1.194) будет нарушено. Это ясно из того, что световой конус будущего с вершиной в точке y_1 касается светового конуса будущего в точке y и целиком лежит внутри него. $\square$

Четырехмерное пространство Минковского, $\mathbb{R}^{1,3} = \mathbb{R}^{1,1} \times \mathbb{R}^2$, можно представить как прямое произведение двумерной плоскости Минковского $\mathbb{R}^{1,1}$ и евклидовой плоскости $\mathbb{R}^2$ с координатами (x^0, x^1) и (x^2, x^3) , соответственно. При таком разбиении можно ввести координаты светового конуса (1.93) на $\mathbb{R}^{1,1}$ и комплексные координаты (1.87) на $\mathbb{R}^2$. Тогда лоренцев интервал примет вид

$$ds^2 = dudv - dzd\bar{z}, \quad (1.196)$$

где $u := x^0 + x^1$, $v := x^0 - x^1$, $z := x^2 + ix^3$ и $\bar{z} := x^2 - ix^3$. Соответствующая метрика в координатах $\{x^a\} = \{u, v, z, \bar{z}\}$ имеет вид

$$\eta = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

В моделях гравитации используется комплексная изотропная тетрада, которая соответствует введенным выше координатам в касательном пространстве к пространству-времени.

Коммутационные соотношения, определяющие алгебру Пуанкаре (1.165)–(1.167), в четырехмерном пространстве-времени можно переписать в другом, эквивалентном, виде, выделив явно пространственные и временные компоненты. Среди координат пространства Минковского, как и ранее, явно выделяем время $\{x^a\} = \{x^0, x^i\}$, $i = 1, 2, 3$. Группа Пуанкаре содержит подгруппу трехмерных вращений, алгебра которой определяется генераторами L_{ij} . В связи с этим вместо генераторов L_{ab} можно рассматривать 6 генераторов

$$J_i := \frac{1}{2} \varepsilon_i^{jk} L_{jk}, \quad N_i := L_{0i},$$

где подъем и опускание латинских индексов $i, j = 1, 2, 3$ осуществляется с помощью евклидовой метрики δ_{ij} . Образующие алгебры Ли группы Лоренца J_i и N_i представляют собой генераторы пространственных вращений и бустов. Нетрудно проверить, что они удовлетворяют следующим коммутационным соотношениям

$$\begin{aligned} [J_i, J_j] &= -\varepsilon_{ij}^{k} J_k, \\ [N_i, N_j] &= \varepsilon_{ij}^{k} J_k, \\ [J_j, N_j] &= -\varepsilon_{ij}^{k} N_k. \end{aligned} \quad (1.197)$$

Последнее коммутационное соотношение показывает, что базисные векторы N_i преобразуются как компоненты ковектора при пространственных вращениях. Остальные коммутационные соотношения (1.165), (1.166) примут вид

$$\begin{aligned} [J_i, P_0] &= 0, \\ [J_i, P_j] &= -\varepsilon_{ijk} P^k, \\ [N_i, P_0] &= P_i, \\ [N_i, P_j] &= \delta_{ij} P_0. \end{aligned} \quad (1.198)$$

Построим гомоморфизм группы комплексных 2×2 -матриц с единичным определителем $\mathrm{SL}(2, \mathbb{C})$ на связную компоненту единицы группы Лоренца $\mathrm{SO}_0(1, 3)$. Это построение аналогично доказательствам существования накрытий $\mathrm{SU}(2) \rightarrow \mathrm{SO}(3)$ и $\mathrm{SL}(2, \mathbb{R}) \rightarrow \mathrm{SO}_0(1, 2)$, проведенным в разделах 1.8 и 1.9.4. Для построения гомоморфизма заметим, что произвольную эрмитову матрицу $A = A^\dagger$ можно представить в виде

$$A = \begin{pmatrix} t + z & x - iy \\ x + iy & t - z \end{pmatrix} = t + x\sigma_1 + y\sigma_2 + z\sigma_3, \quad t, x, y, z \in \mathbb{R}, \quad (1.199)$$

где σ_1, σ_2 и σ_3 – матрицы Паули. Рассмотрим числа t, x, y, z как координаты в пространстве-времени Минковского $\mathbb{R}^{1,3}$. Очевидно, что

$$\det A = t^2 - x^2 - y^2 - z^2.$$

Пусть $M \in \mathrm{SL}(2, \mathbb{C})$ – произвольная комплексная матрица с единичным определителем. Тогда каждому преобразованию

$$A \mapsto A' = MAM^\dagger, \quad (1.200)$$

можно однозначно поставить в соответствие некоторое собственное ортохронное преобразование Лоренца. Действительно, при таком преобразовании матрица A' остается эрмитовой, и ее можно записать в виде

$$A' = \begin{pmatrix} t' + z' & x' - iy' \\ x' + iy' & t' - z' \end{pmatrix}$$

с некоторыми новыми координатами t', x', y', z' в $\mathbb{R}^{1,3}$. Поскольку при этом определитель матрицы не меняется,

$$\det A' = \det A,$$

то каждому элементу $M \in \mathrm{SL}(2, \mathbb{C})$ однозначно ставится в соответствие элемент из группы Лоренца $\mathrm{SO}_0(1, 3)$. Прямая проверка показывает, что отображение

$$\mathrm{SL}(2, \mathbb{C}) \rightarrow \mathrm{SO}_0(1, 3) \quad (1.201)$$

является гомоморфизмом групп. Обратно, по заданному преобразованию Лоренца уравнение (1.200) определяет матрицу M с точностью до знака, причем отображение (1.201) сюръективно [25]. Поскольку группа $\mathbb{SL}(2, \mathbb{C})$ связна, то это устанавливает изоморфизм групп

$$\mathbb{SO}_0(1, 3) \simeq \frac{\mathbb{SL}(2, \mathbb{C})}{\mathbb{Z}_2}. \quad (1.202)$$

При этом тождественному преобразованию Лоренца соответствуют две матрицы 1 и -1 из группы $\mathbb{SL}(2, \mathbb{C})$. Можно доказать (см., например, [32]), что группа $\mathbb{SL}(2, \mathbb{C})$ является односвязной. Это значит, что построенное отображение представляет собой двулистное универсальное накрытие (см. раздел 11). Таким образом справедлива

Теорема 1.9.2. *Существует гомоморфизм групп (1.201), который представляет собой двулистное универсальное накрытие.*

Определение. Точное представление универсальной накрывающей $\mathbb{SL}(2, \mathbb{C})$ называется *спинорным* представлением собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, 3)$. При этом каждому вращению $S \in \mathbb{SO}_0(1, 3)$ ставится в соответствие неупорядоченная пара элементов $\pm M \in \mathbb{SL}(2, \mathbb{C})$. $\square$

Избавиться от неоднозначности спинорного представления, т.е. упорядочить пару элементов $\pm M$, нельзя, т.к. при тождественном преобразовании (пространственное вращение на угол 2π) матрица M изменит знак. Поэтому говорят, что оно является двузначным представлением собственной ортохронной группы Лоренца. Таким образом, спинорное представление связной компоненты единицы группы Лоренца $\mathbb{SO}_0(1, 3)$ является комплексным двумерным. Спинорные представления полной группы Лоренца $\mathbb{O}(1, 3)$ являются комплексными четырехмерными.

Собственные вращения $\mathbb{SO}(3)$ в трехмерном евклидовом пространстве, определяемом уравнением $t = 0$ в пространстве Минковского, образуют подгруппу группы Лоренца $\mathbb{SO}_0(1, 3)$. Пусть $\tilde{S} \in \mathbb{SO}(3)$. Тогда каждой матрице вращений $\tilde{S}$ соответствуют две матрицы $\pm U$ из группы специальных унитарных матриц $\mathbb{SU}(2)$, которая является подгруппой в $\mathbb{SL}(2, \mathbb{C})$. Это следует из параметризации (1.199). Заметим, что для преобразований (1.200) с унитарной матрицей $U \in \mathbb{SU}(2) \subset \mathbb{SL}(2, \mathbb{C})$ сохраняется не только определитель, но и след, $\text{tr } A' = \text{tr } A$. Так как $\text{tr } A = 2t$, то каждому преобразованию из подгруппы унитарных унимодулярных матриц ставится в соответствие такое преобразование Лоренца, которое не затрагивает время, т.е. некоторое вращение $\tilde{S} \in \mathbb{SO}(3)$.

Аналогичное рассуждение применимо к подгруппе вещественных матриц $\mathbb{SL}(2, \mathbb{R}) \subset \mathbb{SL}(2, \mathbb{C})$. Если $M \in \mathbb{SL}(2, \mathbb{R})$, то $M^T = M^\dagger$, и преобразование (1.200) переходит в $A' = MAM^T$, которое отображает вещественные симметричные матрицы на себя. Это соответствует таким преобразованиям пространства Минковского, которые не затрагивают координату y в параметризации (1.199). Следовательно, каждой вещественной унимодулярной матрице M из $\mathbb{SL}(2, \mathbb{R})$ ставится в соответствие некоторое собственное ортохронное преобразование Лоренца из $\mathbb{SO}_0(1, 2)$, которое действует только на координаты t, x, z в пространстве Минковского.

Посмотрим, что представляют из себя группа Лоренца $\mathbb{SO}_0(1, 3)$ и ее универсальная накрывающая $\mathbb{SL}(2, \mathbb{C})$ как многообразия.

Предложение 1.9.7. *Группа специальных комплексных 2×2 матриц $\mathbb{SL}(2, \mathbb{C})$ и собственная ортохронная группа Лоренца $\mathbb{SO}_0(1, 3)$ как многообразия представля-*

ют собой прямые произведения:

$$\mathrm{SL}(2, \mathbb{C}) \approx \mathbb{S}^3 \times \mathbb{R}^3. \quad (1.203)$$

$$\mathrm{SO}_0(1, 3) \approx \mathbb{RP}^3 \times \mathbb{R}^3. \quad (1.204)$$

Доказательство. Согласно полярному разложению матриц (теорема 37.1.6) произвольная матрица $M \in \mathrm{SL}(2, \mathbb{C})$ однозначно представима в виде

$$M = RU, \quad (1.205)$$

где $R = \sqrt{MM^\dagger}$ – положительно определенная эрмитова матрица и U – унитарная матрица. Поскольку $\det M = 1$, то выполнены равенства $\det R = 1$ и $\det U = 1$. Следовательно, $U \in \mathrm{SU}(2)$. Множество эрмитовых 2×2 матриц с единичным определителем имеет вид (1.199) с единственным ограничением

$$\det R = t^2 - x^2 - y^2 - z^2 = 1.$$

Поскольку матрица R положительно определена, то нам необходимо выбрать верхнюю полу однополостного гиперboloида, соответствующую значениям $t > 0$. Она, как многообразие, диффеоморфна трехмерному евклидову пространству $\mathbb{R}^3$. Ранее было установлено, что многообразие группы $\mathrm{SU}(2)$ представляет собой сферу $\mathbb{S}^3$. Таким образом построен диффеоморфизм (1.203).

Собственная ортохронная группа Лоренца $\mathrm{SO}_0(1, 3)$ возникает после отождествления $M \sim -M$. В полярном разложении (1.205) это означает отождествление унитарных матриц $U \sim -U$, т.к. матрица R положительно определена. Это отождествление соответствует накрытию трехмерного проективного пространства $\mathbb{RP}^3$ трехмерной сферой $\mathbb{S}^3$ (см. раздел 1.8). Следовательно, существует диффеоморфизм (1.204). $\square$

Существование диффеоморфизмов (1.203) и (1.204) согласуется с общей теоремой 8.11.10. Тем самым мы показали, что подгруппы $\mathrm{SU}(2)$ и $\mathrm{SO}(3)$ являются максимальными компактными подгруппами в $\mathrm{SL}(2, \mathbb{C})$ и $\mathrm{SO}_0(1, 3)$, соответственно.

Определим оператор пространственного отражения (четности) в пространстве Минковского как отражение всех пространственных координат

$$P : (x^0, x^1, x^2, x^3) \mapsto (x^0, -x^1, -x^2, -x^3).$$

Тогда внешние автоморфизмы (1.159), вызванные обращением времени и преобразованием четности, совпадают, т.к. $P = -T$. При инверсии пространственных координат матрица преобразований Лоренца меняется по правилу (1.160):

$$P : S \mapsto PSP^{-1} = (S^T)^{-1}.$$

Пусть пространственному вращению $\tilde{S} \in \mathrm{SO}(3)$ соответствуют две комплексные матрицы $\pm \tilde{A} \in \mathrm{SU}(2)$. Тогда внешнему автоморфизму (1.160) соответствует автоморфизм в группе $\mathrm{SL}(2, \mathbb{C})$

$$P : \pm \tilde{A} \mapsto \pm (\tilde{A}^T)^{-1}.$$

Последний будет внутренним, т.к. группа $\mathrm{SL}(2, \mathbb{C})$ связна. Это значит, что представление группы $\mathrm{SL}(2, \mathbb{C})$ является спинорным представлением не только для связной компоненты единицы $\mathrm{SO}_0(1, 3)$, но и группы $\mathrm{O}_0(1, 3)$, включающей преобразование четности.

Группа Лоренца является некомпактной, ее представления могут быть неунитарны, а среди ее неприводимых представлений есть бесконечномерные. Известно, что все конечномерные неприводимые представления, за исключением тривиального единичного, группы Лоренца неунитарны [25]. Это существенно отличает ее от группы вращений в евклидовом пространстве и представляет существенные трудности при построении физических моделей, где требуется положительная определенность энергии и скалярного произведения в гильбертовом пространстве соответствующей квантовой теории. Из изоморфизма (1.202) следует, что точное представление группы $\mathrm{SL}(2, \mathbb{C})$ является двузначным или спинорным представлением группы $\mathrm{SO}_0(1, 3)$. При этом представление собственной ортохронной группы Лоренца $\mathrm{SO}_0(1, 3)$ однозначно или двузначно одновременно с порожденным им представлением группы вращений $\mathrm{SO}(3)$.

Построим параметризацию группы $\mathrm{SL}(2, \mathbb{C})$ элементами ее алгебры. Алгебра Ли $\mathfrak{sl}(2, \mathbb{C})$ состоит из комплексных 2×2 -матриц с нулевым следом (6 вещественных параметров). Алгебру Ли $\mathfrak{sl}(2, \mathbb{C})$ можно рассматривать как трехмерное векторное пространство над полем комплексных чисел. В качестве базиса выберем матрицы Паули σ_i (см. приложение 37.2). Тогда произвольный элемент алгебры Ли $\mathfrak{sl}(2, \mathbb{C})$ параметризуется вектором $\mathbf{z}$ в трехмерном комплексном пространстве

$$\mathbf{z} = z^k \sigma_k, \quad \mathbf{z} = \{z^k\} \in \mathbb{C}^3, \quad k = 1, 2, 3.$$

Предложение 1.9.8. *Каждому элементу алгебры Ли $\mathbf{z} = \{z^k\} \in \mathfrak{sl}(2, \mathbb{C}) \approx \mathbb{C}^3$ соответствует элемент группы Ли $S = \{S_A^B\} \in \mathrm{SL}(2, \mathbb{C})$*

$$\begin{aligned} S_A^B(\mathbf{z}) &= \left(e^{iz^k \sigma_k / 2} \right)_A^B = \delta_A^B \cos \frac{\sqrt{\mathbf{z}^2}}{2} + i \frac{z^k \sigma_{kA}^B}{\sqrt{\mathbf{z}^2}} \sin \frac{\sqrt{\mathbf{z}^2}}{2} = \\ &= \begin{pmatrix} \cos \frac{\sqrt{\mathbf{z}^2}}{2} + i \frac{z^3}{\sqrt{\mathbf{z}^2}} \sin \frac{\sqrt{\mathbf{z}^2}}{2} & i \frac{z^1 - iz^2}{\sqrt{\mathbf{z}^2}} \sin \frac{\sqrt{\mathbf{z}^2}}{2} \\ i \frac{z^1 + iz^2}{\sqrt{\mathbf{z}^2}} \sin \frac{\sqrt{\mathbf{z}^2}}{2} & \cos \frac{\sqrt{\mathbf{z}^2}}{2} - i \frac{z^3}{\sqrt{\mathbf{z}^2}} \sin \frac{\sqrt{\mathbf{z}^2}}{2} \end{pmatrix}, \end{aligned} \quad (1.206)$$

где $\mathbf{z}^2 := (z^1)^2 + (z^2)^2 + (z^3)^2$ – комплексное число.

Доказательство. Формальное переписывание (1.144), т.к. все разложения равномерно сходятся в произвольном ограниченном шаре в $\mathbb{C}^3$. $\square$

Многообразие группы $\mathrm{SL}(2, \mathbb{C})$ возникает после отождествления точек в $\mathbb{C}^3$, т.к. некоторым элементам алгебры соответствует один и тот же элемент группы. В рассматриваемом случае это отождествление выглядит громоздко, и мы не будем его проводить.

1.10 Специальная теория относительности

Специальная теория относительности – это учение о пространстве-времени, в котором строятся современные модели всех взаимодействий за исключением гравитационных. В двух словах суть специальной теории относительности состоит в утверждении, что пространство-время представляет собой пространство Минковского $\mathbb{R}^{1,3}$, в котором определена метрика Лоренца $\eta_{ab} = \mathrm{diag}(+ - - -)$, и все законы Природы записываются в виде некоторой системы уравнений, которые ковариантны относительно преобразований Лоренца.

Мы начнем с описания основных свойств нерелятивистских моделей, а затем обсудим некоторые свойства пространства Минковского, которые затем будут обобщены в теории тяготения Эйнштейна. Кроме того, рассмотренные примеры из специальной теории относительности помогают развить определенную интуицию, которая полезна при исследовании многообразий общего вида, на которых задана метрика лоренцевой сигнатуры.

1.10.1 Нерелятивистские модели

До создания специальной теории относительности модели математической физики были нерелятивистскими и строились в пространстве-времени, в котором время играет выделенную роль.

Определение. *Галилеевым* пространством-временем называется топологически тривиальное четырехмерное многообразие $\mathbb{M} = \mathbb{R} \times \mathbb{R}^3$, где первый сомножитель $\mathbb{R}$ соответствует *времени* t , а второй $\mathbb{R}^3$ – *пространству*, со следующей структурой. Мы предполагаем, что в пространстве $\mathbb{R}^3$ задана евклидова метрика. Пусть x^i , $i = 1, 2, 3$, – декартовы координаты в пространстве $\mathbb{R}^3$, в которых метрика имеет диагональный вид $\delta_{ij} = \text{diag}(++ +)$. Каждая точка пространства-времени $x = \{t, x^i\} \in \mathbb{M}$ называется *событием* и говорит о том, что нечто произошло в момент времени t в точке пространства $\{x^i\}$. События, соответствующие фиксированному значению t , называются *одновременными*. Пусть задано два произвольных события $x_1, x_2 \in \mathbb{M}$ с координатами $x_1 = \{t_1, x_1^i\}$ и $x_2 = \{t_2, x_2^i\}$. Тогда временной интервал между этими событиями определяется разностью

$$\Delta t := t_2 - t_1. \quad (1.207)$$

Если $t_2 > t_1$, то мы говорим, что событие x_2 произошло позже события x_1 . В противном случае, $t_2 < t_1$, событие x_2 произошло раньше события x_1 . Если $t_2 = t_1$, то мы говорим, что события произошли одновременно. Пространственное расстояние между двумя одновременными событиями $x_1 = \{t, x_1^i\}$ и $x_2 = \{t, x_2^i\}$ равно евклидову расстоянию в $\mathbb{R}^3$:

$$\Delta x := \sqrt{\delta_{ij}(x_2^i - x_1^i)(x_2^j - x_1^j)}. \quad (1.208)$$

Система координат t, x^i называется *инерциальной* и обозначается O . Кроме того, мы предполагаем, что галилеево пространство-время снабжено естественной линейной и аффинной структурой, такой же как и евклидово пространство $\mathbb{R}^4$ в разделах 1.3.3 и 1.3.4. $\square$

Одновременные события составляют пространственные сечения пространства-времени $\mathbb{M}$, соответствующие фиксированному значению времени t . Из определения следует, что понятия “раньше” и “позже” не зависят от выбора событий в каждом сечении, т.е. временной интервал определен для двух произвольных сечений t_1 и t_2 . Другими словами, на оси времени $t \in \mathbb{R}$ задана обычная топологическая евклидова метрика $l(t_1, t_2) = |t_2 - t_1|$, которая определяет расстояние между двумя пространственными сечениями. Единственное отличие – это наличие знака у временного интервала, который соответствует понятиям “раньше” и “позже”.

Подчеркнем, что пространственное расстояние можно определить только для одновременных событий. Если две системы координат движутся друг относительно друга с ненулевой скоростью и события не являются одновременными, то, как легко

видеть, пространственное расстояние (1.208) между этими событиями, измеренное в этих системах координат, будет различным.

Данное выше определение галилеева пространства-времени содержит одну инерциальную систему отсчета O . Расширим это понятие. Временной интервал между двумя произвольными событиями и пространственное расстояние между двумя одновременными событиями инвариантны относительно преобразований Галилея (1.174). Обратное утверждение в общем случае неверно.

Обсудим это. Сохранение временного интервала оставляет возможность только для сдвигов вдоль оси времени. Эти сдвиги не могут зависеть от точки пространства, т.к. мы требуем сохранения интервала (1.207) не для отдельных событий, а для сечений.

Пространственная метрика (1.208) инвариантна только относительно преобразований из неоднородной группы вращений $\mathbb{IO}(3)$. В общем случае преобразований пространства-времени $\mathbb{R} \times \mathbb{R}^3$ и матрица вращений, и вектор сдвига пространства могут зависеть от времени, причем произвольно. Однако, если ограничиться линейными преобразованиями, то остаются только преобразования из группы Галилея. Снабдим галилеево пространство-время обычной структурой линейного пространства из $\mathbb{R}^4$. Тогда два произвольных различных события определяют единственную прямую (1.25), проходящую через эти события. При преобразованиях Галилея линейная структура сохраняется и прямые переходят в прямые. Верно также обратное утверждение: любое преобразование пространства-времени, которое сохраняет временной интервал (1.207), пространственную метрику (1.208) и линейную структуру пространства-времени, называется преобразованием Галилея.

Определение. Любая система координат в галилеевом пространстве-времени $\mathbb{M} = \mathbb{R} \times \mathbb{R}^3$, связанная с системой координат O преобразованием из группы Галилея, называется *инерциальной*. $\square$

Точечные частицы движутся в пространстве по траекториям $\{x^i(t)\} \in \mathbb{R}^3$. Эти три функции задают мировую линию частицы в пространстве-времени $\{t, x^i(t)\} \in \mathbb{M}$. Мировые линии частиц вводятся точно таким же образом и в специальной, и в общей теории относительности. Различие заключается лишь в том, что время в нерелятивистской механике имеет абсолютный характер, и его естественно выбрать в качестве параметра вдоль мировых линий всех частиц. Такой выбор общепринят и нагляден с физической точки зрения.

Определение. Если в некоторой инерциальной системе координат мировая линия частицы представляет собой прямую линию вида $\{t, x^i = v^i t + x_0^i\} \in \mathbb{M}$, где v^i, x_0^i – некоторые постоянные, то такая частица называется *свободной*. Точка $\{x_0^i\} \in \mathbb{R}^3$ является точкой пространства, в которой частица расположена в начальный момент времени $t = 0$, а векторное поле $\mathbf{v} = \{v^i\} \in \mathbb{T}(\mathbb{R}^3)$, определенное вдоль траектории частицы, называется *наблюдаемой скоростью*. $\square$

Из определения инерциальной системы координат следует, что, если мировая линия частицы является прямой в какой-либо инерциальной системе отсчета, то она будет прямой и в любой другой инерциальной системе.

Замечание. В физической литературе понятие инерциальной системы координат часто определяют следующим образом. Систему координат называют инерциальной, если свободная частица движется в ней равномерно и прямолинейно. В таком определении не сказано, что значит свободная частица. Поэтому мы пошли другим путем.

Сначала было построено пространство-время и инерциальные системы отсчета. Это позволило дать определение свободной частицы. $\square$

Замечание. В пространстве-времени $\mathbb{M}$ не существует метрики, инвариантной относительно преобразований Галилея, потому что эта группа не является полупростой. Поэтому говорить про экстремали не имеет смысла. В галилеевом пространстве-времени заданы естественные линейная и аффинная структуры, и параллельный перенос мы отождествляем с трансляциями. Это определяет связность на $\mathbb{M}$, которая в инерциальной системе отсчета имеет нулевые компоненты. Эта связность плоская, а геодезические являются прямыми линиями. $\square$

Линейность функций $x^i(t)$ означает, что свободная частица движется в пространстве равномерно и прямолинейно. Это утверждение известно как *первый закон Ньютона*. Мировая линия свободной частицы в пространстве-времени в любой инерциальной системе координат представляет собой прямую линию. Мы говорим, что частица не является свободной, если ее мировая линия отличается от прямой. В этом случае траектория частицы имеет ненулевое ускорение $\mathbf{a} := d\mathbf{v}/dt$. Мы говорим, что на частицу действует сила $\mathbf{F}$, которая, по определению, пропорциональна ускорению:

$$\mathbf{F} = m\mathbf{a}, \quad (1.209)$$

где коэффициент пропорциональности m называется *инертной массой*. Это соотношение называется *вторым законом Ньютона*.

Если имеется две взаимодействующих частицы, то между ними действует сила. Пусть $\mathbf{F}$ – сила, действующая на первую частицу со стороны второй. Тогда сила, действующая на вторую частицу со стороны первой, равна $-\mathbf{F}$. Это утверждение составляет *третий закон Ньютона*. Коротко говорят: “Сила действия равна силе противодействия”.

Для механики Галилея справедливо следующее правило сложения скоростей. Если частица движется относительно инерциальной системы координат с постоянной скоростью $\mathbf{v} = \{v^i\}$, то ее скорость движения относительно другой инерциальной системы координат, движущейся со скоростью $\mathbf{V} = \{V^i\}$ относительно первой, равна разности $\mathbf{v} - \mathbf{V} = \{v^i - V^i\}$.

Модели математической физики, которые инвариантны или ковариантны относительно преобразований Галилея называются *нерелятивистскими*.

Пример 1.10.1. В теории тяготения Ньютона две точечные частицы с массами m_1 и m_2 , находящиеся, соответственно, в точках $\mathbf{x}_1$ и $\mathbf{x}_2$ (в инерциальной системе отсчета), испытывают гравитационное притяжение. При этом сила, действующая со стороны второй частицы на первую, равна

$$\mathbf{F}_{12} = G \frac{m_1 m_2}{r^2} \frac{\mathbf{r}}{|\mathbf{r}|}, \quad (1.210)$$

где G – гравитационная постоянная, $\mathbf{r} := \mathbf{x}_2 - \mathbf{x}_1$ и $|\mathbf{r}|$ – расстояние между частицами. Выражение (1.210) называется законом *всемирного тяготения* или *законом тяготения Ньютона*. Теория тяготения Ньютона является нерелятивистской моделью, в которой время имеет абсолютный характер. Она инвариантна относительно преобразований Галилея, поскольку сила гравитационного взаимодействия двух тел в каждый момент времени зависит только от их масс и расстояния между ними. Массы частиц m_1 и m_2 называются *гравитационными массами*. В механике Ньютона инертная и гравитационная массы частиц считаются равными. В этом случае из

закона всемирного тяготения и второго закона Ньютона (1.209) следует, что ускорение, которая испытывает частица массы m_1 в поле тяжести частицы m_2 , не зависит от ее массы. Это следствие проверено в поле тяжести Земли и Солнца с высокой степенью точности. $\square$

В классической механике взаимодействие частиц описывается энергией потенциального взаимодействия, которая является функцией от координат частиц. Тем самым изменение положения одной из взаимодействующих частиц отразится на других частицах в тот же момент времени. Это соответствует предположению о том, что взаимодействие распространяется с бесконечной скоростью. Опыт, однако, показывает, что мгновенных взаимодействий в природе не существует. Если сдвинуть одну частицу, то это отразится на других частицах только через некоторое время, которое определяется скоростью распространения взаимодействий.

Электромагнитные взаимодействия распространяются со скоростью света

$$c = 2,998 \cdot 10^{10} \text{ см/с} \approx 300000 \text{ км/с}.$$

В 1887 году Майкельсон и Морли [33] экспериментально установили с точностью до 5 км/с, что скорость света постоянна и не зависит от того, движется ли он параллельно траектории Земли или в перпендикулярном направлении. Точность этого результата сейчас доведена до 1 км/с [34].

Постоянство скорости света находится в явном противоречии с правилом сложения скоростей в механике Галилея и с предположением о мгновенности распространения взаимодействий. Это и привело к созданию специальной теории относительности.

1.10.2 Релятивистские модели

Специальная теория относительности изменила наше представление о структуре пространства-времени. В релятивистских моделях отсутствует понятие одновременных событий и пространственного расстояния между двумя одновременными событиями. Эти обстоятельства являются наиболее трудными в процессе понимания релятивистских моделей, т.к. наша интуиция основана на механике Ньютона.

Определение. *Пространством Минковского $\mathbb{R}^{1,3}$ называется четырехмерное топологически тривиальное многообразие с заданной лоренцевой метрикой в инерциальной системе отсчета и естественными линейной и аффинной структурами евклидова пространства.* $\square$

В специальной теории относительности мы предполагаем, что все законы природы формулируются в пространстве Минковского. Как линейное пространство пространство-время $\mathbb{R}^{1,3}$ такое же, как и в нерелятивистских моделях. С топологической точки зрения это просто прямое произведение, $\mathbb{R}^{1,3} = \mathbb{R} \times \mathbb{R}^3$. Точка пространства-времени, как и раньше, называется событием, и каждая точечная частица движется вдоль своей мировой линии в $\mathbb{R}^{1,3}$. Отличие заключается в том, что вместо понятий одновременности и пространственного расстояния между одновременными событиями, мы постулируем существование инвариантной метрики в пространстве Минковского $\mathbb{R}^{1,3}$. В декартовой системе координат x^a , $a = 0, 1, 2, 3$, лоренцева метрика, по определению, имеет диагональный вид $\eta_{ab} := \text{diag}(+, -, -, -)$.

Как и в нерелятивистских моделях, в специальной теории относительности вводится понятие инерциальной системы координат. Однако ее определение другое.

Определение. Система координат в пространстве Минковского $\mathbb{R}^{1,3}$, в которой метрика имеет вид $\eta_{ab} = \text{diag}(+ - - -)$, называется *инерциальной*. Если мировая линия частицы в инерциальной системе координат представляет собой прямую линию, то она называется *свободной*. $\square$

Замечание. Напомним, что прямые линии в пространстве Минковского и только они являются экстремальными. Поэтому в определении свободной частицы “прямую линию” можно заменить на “экстремаль”. $\square$

Замечание. Данное определение инерциальной системы координат по своей сути совпадает с определением декартовой системы координат, которое было дано в разделе 1.3.1 в случае евклидовой метрики. $\square$

Инерциальные системы координат определены неоднозначно. Любые две инерциальные системы координат связаны между собой преобразованием из полной группы Пуанкаре, которая включает в себя преобразования Лоренца (собственные и несобственные), сдвиги, преобразование четности и обращение времени (см. раздел 1.9). Обратно, если некоторая система координат связана с инерциальной преобразованием из группы Пуанкаре, то она сама является инерциальной. Это следует из определений группы Пуанкаре и инерциальной системы координат.

Вторым постулатом специальной теории относительности является предположение о том, что все физические явления описываются некоторым набором полей на пространстве Минковского $\mathbb{R}^{1,3}$, которые преобразуются по какому то, возможно, приводимому представлению группы Пуанкаре.

В физической литературе принято формулировать

Принцип относительности. Все физические законы природы (уравнения движения, равновесия и т.д.) инвариантны или ковариантны относительно преобразований из группы Пуанкаре. $\square$

Пример 1.10.2. Уравнение Клейна–Гордона–Фока для скалярного поля (25.3) в пространстве Минковского инвариантно относительно преобразований из группы Пуанкаре. Уравнение Дирака (27.72) для спинорного поля и уравнения Максвелла (25.145) для электромагнитного поля в пространстве Минковского ковариантны относительно преобразований из группы Пуанкаре. $\square$

Замечание. Часто принцип относительности формулируют следующим образом: “Законы природы выглядят одинаково во всех инерциальных системах отсчета”. В такой формулировке за кадром остается определение инерциальной системы отсчета и смысл термина “одинаково”. $\square$

Приведенные рассуждения приводят к следующему итогу. В специальной теории относительности имеется три основных постулата:

1. Пространство-время, в котором происходят все окружающие нас явления, представляет собой четырехмерное пространство Минковского $\mathbb{R}^{1,3}$.
2. Модели математической физики строятся из соответствующего набора полей на пространстве Минковского $\mathbb{R}^{1,3}$, которые преобразуются по некоторому представлению группы Пуанкаре.
3. Все законы природы описываются некоторыми уравнениями, которые либо инвариантны, либо ковариантны относительно преобразований из группы Пуанкаре.

Мы всегда предполагаем, что точечные частицы, на которые не действуют никакие силы, кроме гравитационных, движутся в пространстве-времени вдоль экстремалей. В специальной теории относительности гравитационные эффекты не учитываются. Поэтому свободные частицы, на которые не действуют никакие силы, движутся в пространстве Минковского вдоль экстремалей. В пространстве Минковского экстремали совпадают с геодезическими и являются прямыми линиями. Поэтому свободные точечные частицы в специальной теории относительности в инерциальной системе координат, как и в нерелятивистских моделях, движутся равномерно и прямолинейно.

Мы говорим, что пространство-время специальной теории относительности однородно и изотропно. Математически это значит, что метрика инвариантна относительно сдвигов из группы Пуанкаре и вращений пространства Минковского, т.е. преобразований Лоренца. Другими словами, физические свойства пространства-времени не зависят от выбора начала отсчета декартовой системы координат и направления.

Понятие инерциальной системы координат является математической абстракцией. Поскольку во вселенной существуют материальные тела, то гравитационное поле не устранимо. Поэтому системы координат, которые используются в экспериментах, можно считать инерциальными только с определенной степенью точности.

Пример 1.10.3. Система координат, покоящаяся относительно Земли, с хорошей степенью точности является инерциальной при изучении электромагнитных явлений. $\square$

Несмотря на то, что инерциальных систем отсчета, строго говоря, в природе не существует, это понятие чрезвычайно важно в современной теоретической физике. Модели математической физики (например, электродинамика), сформулированные в пространстве Минковского, приводят к следствиям, которые находятся в прекрасном согласии с экспериментами. В настоящее время известно только об одном случае несогласия специальной теории относительности с опытом: нарушение четности в слабых взаимодействиях. Поэтому считается, что специальная теория относительности находится в прекрасном согласии с экспериментом, и все взаимодействия, кроме гравитационных, инвариантны относительно трансляций и преобразований Лоренца.

Наиболее успешные модели математической физики: модели электромагнитных, слабых, сильных (квантовая хромодинамика) взаимодействий, а также их объединения, не претендующие на описание гравитационных взаимодействий, построены в пространстве Минковского. С этой целью выбирается некоторое представление группы группы Пуанкаре и строится инвариантное действие в пространстве Минковского, которое приводит к ковариантным уравнениям движения. В последние десятилетия такой подход оказался самым распространенным и успешным.

Обсудим понятие одновременности и пространственного расстояния в пространстве Минковского. Зафиксируем два события $x_1 = \{ct_1, x_1^i\}$ и $x_2 = \{ct_2, x_2^i\}$ в $\mathbb{R}^{1,3}$, где мы ввели скорость света c , чтобы измерять время в секундах, а не в метрах. Если сказать, что эти события разделены временным интервалом $\Delta t := t_2 - t_1$, то такое определение будет некорректным, т.к. Δt зависит от выбора инерциальной системы отсчета. Это следует прямо из выражения для преобразований координат (1.118) (лоренцевых бустов в плоскости x^0, x^1). Поэтому понятие одновременности в специальной теории относительности отсутствует. Как следствие, нельзя также корректно ввести понятие пространственного расстояния между одновременными событиями, т.к. одновременные события не определены. Вместо этого вводится понятие интер-

вала между двумя событиями

$$\Delta s^2 = (x_2^0 - x_1^0)^2 - (x_2^1 - x_1^1)^2 - (x_2^2 - x_1^2)^2 - (x_2^3 - x_1^3)^2. \quad (1.211)$$

Кроме этого, введение скорости света полезно для определения нерелятивистского предела. Данное определение интервала корректно, т.к. не зависит от выбора инерциальной системы отсчета. Интервал для двух событий может быть положительным, равен нулю или отрицательным. Мы говорим, что два события причинно связаны, если интервал между ними положителен, $\Delta s^2 > 0$. В этом случае существует такая инерциальная система координат, в которой оба события происходят в одной точке. Если свет испущен в точке x_1 и получен в точке x_2 , то для этих событий интервал равен нулю, $\Delta s^2 = 0$. Если интервал отрицателен, $\Delta s^2 < 0$, то события являются причинно не связанными. Для этих событий существует такая инерциальная система координат, в которой оба события происходят одновременно.

Рассмотрим две инерциальные системы координат O и O' с декартовыми координатами t, x, y, z и t', x', y', z' , соответственно. Пусть начала систем координат совпадают и система координат O' движется со скоростью V относительно O вдоль оси x . Тогда события в этих системах координат связаны между собой преобразованием Лоренца (1.118):

$$t' = \frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}}, \quad x' = \frac{x - Vt}{\sqrt{1 - V^2/c^2}}, \quad y' = y, \quad z' = z. \quad (1.212)$$

Обратные преобразования координат от системы координат O' к O получаются простой заменой $V \rightarrow -V$:

$$t = \frac{t' + Vx'/c^2}{\sqrt{1 - V^2/c^2}}, \quad x = \frac{x' + Vt'}{\sqrt{1 - V^2/c^2}}, \quad y = y', \quad z = z'. \quad (1.213)$$

Напомним, что эти преобразования координат называются бустами.

Аналогичные формулы преобразования координат справедливы, если система координат O' движется относительно O либо вдоль оси y , либо оси z .

В приложениях иногда необходимо знать выражение для преобразований Лоренца в общем случае, когда система координат O' движется относительно O равномерно и прямолинейно, но в произвольном направлении. Эти формулы легко получить из (1.212) в виде, который инвариантен относительно пространственных вращений. Пусть $\mathbf{V} = \{V^i\} \neq 0$ – скорость движения штрихованной системы координат. Введем векторное обозначение $\mathbf{x} = \{x^i\} = \{x, y, z\}$ для пространственных координат событий и аналогичные обозначения в штрихованной системе координат. Введем также обычное евклидово скалярное произведение пространственных векторов:

$$\mathbf{V}^2 := -V^i V_i, \quad (\mathbf{x}, \mathbf{V}) := -x^i V_i.$$

Знак минус в этих выражениях следует из того, что при опускании пространственных индексов знак компонент меняется: $V_i = V^j \eta_{ji} = -V^i$. Тогда радиус-вектор можно разложить на составляющие, $\mathbf{x} = \mathbf{x}_{\parallel} + \mathbf{x}_{\perp}$, которые параллельны и перпендикулярны вектору скорости:

$$\mathbf{x}_{\parallel} := \frac{(\mathbf{x}, \mathbf{V})\mathbf{V}}{\mathbf{V}^2}, \quad \mathbf{x}_{\perp} := \mathbf{x} - \frac{(\mathbf{x}, \mathbf{V})\mathbf{V}}{\mathbf{V}^2}.$$

Аналогично

$$\mathbf{x}'_{\parallel} := \frac{(\mathbf{x}', \mathbf{V})\mathbf{V}}{\mathbf{V}^2}, \quad \mathbf{x}'_{\perp} := \mathbf{x}' - \frac{(\mathbf{x}', \mathbf{V})\mathbf{V}}{\mathbf{V}^2}.$$

Из преобразований Лоренца (1.212) следует, что перпендикулярная составляющая радиус-вектора не меняется, а параллельная меняется так же, как координата x . Отсюда вытекает общее правило преобразования координат:

$$t' = \frac{t - (\mathbf{x}, \mathbf{V})/c^2}{\sqrt{1 - \mathbf{V}^2/c^2}}, \quad \mathbf{x}'_{\parallel} = \frac{\mathbf{x}_{\parallel} - \mathbf{V}t}{\sqrt{1 - \mathbf{V}^2/c^2}}, \quad \mathbf{x}'_{\perp} = \mathbf{x}_{\perp}. \quad (1.214)$$

Для полного радиуса-вектора получаем следующую формулу преобразования

$$\mathbf{x}' = \mathbf{x} + \frac{(\mathbf{x}, \mathbf{V})\mathbf{V}}{\mathbf{V}^2} \left(\frac{1}{\sqrt{1 - \mathbf{V}^2/c^2}} - 1 \right) - \frac{\mathbf{V}t}{\sqrt{1 - \mathbf{V}^2/c^2}}$$

или, в компонентах,

$$x'^i = x^i - \frac{x^j V_j V^i}{\mathbf{V}^2} \left(\frac{1}{\sqrt{1 - \mathbf{V}^2/c^2}} - 1 \right) - \frac{V^i t}{\sqrt{1 - \mathbf{V}^2/c^2}}. \quad (1.215)$$

В общем случае координаты события преобразуются по правилу

$$t' = t S_0^0 + x^i S_i^0, \quad x'^i = t S_0^i + x^j S_j^i,$$

где $\{S_a^b\} = \{S_0^0, S_i^0, S_0^i, S_i^j\}$ – матрица преобразований Лоренца. Сравнение этого выражения с формулами (1.214) и (1.215) дает довольно симметричные выражения:

$$\begin{aligned} S_0^0 &= \frac{1}{\sqrt{1 - \mathbf{V}^2/c^2}}, \\ S_i^0 &= \frac{V_i}{c^2 \sqrt{1 - \mathbf{V}^2/c^2}}, \\ S_0^i &= -\frac{V^i}{\sqrt{1 - \mathbf{V}^2/c^2}}, \\ S_i^j &= \delta_i^j - \frac{V_i V^j}{\mathbf{V}^2} \left(\frac{1}{\sqrt{1 - \mathbf{V}^2/c^2}} - 1 \right). \end{aligned} \quad (1.216)$$

Нетрудно проверить инвариантность лоренцевой метрики η_{ab} относительно этих преобразований:

$$\eta_{ab} = S_a^c S_b^d \eta_{cd}.$$

Эти формулы для лоренцевых бустов получены в [35].

Полученные общие формулы для лоренцевых бустов (1.216) будут использованы в разделе 24.2 при АДМ параметризации репера.

В релятивистской механике скорость света c является универсальной постоянной, и не зависит от выбора инерциальной системы отсчета. Если отношение $\mathbf{V}^2/c^2$ стремится к нулю, что соответствует стремлению скорости света к бесконечности, то лоренцевы бусты (1.212) стремятся к преобразованиям Галилея.

1.10.3 Замедление времени и лоренцево сокращение

Сначала выведем формулу для замедления времени в движущейся системе отсчета. Пусть часы покоятся в системе координат O' и, следовательно, движутся равномерно и прямолинейно в системе O . Время t' , которое показывают часы в той системе

координат, где они покоятся, называется собственным временем. Введем для него специальное обозначение $t_0 = t'$. Это время совпадает с длиной мировой линии часов, деленной на c , которая является прямой, параллельной оси времени. Поскольку для часов $x' = \text{const}$, то из первой формулы преобразований Лоренца (1.213) следует, что для произвольных промежутков времени справедливо равенство

$$\Delta t_0 = \Delta t \sqrt{1 - V^2/c^2} \quad \Leftrightarrow \quad \Delta t = \frac{\Delta t_0}{\sqrt{1 - V^2/c^2}}.$$

Это соотношение показывает, что в движущейся системе координат время течет медленнее. Это свойство называется эффектом *замедления времени*.

Эффект замедления времени в движущейся системе координат породил множество “парадоксов”.

Парадокс близнецов. Допустим, что в инерциальной системе координат в некоторый момент времени родились два близнеца, событие A на рис.1.11. Первый близнец остался на месте, а второго посадили в ракету и отправили в космическое путешествие с большой скоростью. Через некоторое время ракета вернулась, и близнецы встретились снова, событие B . На рисунке мировые линии близнецов помечены цифрами 1 и 2, соответственно. С точки зрения первого близнеца ракета быстро двигалась и из-за замедления времени второй близнец должен оказаться более молодым, чем он сам. С точки зрения второго близнеца ситуация прямо противоположная: он покоился относительно ракеты, а двигался первый близнец. Поэтому именно первый близнец должен оказаться моложе.

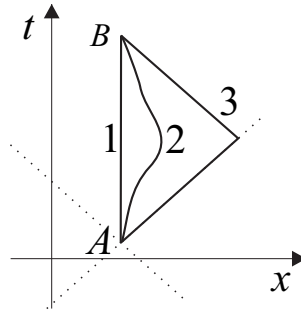


Рис. 1.11: Мировые линии близнецов 1 и 2. Мировая линия первого близнеца “длиннее” мировой линии второго.

Чтобы разъяснить парадокс, часто говорят, что рассуждения второго близнеца неверны, т.к. его система координат не может быть инерциальной (ракета должна замедлить свой полет и повернуть обратно). Однако это замечание ничего не объясняет. Действительно, если оба близнеца отправятся в космические путешествие в разные стороны, а потом вернуться и встретятся. Кто из них будет моложе? Они оба находились в неинерциальных системах отсчета, и с этой точки зрения совершенно равноправны. Для того, чтобы разрешить парадокс, необходимо дать

Определение. Возрастом человека называется промежуток собственного времени, который прошел с момента рождения. Или, эквивалентно, длина мировой линии человека от момента рождения, деленная на скорость света c . $\square$

Это определение корректно, т.к. инвариантно относительно произвольных преобразований координат. Поэтому возраст можно вычислить в любой, в том числе неинерциальной системе отсчета.

Теперь все становится на свои места. Поскольку на плоскости x, t задана метрика Лоренца, то длина любого наклонного отрезка между двумя моментами времени, меньше длины вертикального отрезка. Из рисунка ясно, что длина (собственное время) вертикального отрезка между точками A и B , который является мировой линией первого близнеца, максимальна. Поэтому длина любой (времениподобной) мировой линии, соединяющей точки A и B , меньше длины вертикального отрезка. Заметим, что если бы на плоскости x, t была задана евклидова метрика, то длина вертикального отрезка была бы наименьшей среди всех кривых, соединяющих точки A и B .

Разница в возрасте может оказаться большой, если ракета движется достаточно быстро. В предельном случае, когда ракета удаляется от первого близнеца со скоростью света, затем отражается и возвращается обратно также со скоростью света, второй близнец за время своего путешествия вообще не состарится (траектория 3 на рисунке). $\square$

Теперь предположим, что в движущейся системе координат O' на оси x' покоится линейка длины $l_0 := x'_2 - x'_1$. Собственная длина линейки l_0 измеряется при постоянном времени $t' = \text{const}$ в сопутствующей системе координат O' . В системе координат O линейка равномерно движется и ее длина измеряется при постоянном времени t . Из второй формулы в (1.212) вытекает соотношение между длинами:

$$l_0 = \frac{l}{\sqrt{1 - V^2/c^2}} \quad \Leftrightarrow \quad l = l_0 \sqrt{1 - V^2/c^2}.$$

Отсюда следует, что движущаяся линейка в покоящейся системе координат имеет длину $l < l_0$ и выглядит короче. Это явление называется *лоренцевым сокращением*.

Из формул для лоренцевых преобразований (1.212) следует, что поперечные наблюдаемые размеры тел вдоль осей y и z не испытывают никаких сокращений. Отсюда следует, что объем движущегося тела W уменьшается так же, как и его длина:

$$W := \Delta x \Delta y \Delta z = W_0 \sqrt{1 - V^2/c^2},$$

где W_0 – собственный объем в системе координат, связанной с телом.

Сокращение длин при движении, так же, как и замедление времени, породило много “парадоксов”.

Парадокс машины и гаража. Предположим, что машина и гараж имеют одинаковую собственную длину L . Водитель на большой скорости V въезжает в гараж, а дежурный по гаражу закрывает ворота в тот момент, когда задний бампер машины войдет в гараж. С точки зрения дежурного никаких проблем не возникает: машина испытывает лоренцево сокращение и, следовательно, без труда поместится в гараже. Однако с точки зрения водителя ситуация противоположная: лоренцево сокращение испытывает гараж и поэтому машина никак в гараже поместиться не может.

Разрешение “парадокса” заключается в отсутствии понятия одновременности в специальной теории относительности. Для простоты мы предполагаем, что машина свободно пробивает заднюю стену гаража и продолжает движение. На рисунках 1.12а, б показаны пространственно-временные диаграммы событий с точки зрения дежурного по гаражу и водителя. Для удобства, мы выбрали общее начало координат (точка O) таким образом, что оно соответствует моменту закрытия ворот. С точки зрения дежурного гараж покоится и ему соответствует вертикальная полоса, границы которой пересекают ось x в точках O и B . Машине соответствует затемненная наклонная полоса, пересекающая ось x в точках O и B . В системе координат дежурного машина испытывает лоренцево сокращение и поэтому событие A расположено

левее B . Это значит, что в момент закрытия ворот машина будет расположена в гараже. Через некоторое время передним бампер машины достигнет стены гаража и пробыет его, событие C .

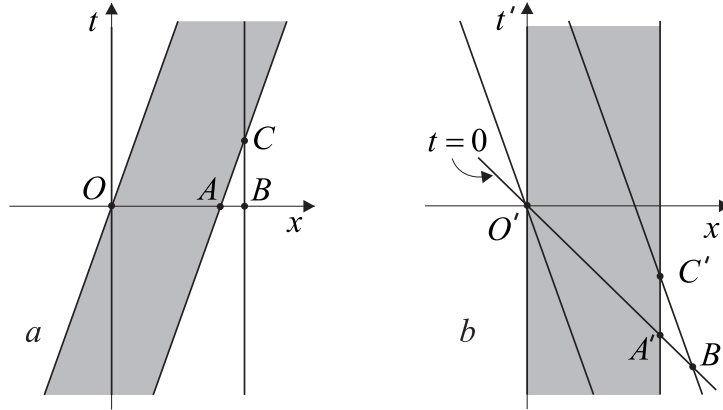


Рис. 1.12: Пространственно-временные диаграммы машины и гаража в различных системах координат. Общее начало систем координат, точка O , соответствует событию закрытия ворот. Затемненная полоса соответствует движению машины. Машина и гараж с точки зрения дежурного (a). Машина и гараж с точки зрения водителя (b).

Нетрудно проверить, что события A , B и C в нештрихованной системе координат, связанной с гаражом, на рис.1.12а имеют следующие координаты:

$$\begin{aligned} A &= \left(0, L\sqrt{1 - V^2/c^2}\right), \\ B &= (0, L) \\ C &= \left(\frac{L}{V} \left(1 - \sqrt{1 - V^2/c^2}\right), L\right). \end{aligned}$$

С точки зрения водителя гараж равномерно приближается к машине со скоростью V . Используя формулы для лоренцевых преобразований координат (1.212), вычислим координаты событий A , B и C в штрихованной системе координат, привязанной к водителю:

$$\begin{aligned} A' &= \left(-\frac{VL}{c^2}, L\right), \\ B' &= \left(-\frac{VL}{c^2\sqrt{1 - V^2/c^2}}, \frac{L}{\sqrt{1 - V^2/c^2}}\right) \\ C' &= \left(-\frac{L}{V} \left(1 - \sqrt{1 - V^2/c^2}\right), L\right). \end{aligned}$$

На рис.1.12b показана схема событий с точки зрения водителя. Машине и гаражу соответствуют вертикальная и наклонная полосы. С точки зрения водителя передний бампер машины сначала пробивает стену гаража, событие C' , и лишь после этого закрываются ворота, событие O' .

Таким образом, последовательность событий с точки зрения дежурного и водителя прямо противоположная. С точки зрения дежурного сначала закрываются ворота

и лишь потом машина пробивает стену. Водитель же видит, что сначала машина пробивает стену, и только потом закрываются ворота.

Заметим, что, если скорость света равна бесконечности, то все три события A , B и C имеют одинаковые координаты $(0, L)$ в обеих системах отсчета. $\square$

Рассмотренный пример наглядно показывает, что понятие одновременности событий в специальной теории относительности отсутствует. К этому трудно привыкнуть, т.к. наша интуиция основана на повседневной жизни, где скорости движения частиц малы по сравнению со скоростью света.

1.10.4 Сложение скоростей и эффект Доплера

Пусть, как и ранее, штрихованная система координат O' движется с постоянной скоростью V вдоль оси x в покоящейся системе координат O . Следовательно, координаты событий в этих системах отсчета связаны преобразованиями Лоренца (1.212). Предположим, что в штрихованной системе координат точечная частица движется равномерно и прямолинейно с постоянной скоростью $v' = (v'_x, v'_y, v'_z)$:

$$x' = v'_x t', \quad y' = v'_y t', \quad z' = v'_z t'.$$

Тогда в покоящейся системе координат мировая линия частицы будет прямой линией, заданной уравнениями:

$$\begin{aligned} t &= \frac{1 + v'_x V/c^2}{\sqrt{1 - V^2/c^2}} t', & y &= v'_y t', \\ x &= \frac{v'_x + V}{\sqrt{1 - V^2/c^2}} t', & z &= v'_z t', \end{aligned}$$

где время t' рассматривается, как параметр вдоль мировой линии частицы. Отсюда вытекает, что скорость частицы, наблюдаемая в покоящейся системе координат, имеет следующие компоненты:

$$\begin{aligned} v_x &:= \frac{x}{t} = \frac{v'_x + V}{1 + v'_x V/c^2}, \\ v_y &:= \frac{y}{t} = \frac{v'_y}{1 + v'_x V/c^2} \sqrt{1 - V^2/c^2}, \\ v_z &:= \frac{z}{t} = \frac{v'_z}{1 + v'_x V/c^2} \sqrt{1 - V^2/c^2}. \end{aligned} \tag{1.217}$$

Это и есть правило сложения скоростей в специальной теории относительности. Из полученных формул сразу следуют правила сложения скоростей в частных случаях, когда скорости параллельны или перпендикулярны друг другу.

Обозначим углы, которые составляют скорости v' и v с осями x' и x , соответственно, через α' и α , т.е.

$$v'_x = v' \cos \alpha', \quad v_x = v \cos \alpha.$$

Здесь мы предполагаем, что углы меняются в интервале: $0 \leq \alpha', \alpha \leq \pi$. Для модуля скорости $v := \sqrt{v_x^2 + v_y^2 + v_z^2}$, наблюдаемой в покоящейся системе отсчета, из (1.217) следует равенство

$$v = \frac{\sqrt{v'^2 + V^2 + 2v'V \cos \alpha' - (v'V \sin \alpha'/c)^2}}{1 + v'V \cos \alpha'/c^2}. \tag{1.218}$$

Далее, простые вычисления дают

$$\sqrt{1 - v^2/c^2} = \frac{\sqrt{1 - v'^2/c^2} \sqrt{1 - V^2/c^2}}{1 + v'V \cos \alpha'/c^2}. \quad (1.219)$$

Предложение 1.10.1. *Если скорость частицы в какой-либо инерциальной системе отсчета меньше скорости света, $v' < c$, то она будет меньше скорости света и в любой другой инерциальной системе координат, $v < c$. Если в движущейся системе координат распространяется свет, $v' = c$, то в покоящейся системе координат его скорость такая же, $v = c$.*

Доказательство. Если $v' < c$ и $V < c$, то знаменатель в (1.219) положителен, т.к. $\cos \alpha'$ ограничен сверху единицей. Следовательно, правая часть (1.219) положительна. Оценим ее сверху. Максимальное значение принимается, когда $\cos \alpha' = -1$, т.е. частица движется в сторону, противоположную движению штрихованной системы координат. Сверху правая часть (1.219) ограничена единицей, поскольку справедливо неравенство

$$\left(1 - \frac{v'^2}{c^2}\right) \left(1 - \frac{V^2}{c^2}\right) \leq \left(1 - \frac{v'V}{c^2}\right),$$

которое просто проверяется. Таким образом, правая часть лежит в пределах от 0 до 1. Поэтому $0 \leq v < c$. При этом $v = 0$ тогда и только тогда, когда $v' = V$ и $\cos \alpha' = -1$.

Если $v' = c$, то из уравнения (1.218) вытекает, что $v = c$. $\square$

Из формулы (1.219) следует, что если частица движется в штрихованной системе координат со скоростью V в противоположную сторону, $\cos \alpha' = -1$, то ее скорость в неподвижной системе координат равна нулю. Это означает, что для произвольной частицы, движущейся со скоростью, меньшей скорости света, всегда найдется сопутствующая система координат, в которой она покоится. Ясно также, что для света сопутствующей системы координат не существует.

В разделе 1.7 была введена скорость относительного движения двух инерциальных систем отсчета (1.119). Из этой формулы вытекает, что скорость относительного движения двух инерциальных систем отсчета не превосходит скорости света. Вместе с предложением 1.10.1 это означает, что специальная теория относительности описывает движение частиц со скоростью, меньшей скорости света, самосогласованным образом. Если не постулировать невозможность движения частиц со скоростью, превышающей скорость света, то специальная теория относительности такое движение допускает. Ясно, что для таких частиц не может существовать сопутствующей системы координат. Частицы, движущиеся со скоростью, превышающей скорость света называются *тахionsами*. Такие частицы, если они существуют, представляют большие теоретические трудности. Например, возникает проблема с положительной определенностью канонического гамильтониана. Тем не менее они возникают в различных моделях математической физики. В настоящее время тахионы экспериментально не обнаружены, и принято считать, что в природе они отсутствуют.

Из правила сложения скоростей (1.217) следует также соотношение между углами:

$$\operatorname{tg} \alpha = \frac{\sqrt{v_y^2 + v_z^2}}{v_x} = \frac{v' \sin \alpha'}{v' \cos \alpha' + V} \sqrt{1 - V^2/c^2}.$$

Теперь кратко обсудим два эффекта, которые подтверждаются экспериментально.

Если в системе отсчета O' распространяется луч света, т.е. $v' = c$, то для угла, наблюдаемого в неподвижной системе координат, получаем выражение

$$\operatorname{tg} \alpha = \frac{\sin \alpha'}{\cos \alpha' + V/c} \sqrt{1 - V^2/c^2}. \quad (1.220)$$

Эту формулу можно переписать в более симметричном виде:

$$\sin \alpha = \sin \alpha' \frac{\sqrt{1 - V^2/c^2}}{1 + \frac{V}{c} \cos \alpha'}$$

или

$$\operatorname{tg} \frac{\alpha}{2} = \operatorname{tg} \frac{\alpha'}{2} \sqrt{\frac{1 - V/c}{1 + V/c}}.$$

Из последней формулы, в частности, следует, что $\alpha' > \alpha$ при $V > 0$. Изменение угла, под которым наблюдается луч света, в движущейся и покоящейся системах координат, называется *абберацией света*.

Замечание. Если источник света покоится в системе координат O , а наблюдатель движется вместе с системой отсчета O' , то бесконечно удаленный источник света будет казаться ему смещенным на угол $\alpha' - \alpha$. Для наблюдательной астрономии это означает, что звезды на небе будут описывать эллипсы в соответствии с годовым вращением Земли вокруг Солнца. Этот эффект действительно наблюдается. $\square$

В качестве еще одного приложения преобразований Лоренца рассмотрим *эффект Доплера*. Допустим, что луч света (плоская электромагнитная волна) распространяется в покоящейся системе координат O в плоскости x, y под углом α к оси x . В такой волне напряженность электрического поля $E(t, x)$ меняется по правилу

$$E(t, x) = E_0 e^{i\varphi},$$

где $E_0 = \text{const}$ – амплитуда электрического поля, и

$$\varphi(t, x, y) = \omega \left(t - \frac{x \cos \alpha + y \sin \alpha}{c} \right)$$

– фаза электромагнитной волны. В приведенной формуле ω – это частота волны, которая наблюдается в покоящейся системе координат. Электромагнитное поле, которое описывает плоскую волну удовлетворяет уравнениям Максвелла. Поскольку уравнения Максвелла ковариантны относительно преобразований Лоренца и напряженность E подчиняется тензорному закону преобразования, то фаза волны φ является скалярным полем на пространстве Минковского. Следовательно, фазы волны в покоящейся и движущейся системах координат совпадают. Для штрихованной системы координат, которая равномерно движется вдоль оси x и была описана ранее, равенство фаз имеет вид

$$\omega \left(t - \frac{x \cos \alpha + y \sin \alpha}{c} \right) = \omega' \left(t' - \frac{x' \cos \alpha' + y' \sin \alpha'}{c} \right),$$

где ω' – частота электромагнитной волны в движущейся системе координат. Полученное уравнение должно выполняться при всех значениях t, x, y и t', x', y' . Подставляя

выражения для t, x, y из преобразований Лоренца (1.213) и приравнивая коэффициенты при t', x' и y' , получаем формулы преобразования для частоты

$$\omega' = \omega \frac{1 - \frac{V}{c} \cos \alpha}{\sqrt{1 - V^2/c^2}} \quad (1.221)$$

и углов

$$\cos \alpha' = \frac{\cos \alpha - V/c}{1 - \frac{V}{c} \cos \alpha}, \quad \sin \alpha' = \sin \alpha \frac{\sqrt{1 - V^2/c^2}}{1 - \frac{V}{c} \cos \alpha}.$$

Из последних двух равенств следует соотношение между углами:

$$\operatorname{tg} \alpha' = \frac{\sin \alpha}{\cos \alpha - V/c} \sqrt{1 - V^2/c^2}.$$

Отличие в знаках в знаменателе полученной формулы и равенства (1.220) обусловлено тем, что углы α и α' поменялись местами, что соответствует изменению знака скорости, $V \rightarrow -V$.

Изменение частоты света (1.221), связанное с движением системы отсчета, называется *эффектом Доплера*. Пусть источник света находится в покоящейся системе отсчета O , и свет распространяется вдоль оси x . Допустим, что наблюдатель находится в движущейся системе отсчета O' и удаляется от источника света, т.е. $V > 0$ и $\cos \alpha = 1$. Тогда наблюдаемая им частота света уменьшится. Этот эффект называется *красным смещением*. Если же, наоборот, наблюдатель приближается к источнику света, $V > 0$ и $\cos \alpha = -1$, то частота света увеличится. Такое поведение частоты называется *голубым смещением*.

Замечание. Изучение спектра далеких галактик показывает, что он смещен в красную сторону. Это приводит к выводу о том, что галактики удаляются от Земли и вселенная расширяется. $\square$

1.10.5 Равноускоренное движение

В настоящем разделе мы определим 4-скорость и наблюдаемую скорость точечной частицы, а также ее ускорение. Затем рассмотрим равноускоренное движение в специальной теории относительности и сравним его с равноускоренным движением в механике Ньютона.

Обозначим инерциальные координаты в пространстве Минковского $\mathbb{R}^{1,3}$ через x^α , $\alpha = 0, 1, 2, 3$. Пусть задана мировая линия частицы

$$\mathbb{R} \ni \sigma \mapsto \{q^\alpha(\sigma)\} \in \mathbb{R}^{1,3},$$

где σ – некоторый параметр. По определению, мировая линия частицы времениподобна, а ее длина s при $\sigma \in [0, \tau]$ равна интегралу

$$s := \int_0^\tau d\sigma \sqrt{\eta_{\alpha\beta} \frac{dq^\alpha}{d\sigma} \frac{dq^\beta}{d\sigma}}.$$

Выберем длину траектории s , которая пропорциональна собственному времени частицы, в качестве параметра вдоль мировой линии. Тогда мировая линия задается четырьмя функциями $q^\alpha(s)$.

Для времениподобных кривых в качестве параметра вдоль кривой можно выбрать также время x^0 . Действительно, поскольку для произвольной траектории частицы $dq^0/ds > 0$, то параметр s можно рассматривать как некоторую функцию координаты $s = s(x^0)$. Тогда мировая линия частицы задается функциями

$$q^0 = x^0, \quad q^\mu = q^\mu(q^0), \quad \mu = 1, 2, 3.$$

Замечание. Физическая интерпретация приведенных параметризаций мировых линий следующая. Время для наблюдателя, который движется вместе с частицей, совпадает с собственным временем s/c . Поэтому проведенные им измерения соответствуют сопутствующей системе координат. Во втором случае x^0 – это время наблюдателя, покоящегося в выбранной исходной инерциальной системе отсчета. Поэтому его измерения соответствуют внешнему наблюдателю. $\square$

Определение. 4-вектор u с компонентами

$$u^\alpha := c \frac{dq^\alpha}{ds} \quad (1.222)$$

называется 4-*скоростью* или просто *скоростью* частицы. 4-вектор w с компонентами

$$w^\alpha := c \frac{du^\alpha}{ds} = c^2 \frac{d^2 q^\alpha}{ds^2} \quad (1.223)$$

называется 4-*ускорением* или просто *ускорением* частицы. *Наблюдаемой скоростью* частицы $\mathbf{v}$ в системе координат x^α называется 3-вектор с компонентами

$$v^\mu := \frac{dq^\mu}{dq^0} = \frac{u^\mu}{u^0}, \quad \mu = 1, 2, 3. \quad (1.224)$$

3-вектор $\mathbf{a}$ с компонентами

$$a^\mu := \frac{dv^\mu}{dq^0} = \frac{d^2 q^\mu}{(dq^0)^2}. \quad (1.225)$$

называется *наблюдаемым ускорением* частицы. $\square$

Скорость (1.222) и ускорение (1.223) точечной частицы являются 4-векторами относительно преобразований из группы Лоренца $\mathbb{O}(1, 3)$, действующей в пространстве Минковского $\mathbb{R}^{1,3}$. Наблюдаемая скорость (1.224) и наблюдаемое ускорение (1.225) преобразуются как векторы только относительно подгруппы вращений $\mathbb{O}(3)$, действующей на пространственных сечениях $x^0 = \text{const}$. Они нековариантны относительно лоренцевых преобразований и не являются компонентами каких-либо 4-векторов.

Мы предполагаем, что траектория частицы времениподобна и направлена в будущее. Поэтому временная компонента скорости положительна, $u^0 > 0$.

В настоящем разделе полезно следить за размерностями различных величин:

$$\begin{aligned} [x^0] &= \text{сек}, & [x^\mu] &= [s] = \text{см}, \\ [u^0] &= 1, & [u^\mu] &= [v^\mu] = \frac{\text{см}}{\text{сек}}, \\ [w^0] &= \frac{1}{\text{сек}}, & [w^\mu] &= [a^\mu] = \frac{\text{см}}{\text{сек}^2}. \end{aligned}$$

Из определения скорости (1.222) сразу следует равенство

$$u^\alpha u_\alpha = c^2 (u^0)^2 - \mathbf{u}^2 = c^2, \quad (1.226)$$

где $\mathbf{u}^2 = -u^\mu u_\mu \geq 0$ – обычный квадрат трехмерного вектора в евклидовом пространстве. Дифференцирование этого равенства по s приводит к соотношению

$$u^\alpha w_\alpha = 0,$$

т.е. 4-ускорение частицы всегда ортогонально ее 4-скорости.

Умножим равенство (1.226) на $m^2 c^2$, где $m > 0$ – масса частицы, $[m] = \Gamma$, и перепишем его в виде

$$E = mc^2 \sqrt{1 + \frac{\mathbf{u}^2}{c^2}},$$

где введено обозначение $E := mc^2 u^0 > 0$. В нерелятивистском пределе, когда пространственные компоненты скорости малы по сравнению со скоростью света, $\mathbf{u}^2/c^2 \ll 1$, квадратный корень можно разложить в ряд Тейлора. В первом приближении получим равенство

$$E = mc^2 + \frac{m\mathbf{u}^2}{2}.$$

Поскольку второе слагаемое совпадает с выражением для кинетической энергии точечной частицы в механике Ньютона, то отсюда следует, что величина E представляет собой энергию частицы в специальной теории относительности. В сопутствующей системе координат, где $\mathbf{u} = 0$, получаем знаменитую формулу Эйнштейна

$$E = mc^2,$$

которая дает выражение для энергии покоя точечной частицы. Формулу (1.226) можно переписать также в виде

$$\frac{E^2}{c^2} - \mathbf{p}^2 = m^2 c^2,$$

где введен пространственный импульс частицы $\mathbf{p} := m\mathbf{u}$. Если ввести 4-импульс частицы

$$p^\alpha := mu^\alpha = (mu^0, m\mathbf{u}) = \left(\frac{E}{c^2}, \mathbf{p} \right),$$

то равенство (1.226) примет вид

$$p^2 = p^\alpha p_\alpha = m^2 c^2.$$

Согласно определению (1.224) для наблюдаемой скорости выполнено равенство $u^\mu = u^0 v^\mu$. Поэтому из уравнения (1.226) вытекает соотношение

$$u^0 = \frac{1}{\sqrt{1 - \mathbf{v}^2/c^2}}. \quad (1.227)$$

Заметим, что в нерелятивистском пределе, который мы определили как $\mathbf{u}^2/c^2 \rightarrow 0$, из равенства (1.226) вытекает, что при этом $u^0 \rightarrow 1$. Поэтому из соотношения $u^\mu = u^0 v^\mu$ следует, что пределы $\mathbf{u}^2 \rightarrow 0$ и $\mathbf{v}^2 \rightarrow 0$ эквивалентны.

Теперь найдем формулы, связывающие 4-ускорение с наблюдаемым ускорением частицы. Для этого в определении 4-ускорения надо заменить производную

$$c \frac{d}{ds} = u^0 \frac{d}{dq^0}$$

и воспользоваться формулой (1.227). В результате получим равенства

$$\begin{aligned} w^0 &= \frac{(\mathbf{v}, \mathbf{a})}{c^2(1 - \mathbf{v}^2/c^2)^2}, \\ \mathbf{w} &= \mathbf{a} \frac{1}{1 - \mathbf{v}^2/c^2} + \mathbf{v} \frac{(\mathbf{v}, \mathbf{a})}{c^2} \frac{1}{(1 - \mathbf{v}^2/c^2)^2} \end{aligned} \quad (1.228)$$

В нерелятивистском пределе, очевидно, $\mathbf{w} = \mathbf{a}$ и $w^0 = 0$.

Для определения равноускоренного движения нам понадобятся формулы преобразования компонент наблюдаемого ускорения при переходе от покоящейся системы координат к движущейся. Как было отмечено, компоненты наблюдаемого ускорения не являются компонентами какого-либо 4-вектора, и поэтому при лоренцевых бустах преобразуются не по тензорным правилам. Пусть система координат O' связана с частицей, т.е. в момент времени $ct = x^0$, ее скорость совпадает с мгновенной скоростью частицы, $\mathbf{V} = \mathbf{v}$. Тогда в штрихованной системе координат скорость $\mathbf{v}' = 0$, и формулы (1.228) принимают вид

$$\mathbf{w}' = \mathbf{a}', \quad w'^0 = 0. \quad (1.229)$$

Для определенности, предположим, что скорость $\mathbf{v}$ направлена вдоль оси x , т.е. $\mathbf{v} = (v^1, 0, 0)$ и $\mathbf{v}^2 = (v^1)^2$. Тогда из формул (1.228) следуют соотношения:

$$w^1 = \frac{a^1}{1 - \mathbf{v}^2/c^2} + v^1 w^0, \quad w^2 = \frac{a^2}{1 - \mathbf{v}^2/c^2}, \quad w^3 = \frac{a^3}{1 - \mathbf{v}^2/c^2}. \quad (1.230)$$

Теперь вспомним, что 4-ускорение w – лоренцев вектор. Его компоненты преобразуются, как компоненты вектора (x^0, x^1, x^2, x^3) . Поэтому компоненты 4-ускорения в штрихованной и нештрихованной системах координат связаны преобразованием Лоренца (1.213):

$$w^0 = \frac{v^1 w'^1}{c \sqrt{1 - \mathbf{v}^2/c^2}}, \quad w^1 = \frac{w'^1}{\sqrt{1 - \mathbf{v}^2/c^2}}, \quad w^2 = w'^2, \quad w^3 = w'^3,$$

где мы учли, что $w'^0 = 0$. Подставляя в эти формулы соотношения между 4-ускорениями и наблюдаемыми ускорениями (1.229) и (1.230) окончательно получаем формулы преобразования наблюдаемых ускорений:

$$a^1 = a'^1(1 - \mathbf{v}^2/c^2)^{3/2}, \quad a^2 = a'^2(1 - \mathbf{v}^2/c^2), \quad a^3 = a'^3(1 - \mathbf{v}^2/c^2). \quad (1.231)$$

Эти формулы преобразования, как и следовало ожидать, отличаются от преобразований Лоренца.

Рассмотрим равноускоренное движение.

Определение. Движение точечной частицы называется *равноускоренным*, если наблюдаемое ускорение $\mathbf{a}'$ в системе координат O' , движущейся вместе с данной частицей, постоянно, $\mathbf{a}' = \text{const}$. $\square$

Теперь найдем мировую линию частицы, которая движется равноускоренно в покоящейся системе координат. Пусть система отсчета O фиксирована. Поскольку частица движется равноускоренно, то штрихованная система координат O' , связанная с частицей, меняется с течением времени.

Рассмотрим случай прямолинейного движения, когда ускорение $\mathbf{a}'$ постоянно и параллельно скорости $\mathbf{v}$. Для этого достаточно, чтобы ускорение было постоянным и частица в начальный момент времени покоилась. Для определенности, предположим, что ускорение и скорость параллельны оси $x = x^1$, т.е. мировая линия частицы лежит в плоскости t, x . Положим $\mathbf{a}' = (b, 0, 0)$, где $b = \text{const}$, и $\mathbf{v} = (v(t), 0, 0)$. Тогда из (1.231) следует одно дифференциальное уравнение

$$\frac{dv}{dt} = b(1 - v^2/c^2)^{3/2}.$$

Оно легко интегрируется:

$$b(t - t_0) = \frac{v}{\sqrt{1 - v^2/c^2}},$$

где t_0 – постоянная интегрирования. Подставляя равенство $v = dx/dt$, получаем дифференциальное уравнение на функцию $x(t)$. Интегрирование полученного уравнения дает уравнение на мировую линию частицы:

$$(x - x_0)^2 - c^2(t - t_0)^2 = \frac{c^4}{b^2},$$

где x_0 – вторая постоянная интегрирования. Таким образом, мировая линия частицы при прямолинейном равноускоренном движении представляет собой гиперболу в плоскости движения t, x . При $t_0 = 0$ и $x_0 = 0$ мировая линия показана на рис.1.13. Для сравнения показана также мировая линия частицы в механике Ньютона, которая задана хорошо известным уравнением

$$x = \frac{c^2}{b} + \frac{1}{2}bt^2$$

и представляет собой параболу.

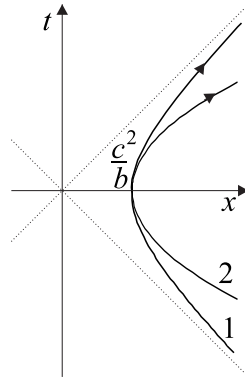


Рис. 1.13: Мировые линии частиц при прямолинейном равноускоренном движении в специальной теории относительности (1) и в механике Ньютона (2).

При равноускоренном движении наблюдаемая скорость частицы при $t \rightarrow \infty$ в специальной теории относительности стремится к скорости света, а в механике Ньютона – к бесконечности.

Глава 2

Многообразия и тензорные поля

В настоящей главе начинается изложение некоторых аспектов современной дифференциальной геометрии. Изложение часто носит координатный характер, т.е. все соотношения записываются в локальной системе координат и приводятся правила преобразований геометрических объектов при переходе от одной системы координат к другой. Параллельно даются также глобальные определения, не зависящие от выбора системы координат, и показывается связь координатного и инвариантного описания. Глобальный бескоординатный подход компактен в обозначениях, прозрачен и удобен для определения геометрических объектов. Координатный подход является более громоздким, однако это не является признаком меньшей строгости. Он незаменим в моделях математической физики, основанных на дифференциальной геометрии, где требуется проведение расчетов.

2.1 Многообразия

Базовым понятием дифференциальной геометрии является дифференцируемое многообразие $\mathbb{M}$, которое обобщает понятие евклидова пространства $\mathbb{R}^n$. Топологически нетривиальные многообразия не покрываются одной системой координат, однако локально устроены так же, как и евклидовы пространства. Это позволяет использовать математический анализ для построения и анализа многих важных моделей современной математической физики.

Определение. Топологическое хаусдорфово пространство $\mathbb{M}$ со счетной базой, каждая точка которого имеет окрестность, гомеоморфную открытому n -мерному шару единичного радиуса $\mathbb{B}_1^n$ в евклидовом пространстве $\mathbb{R}^n$, называется n -мерным *топологическим многообразием*. Число n называется *размерностью* многообразия. Мы пишем $\dim \mathbb{M} = n$. □

Замечание. В определении многообразия требование хаусдорфовости существенно и исключает случаи подобные примерам на с. 28. Согласно теореме 1.3.11, любая сходящаяся последовательность в хаусдорфовом пространстве имеет не более одного предела.

Предположение о счетности базы также существенно, потому что это требование обеспечивает паракомпактность многообразия, которая в рассматриваемом случае достаточна для существования разбиения единицы. В свою очередь, существование разбиения единицы важно, т.к. позволяет определять геометрические объекты на

всем многообразии, исходя из их задания в локальных координатах. Оно часто используется для доказательства теорем существования.

Выбор шаров единичного радиуса $\mathbb{B}_1^n$ сделан для определенности и не является существенным, т.к. шар единичного радиуса можно заменить на произвольную гомеоморфную ему область, в частности, на все евклидово пространство $\mathbb{R}^n$.

В настоящей монографии под областью, как правило, мы понимаем произвольное открытое подмножество, гомеоморфное открытому шару в $\mathbb{R}^n$, которое, следовательно, можно покрыть одной системой координат. □

Замечание. Поскольку многообразие, по определению, является хаусдорфовым топологическим пространством, то для многообразий понятие компактного пространства и компакта совпадают. □

Окрестность произвольной точки многообразия $\mathbb{M}$ устроена так же, как и окрестность точки в евклидовом пространстве $\mathbb{R}^n$. Однако в отличие от евклидова пространства, которое, по определению, может быть покрыто одной картой, многообразие в общем случае одной картой не покрывается. Поэтому на многообразии общего вида нельзя ввести структуру векторного или аффинного пространства.

Предложение 2.1.1. *Многообразие представляет собой объединение конечного или счетного числа областей, $\mathbb{M} = \bigcup_i U_i$, каждая из которых гомеоморфна n -мерному шару, и, следовательно, всему $\mathbb{R}^n$.*

Доказательство. Поскольку каждая точка имеет окрестность, гомеоморфную $\mathbb{R}^n$, то все многообразие можно покрыть, возможно, несчетным числом карт U_i . Выберем счетную базу на $\mathbb{M}$, которая также является покрытием. Каждая координатная окрестность U_i является объединением счетного числа элементов базы и, поэтому, на каждом элементе базы задан гомеоморфизм в $\mathbb{R}^n$. Теперь можно выбрать базу топологии в качестве координатного покрытия, а она – счетна. □

В приложениях рассматриваются, как правило, многообразия, которые покрываются конечным числом карт.

Замечание. Поскольку многообразие представляет собой объединение не более, чем счетного числа областей, гомеоморфных n -мерному шару, то любое многообразие является сепарабельным топологическим пространством. □

Согласно данному определению, многообразие, как объединение открытых множеств, не имеет границы, которую в случае многообразий принято называть краем (см. раздел 2.3).

Поскольку каждая окрестность многообразия гомеоморфна шару в евклидовом пространстве, то всякое многообразие является локально связным. Согласно теореме 1.3.15 всякое многообразие представляет собой объединение связных компонент, каждая из которых является одновременно открытой и замкнутой. Число этих компонент может быть не более, чем счетным, т.к. мы предполагаем счетность базы многообразия. В дальнейшем, однако, мы иногда будем рассматривать многообразия, состоящие из несчетного числа компонент.

Определение. Под 0-мерным многообразием мы будем понимать счетное множество точек с дискретной топологией. Одномерное и двумерное многообразия называются, соответственно, *кривой* и *поверхностью*. □

Размерность многообразия является топологическим инвариантом: два гомеоморфных многообразия имеют одинаковую размерность. Это утверждение является прямым следствием теоремы 1.4.1 об инвариантности размерности.

Данное выше определение задает *топологическое многообразие*, т.к. в нем говорится только о непрерывности. Теперь перейдем к описанию дифференцируемых многообразий.

Определение. Из определения многообразия следует, что существует гомеоморфизм (биекция, непрерывная в обе стороны)

$$\varphi_i : \mathbb{M} \supset \mathbb{U}_i \rightarrow \varphi_i(\mathbb{U}_i) \subset \mathbb{R}^n,$$

области $\mathbb{U}_i$ на ее образ $\varphi_i(\mathbb{U}_i)$ (суммирования нет) в евклидовом пространстве $\mathbb{R}^n$. Поскольку в евклидовом пространстве есть система координат, например, декартова, то данный гомеоморфизм можно записать в виде

$$\varphi_i : \mathbb{M} \supset \mathbb{U}_i \ni x \mapsto \varphi_i(x) = (x^1, \dots, x^n) \in \varphi_i(\mathbb{U}_i) \subset \mathbb{R}^n,$$

где точку многообразия x и ее координаты $(x^1, \dots, x^n)$ мы обозначили одной и той же буквой. Области $\mathbb{U}_i$, покрывающие многообразие, называются *координатными окрестностями*, а набор чисел $(x^1, \dots, x^n)$ – *локальными координатами*. Пара $(\mathbb{U}_i, \varphi_i)$ называется *картой*.

Если две карты пересекаются, $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$, то произвольная точка из пересечения $x \in \mathbb{U}_i \cap \mathbb{U}_j$ имеет свой набор координат в каждой карте. Отображение областей евклидова пространства,

$$f_{ji} := \varphi_j \circ \varphi_i^{-1} : \mathbb{R}^n \supset \varphi_i(\mathbb{U}_i \cap \mathbb{U}_j) \rightarrow \varphi_j(\mathbb{U}_i \cap \mathbb{U}_j) \subset \mathbb{R}^n,$$

задается набором n функций $x^{\alpha'}(x)$ от n переменных (1.66), где x^α и $x^{\alpha'}$ ($\alpha, \alpha' = 1, \dots, n$) – координаты, соответственно, на $\mathbb{U}_i$ и $\mathbb{U}_j$. Они называются *функциями склейки*, поскольку склеивают между собой различные карты. Совокупность всех карт, покрывающих многообразие, $\mathbb{M} = \bigcup_i \mathbb{U}_i$, называется *координатным покрытием* или *атласом* $\{\mathbb{U}_i, \varphi_i\}$, $i \in I$ многообразия $\mathbb{M}$. Атлас, который не содержится ни в каком другом атласе, называется *полным*. $\square$

При проведении вычислений в одной карте точку многообразия $x \in \mathbb{M}$ и ее координаты $\varphi(x) \in \mathbb{R}^n$ можно отождествлять. Но всегда следует помнить, что точка одна, а координат много.

Замечание. В силу счетности топологической базы многообразия, координатное покрытие всегда можно выбрать счетным. Однако, даже евклидово пространство можно при желании покрыть несчетным числом карт. Это значит, что полный атлас любого многообразия всегда несчетен. $\square$

Мы требуем, чтобы каждая функция склейки $f_{ji} := \varphi_j \circ \varphi_i^{-1}$ заданного атласа определяла некоторый диффеоморфизм, т.е. являлась элементом псевдогруппы преобразований координат $\text{diff}^k(\mathbb{R}^n)$ для некоторого k , определенной в разделе 1.5. Поэтому мы говорим, что атлас $\{\mathbb{U}_i, \varphi_i\}$ совместим с псевдогруппой преобразований координат $\text{diff}^k(\mathbb{R}^n)$.

Определение. Многообразие $\mathbb{M}$ вместе с полным атласом $\{\mathbb{U}_i, \varphi_i\}$ называется *дифференцируемым многообразием* класса C^k , $k \in \mathbb{N}$, если функции склейки (1.66) для всех пересекающихся карт непрерывны вместе со своими частными производными k -того порядка: $f_{ji} := \varphi_j \circ \varphi_i^{-1} \in \text{diff}^k(\mathbb{R}^n)$. Полный атлас $\{\mathbb{U}_i, \varphi_i\}$ называется *дифференцируемой структурой* многообразия $\mathbb{M}$. $\square$

Аналогично определяются гладкие C^∞ и вещественно аналитические многообразия C^ω . На C^ω -многообразиях функции склейки задаются сходящимися степенными рядами. Напомним, что функции склейки имеют ненулевой якобиан и поэтому осуществляют взаимно однозначное отображение областей в $\mathbb{R}^n$ (теорема 1.5.1). Это означает, что обратное преобразование существует, и дифференцируемость обратных функций такая же, как и самих функций склейки. В дальнейшем, если не оговорено противное, под многообразием мы будем понимать C^∞ дифференцируемые многообразия. Многие из рассмотренных ниже утверждений справедливы и при более слабых ограничениях на дифференцируемость функций склейки. Для проверки достаточно следить лишь за числом производных при вычислениях.

Замечание. Отображения φ_i являются гомеоморфизмами. Это означает, что дифференцируемые структуры, согласованы с естественной топологией евклидова пространства $\mathbb{R}^n$, т.е. можно считать, что топология многообразия индуцирована отображениями областей U_i в евклидово пространство $\mathbb{R}^n$ с естественной топологией. $\square$

Замечание. В определение дифференцируемого многообразия входит понятие евклидова пространства $\mathbb{R}^n$. В евклидовом пространстве можно определить много различных структур. Для определения топологического многообразия нам достаточно наличия топологии в $\mathbb{R}^n$. Для определения дифференцируемого многообразия нам достаточно дополнительно предположить наличие структуры поля на вещественной прямой $\mathbb{R}$ (для определения производной необходимо вычитать и делить числа). Конечно, в евклидовом пространстве всегда можно ввести евклидову метрику и структуру векторного пространства. Однако для определения дифференцируемого многообразия это не является необходимым условием. $\square$

Отметим некоторые свойства функций склейки.

Предложение 2.1.2. *Функции склейки удовлетворяют тождествам:*

$$f_{ij} = f_{ji}^{-1}, \quad \forall x \in U_i \cap U_j, \quad (2.1)$$

$$f_{ij}f_{jk}f_{ki} = \text{id}, \quad \forall x \in U_i \cap U_j \cap U_k, \quad (2.2)$$

где id – тождественное отображение.

Доказательство. Прямая проверка. $\square$

Следствие. Справедливо тождество $f_{ii} = \text{id}$.

Предложение 2.1.3. *Любой атлас класса C^k можно дополнить до полного атласа того же класса гладкости, и соответствующий полный атлас единственный. Если на многообразии задано два диффеоморфных атласа $\{U_i, \varphi_i\}$ и $\{V_\alpha, \psi_\alpha\}$, т.е. все отображения*

$$\psi_\alpha \circ \varphi_i^{-1} : \varphi_i(U_i \cap V_\alpha) \rightarrow \psi_\alpha(U_i \cap V_\alpha)$$

являются диффеоморфизмами областей евклидова пространства $\mathbb{R}^n$, как только пересечения $U_i \cap V_\alpha$ не пусты, то они содержатся в одном полном атласе. Любые два атласа, содержащиеся в одном полном атласе связаны между собой диффеоморфизмом.

Доказательство. Пусть $\{\mathbb{W}, \chi\}$ – семейство всех пар таких, что χ есть гомеоморфизм открытого подмножества $\mathbb{W} \subset \mathbb{M}$ на открытое подмножество в евклидовом пространстве $\mathbb{R}^n$ и что отображение

$$\varphi_i \circ \chi^{-1} : \chi(\mathbb{W} \cap U_i) \rightarrow \varphi_i(\mathbb{W} \cap U_i)$$

есть элемент $\text{diff}^k(\mathbb{R}^n)$, как только $\mathbb{W} \cap \mathbb{U}_i$ не пусто. Тогда $\{\mathbb{W}, \chi\}$ есть полный атлас, содержащий $\{\mathbb{U}_i, \varphi_i\}$. Этот атлас единственный по построению.

Если заданы два диффеоморфных атласа $\{\mathbb{U}_i, \varphi_i\}$ и $\{\mathbb{V}_\alpha, \psi_\alpha\}$, то их объединение также является атласом и того же класса гладкости. Поэтому они содержатся в одном полном атласе.

Все атласы, содержащиеся в одном полном атласе, получаются путем отбрасывания некоторого количества карт. Они диффеоморфны между собой, поскольку все функции склейки полного атласа принадлежат $\text{diff}^k(\mathbb{R}^n)$. $\square$

Данное предложение означает, что для задания дифференцируемой структуры на многообразии достаточно задать один атлас, не обязательно полный. Если на многообразии $\mathbb{M}$ можно задать два атласа одного класса гладкости, которые не диффеоморфны между собой, то это означает, что на $\mathbb{M}$ существуют различные дифференцируемые структуры.

Определение. Пусть на многообразии $\mathbb{M}$ введено две дифференцируемые структуры разных классов гладкости $\mathcal{C}^k$ и $\mathcal{C}^l$, где $l > k$, причем полный атлас класса $\mathcal{C}^l$ является одновременно атласом класса $\mathcal{C}^k$. Если $l = \infty$, то ограничение класса гладкости $\mathcal{C}^k$ -многообразия до $\mathcal{C}^\infty$ -многообразия называется *сглаживанием*. При этом сглаживания, которые связаны $\mathcal{C}^k$ автоморфизмами, рассматриваются, как эквивалентные. $\square$

Теорема 2.1.1 (Уитни). Каждое $\mathcal{C}^k$ -многообразие имеет ровно одно (с точностью до $\mathcal{C}^k$ -диффеоморфизмов) сглаживание.

Доказательство. См. [36]. $\square$

В силу этой теоремы во многих случаях можно игнорировать различие между $\mathcal{C}^k$ дифференцируемыми и $\mathcal{C}^\infty$ гладкими структурами.

В низших размерностях дифференцируемая структура на многообразиях единственна. Доказано, что двух- и трехмерные многообразия допускают ровно одну дифференцируемую структуру [37]. Для многообразий более высоких размерностей это не так. Данное утверждение иллюстрирует пример Милнора [38], который показал, что на семимерной сфере $\mathbb{S}^7$, рассматриваемой, как топологическое многообразие, существует 28 недиффеоморфных дифференцируемых структур. Построение этих структур является сложным, и мы его не приводим. Было также доказано, что четырехмерное евклидово пространство $\mathbb{R}^4$ допускает бесконечное множество недиффеоморфных дифференцируемых структур [39]. Существует также теорема о том, что каждое гладкое многообразие имеет единственную вещественно аналитическую структуру [36].

Если в определении многообразия отбросить требование счетности базы, то получим *локально евклидово топологическое пространство*. Существуют такие топологические локально евклидовы пространства, которые вообще не допускают дифференцируемой структуры [40].

Любое открытое подмножество $\mathbb{D}$ дифференцируемого многообразия $\mathbb{M}$ само является многообразием. Дифференцируемая структура на $\mathbb{D}$ состоит из карт $(\mathbb{U}_i \cap \mathbb{D}, \varphi_i|_{\mathbb{D}})$, возникающих после сужения гомеоморфизмов φ_i на $\mathbb{D}$.

Поскольку многообразия являются топологическими пространствами, то их можно умножать как топологические пространства. При этом произведение двух многообразий $\mathbb{M} \times \mathbb{N}$ также является многообразием размерности $m + n$, где $\dim \mathbb{M} = m$

и $\dim \mathbb{N} = n$. Дифференцируемая структура на прямом произведении многообразий $\mathbb{M} \times \mathbb{N}$ строится, как прямое произведение дифференцируемых структур на $\mathbb{M}$ и $\mathbb{N}$. А именно, пусть дифференцируемая структура на многообразии $\mathbb{M}$ определяется атласом $\{\mathbb{U}_i, \varphi_i\}$, а на $\mathbb{N}$ – атласом $\{\mathbb{V}_A, \phi_A\}$. Тогда естественная дифференцируемая структура на топологическом произведении $\mathbb{M} \times \mathbb{N}$ определяется атласом $\{\mathbb{U}_i \times \mathbb{V}_A, \varphi_i \times \phi_A\}$, где отображение

$$\varphi_i \times \phi_A : \mathbb{U}_i \times \mathbb{V}_A \rightarrow \mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$$

определяется естественным образом. Отметим, что этот атлас не будет полным даже если атласы $\{\mathbb{U}_i, \varphi_i\}$ и $\{\mathbb{V}_A, \phi_A\}$ полны.

Определение. Если связное многообразие покрыто совокупностью карт $\{\mathbb{U}_i, \varphi_i\}$ с координатами $\varphi_i(x) = \{x_i^\alpha\}$, причем якобианы функций перехода (1.69) для всех пересекающихся карт $\mathbb{U}_i$ и $\mathbb{U}_j$ положительны,

$$\det \left(\frac{\partial x_i^\alpha}{\partial x_j^\beta} \right) > 0, \quad \forall i, j$$

то многообразие называется *ориентированным*. Многообразие называется *неориентируемым*, если атласа со всеми положительными якобианами функций склейки не существует. При неудачно выбранном атласе на ориентируемом многообразии якобианы могут быть разных знаков, однако атлас с положительными якобианами существует. Такие многообразия называются *ориентируемыми*. Несвязное многообразие $\mathbb{M}$ называется ориентируемым, если ориентируема каждая его компонента. $\square$

Связное ориентируемое многообразие допускает в точности две ориентации. Чтобы поменять ориентацию ориентированного многообразия достаточно заменить каждую карту $(\mathbb{U}_i, \varphi_i)$ ориентированного атласа на карту $(\mathbb{U}_i, \psi_i)$, где гомеоморфизм ψ_i является композицией φ_i и отражения первой (или любой другой) координаты: $(x^1, x^2, \dots, x^n) \mapsto (-x^1, x^2, \dots, x^n)$. (Более подробно ориентация многообразий будет обсуждаться в разделе 10.4.) Ориентация несвязного многообразия – это выбор ориентации на каждой компоненте.

Рассмотрим простейшие примеры многообразий.

Пример 2.1.1. Все евклидово пространство $\mathbb{R}^n$ является простейшим n -мерным многообразием, которое можно покрыть одной картой (а можно и несколькими). Дифференцируемая структура – это полный атлас, содержащий естественную карту $(\mathbb{U} = \mathbb{R}^n, \varphi = \text{id}(\mathbb{R}^n))$. При этом класс гладкости многообразия определяется классом гладкости допустимых преобразований координат. Любое многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, которое покрывается одной картой, диффеоморфно (см. раздел 2.9) $\mathbb{R}^n$ и называется *тривиальным*. $\square$

Пример 2.1.2. Рассмотрим произвольное вещественное векторное пространство $\mathbb{V}$, $\dim \mathbb{V} = n$. Если в векторном пространстве выбран базис $\{e_a\}$, $a = 1, \dots, n$, то каждая точка векторного пространства задается упорядоченным набором вещественных чисел $(x^1, \dots, x^n)$. Тогда его можно отождествить с евклидовым пространством $\mathbb{R}^n$ и рассматривать, как гладкое многообразие. Дифференцируемая структура в $\mathbb{V}$ не зависит от выбора базиса, т.к. замена базиса задается невырожденной матрицей, и соответствует преобразованию координат класса C^∞ . В дальнейшем мы всегда будем считать, что все векторные пространства снабжены естественной дифференцируемой структурой евклидова пространства. $\square$

Пример 2.1.3. Рассмотрим ломаную линию γ на евклидовой плоскости $\mathbb{R}^2$, заданную уравнением $y = |x|$ (см. рис. 2.1a). Топология, индуцированная из $\mathbb{R}^2$, превращает ее в связное хаусдорфово топологическое пространство. Ломаную γ можно покрыть одной картой $(\mathbb{U}, \varphi)$, спроектировав ломаную линию на ось x ,

$$\mathbb{U} = \{(x, y) \in \mathbb{R}^2 : y = |x|\}, \quad \varphi(x, y) = x.$$

Таким образом γ становится одномерным тривиальным многообразием класса $\mathcal{C}^\infty$. Однако вложение $\gamma \hookrightarrow \mathbb{R}^2$ (см. раздел 2.9) является только непрерывным, а не дифференцируемым, и это является причиной излома. $\square$

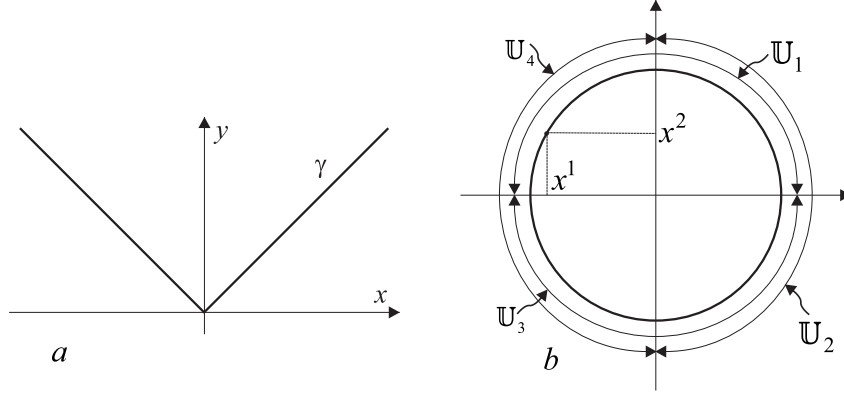


Рис. 2.1: Ломаная прямая γ является $\mathcal{C}^\infty$ многообразием (a). Покрытие окружности S^1 четырьмя картами (b). Показана точка, принадлежащая пересечению $U_1 \cap U_4$.

Пример 2.1.4. n -мерная сфера $\mathbb{S}_r^n \hookrightarrow \mathbb{R}^{n+1}$ радиуса r с центром в начале координат, вложенная в $(n+1)$ -мерное евклидово пространство, задается уравнением

$$\mathbb{S}_r^n := \{x \in \mathbb{R}^{n+1} : (x^1)^2 + (x^2)^2 + \dots + (x^{n+1})^2 = r^2\}, \quad r = \text{const} > 0. \quad (2.3)$$

Само по себе это уравнение задает только множество точек в $\mathbb{R}^{n+1}$, а никак не многообразие. Зададим на нем топологию, сказав, что топология индуцирована вложением $\mathbb{S}_r^n \hookrightarrow \mathbb{R}^{n+1}$. Пусть $N = (0, \dots, 0, r)$ и $S = (0, \dots, 0, -r)$ – северный и южный полюс сферы. Гладкая дифференцируемая структура на $\mathbb{S}_r^n$ – это полный атлас, содержащий две карты: $(\mathbb{S}_r^n \setminus N, \varphi_N)$ и $(\mathbb{S}_r^n \setminus S, \varphi_S)$, где φ_N и φ_S – стереографические проекции из северного и южного полюса. Тогда сфера становится n -мерным компактным ориентируемым многообразием. Это многообразие нетривиально и покрывается не менее, чем двумя картами. $\square$

Пример 2.1.5. В примере 1.2.16, была построена сфера Римана $\overline{\mathbb{C}}$. Она представляет собой вещественное двумерное многообразие, диффеоморфное обычной сфере, $\overline{\mathbb{C}} \approx S^2$. С точки зрения комплексной геометрии она является одномерным голоморфным комплексным многообразием комплексной размерности $\dim_{\mathbb{C}} \overline{\mathbb{C}} = 1$. Действительно, покроем сферу Римана двумя картами: $U_0 := \mathbb{C}$ – окрестность нуля и $U_\infty := \overline{\mathbb{C}} \setminus \{0\}$ – окрестность бесконечности. Они имеют непустое пересечение $U_0 \cap U_\infty = \mathbb{C} \setminus \{0\}$. В качестве координат на U_0 и U_∞ выберем, соответственно, z и $1/z$. В области пересечения функция склейки $f_{0\infty}(z) = 1/z$ голоморфна. Тем самым мы построили голоморфный атлас на $\overline{\mathbb{C}}$.

Как комплексное многообразие сфера Римана диффеоморфна одномерному комплексному проективному пространству (проективной прямой), $\overline{\mathbb{C}} \approx \mathbb{CP}^1$. Напомним,

что проективная прямая – это множество комплексных прямых $az^1 + bz^2 = 0$, где $(z^1, z^2) \in \mathbb{C} \times \mathbb{C}$ и $a, b \in \mathbb{C}$, в двумерном комплексном многообразии $\mathbb{C}^2 := \mathbb{C} \times \mathbb{C}$. Если $b \neq 0$, то из уравнения прямой следует равенство $z^2 = -az^1/b$, т.е. каждая прямая параметризуется комплексным числом $-a/b = z \in \mathbb{C}$. При $a \neq 0$, аналогично, каждая прямая параметризуется числом $-b/a = 1/z \in \mathbb{C}$. Тем самым мы покрыли проективное пространство $\mathbb{CP}^1$ двумя картами. В области пересечения $ab \neq 0$ функция склейки имеет вид $f(z) = 1/z$. Таким образом, сфера Римана и одномерное комплексное проективное пространство имеют одинаковое координатное покрытие и, следовательно, диффеоморфны. $\square$

Пример 2.1.6. Рассмотрим задание окружности $\mathbb{S}^1 \hookrightarrow \mathbb{R}^2$ единичного радиуса с помощью четырех карт:

$$\begin{aligned} \mathbb{U}_1 &= \{x \in \mathbb{S}^1 : x^2 > 0\}, & \varphi_1(x) &= x^1, \\ \mathbb{U}_2 &= \{x \in \mathbb{S}^1 : x^1 > 0\}, & \varphi_2(x) &= x^2, \\ \mathbb{U}_3 &= \{x \in \mathbb{S}^1 : x^2 < 0\}, & \varphi_3(x) &= x^1, \\ \mathbb{U}_4 &= \{x \in \mathbb{S}^1 : x^1 < 0\}, & \varphi_4(x) &= x^2. \end{aligned}$$

Очевидно, что совокупность областей $\{\mathbb{U}_1, \mathbb{U}_2, \mathbb{U}_3, \mathbb{U}_4\}$ является конечным открытым покрытием окружности $\mathbb{S}^1$. Четыре карты $\{(\mathbb{U}_1, \varphi_1), (\mathbb{U}_2, \varphi_2), (\mathbb{U}_3, \varphi_3), (\mathbb{U}_4, \varphi_4)\}$ представляют собой атлас. В области пересечения двух карт $\mathbb{U}_1 \cap \mathbb{U}_4$ функции преобразования координат имеют вид (см. рис. 2.1b)

$$\begin{aligned} x^{2'} &= \sqrt{1 - (x^1)^2} > 0, \\ x^{1'} &= -\sqrt{1 - (x^2)^2} < 0, \end{aligned}$$

где нештрихованные и штрихованные координаты относятся, соответственно, к областям $\mathbb{U}_1$ и $\mathbb{U}_4$. Аналогично выписываются функции склейки для всех других пересечений карт. Все функции склейки являются бесконечно дифференцируемыми. Это значит, что построенный атлас принадлежит классу $\mathcal{C}^\infty$. $\square$

Пример 2.1.7. Тор $\mathbb{T}^n$ представляет собой прямое произведение n окружностей,

$$\mathbb{T}^n := \underbrace{\mathbb{S} \times \mathbb{S} \times \cdots \times \mathbb{S}}_n,$$

и является n -мерным компактным ориентируемым многообразием. Дифференцируемая структура на $\mathbb{T}^n$ задается как на прямом произведении многообразий. $\square$

Пример 2.1.8. Проективным пространством $\mathbb{RP}^n$ над полем вещественных чисел называется множество прямых евклидова пространства $\mathbb{R}^{n+1}$, проходящих через начало координат. Проективное пространство $\mathbb{RP}^n$ представляет собой многообразие размерности n . Его можно представлять себе как сферу $\mathbb{S}^n$ с отождествленными диаметрально противоположными точками. Действительно, любая прямая, проходящая через начало координат, пересекает единичную сферу с центром в начале координат ровно в двух диаметрально противоположных точках. Обратно, любая из этих двух точек однозначно определяет прямую, проходящую через начало координат. Таким образом

$$\mathbb{RP}^n \approx \frac{\mathbb{S}^n}{\mathbb{Z}_2},$$

где циклическая группа $\mathbb{Z}_2$ состоит из двух элементов $\{1, -1\}$. Проективное пространство $\mathbb{RP}^n$ можно представить также в виде полусферы $x \in \mathbb{S}^n, x^n \geq 0$, у которой отождествлены диаметрально противоположные краевые точки, т.е. точки $(n-1)$ -мерной сферы

$$\mathbb{S}^{n-1} = \{x \in \mathbb{S}^n : x^n = 0\}.$$

Рассмотрим замкнутую кривую в проективном пространстве, проходящую через одну из краевых точек. На рис. 2.2, *a*, для наглядности изображена проективная плоскость $\mathbb{RP}^2$ в трехмерном евклидовом пространстве и возможная кривая. Выберем ортонормированный базис вдоль кривой, включающий единичный касательный вектор e_1 . При прохождении через краевую точку p касательный вектор e_1 не меняет ориентации относительно кривой, в то время как все остальные базисные векторы $e_2, e_3, \dots$ меняют направление. Это значит, что ориентация ортонормированного базиса при прохождении вдоль этой замкнутой кривой изменится при четных n и сохранится при нечетных n . Если замкнутая кривая целиком лежит в верхней полусфере, то ориентация базиса вдоль кривой сохраняется. Тем самым мы показали, что проективные пространства четного числа измерений неориентируемы, а нечетного числа измерений – ориентируемы. $\square$

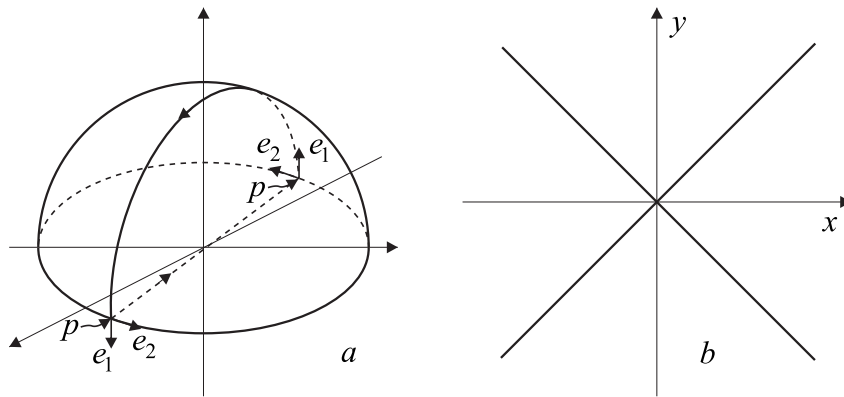


Рис. 2.2: Проективная плоскость $\mathbb{RP}^2$ как полусфера в $\mathbb{R}^3$. Показан замкнутый путь, проходящий через граничную точку p , и перенос базиса вдоль пути (*a*). Две пересекающиеся прямые не являются многообразием (*b*).

Пример 2.1.9. Продемонстрируем отличие топологического пространства от многообразия. Пусть множество точек на евклидовой плоскости состоит из двух пересекающихся прямых $y^2 - x^2 = 0$, изображенных на рис. 2.2, *b*. Пусть топология на этих прямых индуцирована вложением. Тогда это множество представляет собой связное хаусдорфово топологическое пространство. В то же время оно не является многообразием, потому что окрестность точки пересечения прямых нельзя взаимно однозначно отобразить на интервал вещественной прямой $\mathbb{R}$. $\square$

Пример 2.1.10. Рассмотрим прямое произведение двух прямых $\mathbb{M} = \mathbb{R} \times \mathbb{R}_D$. Будем считать, что на первом сомножителе задана естественная топология, а на втором – дискретная, что отмечено индексом D . Тогда, как множество, многообразие $\mathbb{M}$ совпадает с двумерным евклидовым пространством $\mathbb{R}^2$, а как многообразие – нет. База топологии $\mathbb{M}$ состоит из всех интервалов на всех прямых, параллельных оси x и проходящих через все точки оси y (см. рис. 2.3). На каждой прямой база топологии счетна. Рассмотрим открытый диск $\mathbb{D} \subset \mathbb{M}$. Так же как и в евклидовом пространстве $\mathbb{R}^2$,

он является открытым множеством, но на этот раз как объединение несчетного числа интервалов. Рассмотрим отображение f многообразия $\mathbb{M}$ на евклидову плоскость $\mathbb{R}^2$, которое задается простым отождествлением координат: $\mathbb{M} \ni (x, y) \xrightarrow{f} (x, y) \in \mathbb{R}^2$. Это отображение биективно и непрерывно. Однако обратное отображение f^{-1} не является непрерывным, т.к. на евклидовой плоскости не существует открытого множества $U \subset \mathbb{R}^2$, образ которого при обратном отображении $f^{-1}(U)$ лежал бы в интервале. Поэтому отображение f не является гомеоморфизмом и множество $\mathbb{M} = \mathbb{R} \times \mathbb{R}_D$ не является двумерным многообразием. По построению, множество $\mathbb{M}$ является одномерным несвязным многообразием, состоящим из несчетного числа одномерных многообразий – прямых $\mathbb{R}$.

Рассмотренный пример показывает, что на одном и том же множестве точек можно задавать различные структуры. Мы говорим, что на евклидовой плоскости $\mathbb{R}^2$, рассматриваемой, как двумерное многообразие, задана структура слоения, т.е. мы представляем плоскость в виде объединения несчетного числа одномерных подмногообразий – прямых $\mathbb{R}$. Каждая прямая является листом слоения, которые параметризуются точками другой прямой $\mathbb{R}_D$. Более подробно слоения рассмотрены в разделе 2.12. $\square$

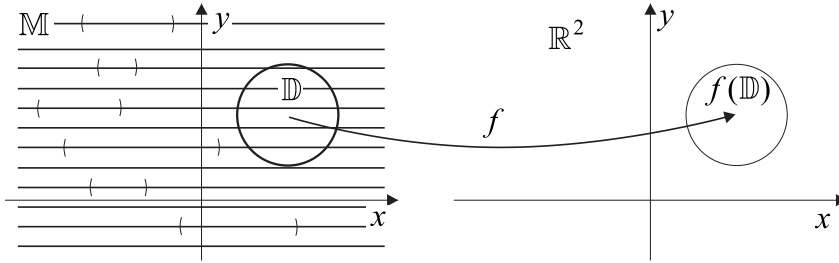


Рис. 2.3: Отображение f объединения $\mathbb{M}$ несчетного числа прямых, параллельных оси x , на евклидову плоскость $\mathbb{R}^2$.

В дальнейшем мы будем изучать различные свойства многообразий. Условимся о терминологии. Будем говорить, что данное свойство выполняется на многообразии $\mathbb{M}$ *глобально*, если оно выполнено во всех точках $x \in \mathbb{M}$. Гораздо чаще встречаются свойства, которые выполнены только *локально*. А именно, для каждой точки $x \in \mathbb{M}$ существует координатная окрестность U_x такая, что данное свойство выполнено на U_x . В этом случае можно говорить, что данное свойство выполнено в фиксированной системе координат. Конечно, любое свойство, выполненное глобально, справедливо и локально, но не наоборот.

2.2 Разбиение единицы

Один из способов задания геометрических структур, например, метрики, на многообразии $\mathbb{M}$ заключается в следующем. Сначала выбирается некоторый атлас, покрывающий $\mathbb{M}$. Затем в каждой карте данного атласа в координатах задается некоторая геометрическая структура. Чтобы задать данную геометрическую структуру на всем многообразии $\mathbb{M}$, ее необходимо склеить в областях пересечения карт. Для этого используется разбиение единицы.

Как уже отмечалось, для существования разбиения единицы нам достаточно рассматривать паракомпактные многообразия. Счетность базы многообразия является достаточным условием паракомпактности. Напомним некоторые определения и

утверждения из общей топологии, которые понадобятся для формулировки теоремы о существовании разбиения единицы.

Определение. Топологическое пространство M является *счетным в бесконечности*, если существует счетное семейство компактных множеств U_i , $i \in \mathbb{N}$, таких, что

$$U_1 \subset \text{int } U_2 \subset U_2 \subset \dots \subset U_i \subset \text{int } U_{i+1} \subset U_{i+1} \dots$$

и

$$M = \bigcup_{i=1}^{\infty} U_i.$$

□

Теорема 2.2.1. *Паракомпактное топологическое пространство является объединением семейства связных паракомпактных топологических пространств, которые являются счетными в бесконечности.*

Доказательство. См., например, [41], глава 1, теорема 1.11. □

Теорема 2.2.2. *Любое покрытие паракомпактного многообразия имеет счетное локально конечное измельчение.*

Доказательство. Существование локально конечного измельчения, которое также является покрытием, входит в определение паракомпактности и его доказывать не надо. Нетривиальность утверждения теоремы в том, что локально конечное покрытие можно выбрать счетным. Доказательство приведено в [42], глава I, лемма 1.9. □

Перед тем как ввести разбиение единицы напомним

Определение. *Носителем* функции f , заданной на многообразии M , называется замыкание множества тех точек, в которых она отлична от нуля. Носитель функции обозначается $\text{supp } f \subset M$. Если носитель функции компактен, то функция называется *финитной*. □

Из определения следует, что носитель произвольной функции f всегда замкнут в M .

Определение. *Разбиением единицы*, подчиненным заданному покрытию $\{U_i\}_{i \in I}$ многообразия M , называется семейство функций $\{f_i\}$, удовлетворяющих следующим условиям:

- 1) $\text{supp } f_i \subset U_i$ для всех i ;
- 2) каждая точка имеет окрестность W такую, что $W \cap \text{supp } f_i = \emptyset$ за исключением конечного числа функций f_i ;
- 3) $0 \leq f_i(x) \leq 1$, $\sum_i f_i(x) = 1$, $\forall x \in M$. □

Замечание. В данном определении покрытие произвольно и не обязательно счетное и локально конечное. Сумма в условии 3) определена, т.к. содержит только конечное число слагаемых в силу условия 2). □

Теорема 2.2.3 (Разбиение единицы). *На любом паракомпактном многообразии класса C^k существует C^k разбиение единицы, подчиненное заданному счетному локально конечному покрытию.*

Доказательство. См., например, [43], глава I, §2, теорема 4. □

Замечание. Разбиение единицы существует как для ориентируемых, так и для неориентируемых многообразий. $\square$

Для заданного покрытия $\{\mathbb{U}_i\}$ существует много разбиений единицы. Пусть $\{f_i\}$ и $\{g_i\}$ – два разбиения единицы, подчиненные одному покрытию $\{\mathbb{U}_i\}$. Тогда очевидна формула

$$\sum_i f_i \sum_j g_j = \sum_{ij} f_i g_j = \sum_j g_j \sum_i f_i,$$

поскольку в каждой точке $x \in \mathbb{M}$ суммы содержат только конечное число слагаемых.

Существование разбиения единицы на многообразии является эффективным средством доказательства существования геометрических структур на многообразии путем склеивания этих структур, заданных в отдельных картах. Например, в теореме 4.1.1 доказано существование римановой метрики на произвольном многообразии.

Замечание. Условие паракомпактности, являющееся достаточным условием существования разбиения единицы можно заменить на вторую аксиому счетности, утверждающую существование счетной базы топологии. Если потребовать, чтобы многообразие $\mathbb{M}$ являлось хаусдорфовым топологическим пространством, удовлетворяющим второй аксиоме счетности, то отсюда будет следовать паракомпактность многообразий и, кроме того, нормальность и метризуемость. (Доказательство метризуемости топологических пространств, удовлетворяющих второй аксиоме счетности, приведено, например, в [9].) $\square$

2.3 Многообразие с краем

Определение. Пусть $f(x) \in C^\infty(M)$ – гладкая функция на гладком многообразии $\mathbb{M}$. Тогда замкнутое множество $\mathbb{A} \subset \mathbb{M}$, выделяемое в многообразии $\mathbb{M}$ неравенством $f(x) \leq 0$ (или $f(x) \geq 0$) называется *гладким многообразием с краем*. Подмногообразие $\partial\mathbb{A} \subset \mathbb{M}$, задаваемое уравнением $f(x) = 0$, называется *краем* $\mathbb{A}$. При этом мы предполагаем, что градиент функции f на крае $\partial\mathbb{A}$ отличен от нуля. $\square$

Если функция положительна, $f > 0$, на $\mathbb{M}$, то край пустой, $\partial\mathbb{A} = \emptyset$, и $\mathbb{A} = \mathbb{M}$. Нетривиальное многообразие с краем получается, если область значений функции f включает нуль.

Пример 2.3.1. Замыкание любой ограниченной открытой области $\mathbb{U} \subset \mathbb{R}^n$ в евклидовом пространстве $\mathbb{R}^n$ является многообразием с краем $\bar{\mathbb{U}}$. При этом краем является граница области: $\partial\bar{\mathbb{U}} = \bar{\mathbb{U}} \setminus \mathbb{U}$. $\square$

Замечание. Условие отличия градиента функции от нуля на крае является достаточным условием того, что край является $(n-1)$ -мерным подмногообразием в $\mathbb{M}$. По сути дела это определение является инвариантным обобщением понятия области и ее границы в евклидовом пространстве $\mathbb{R}^n$. Если многообразие с краем можно покрыть одной картой, то оба понятия в точности совпадают. $\square$

Пример 2.3.2. Ориентируемым многообразием с краем является цилиндр конечной высоты, который получается при склейке двух краев прямоугольника, показанного на рис. 2.4а. При этом направление склеиваемых сторон, которое показано стрелками, сохраняется. Цилиндр можно покрыть двумя картами, которые пересекаются по двум областям, и в обеих областях якобиан перехода положителен. Краем цилиндра является объединение двух окружностей. $\square$

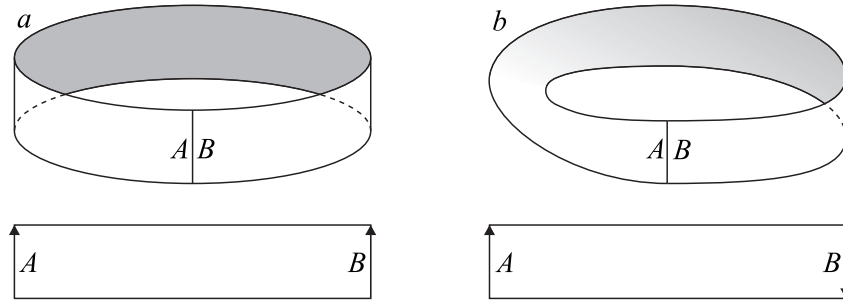


Рис. 2.4: Цилиндр (a) и лента Мебиуса (b) получаются при склеивании двух сторон прямоугольника с сохранением и изменением направления сторон при склейке, как показано стрелками.

Пример 2.3.3. Если перед склейкой прямоугольника изменить направление одной из сторон, как показано стрелками на рис. 2.4b, то получится неориентируемая поверхность с краем, которая называется *листом Мебиуса*. Лист Мебиуса также можно покрыть двумя картами, пересекающимися по двум областям, но в этом случае якобиан перехода в этих областях будет иметь разный знак. У листа Мебиуса край диффеоморфен одной окружности и является связным. Окружность, возникающая при склеивании середин отрезков A и B , называется центральной. $\square$

Если $\mathbb{A}$ – многообразие с краем, то разность $\mathbb{A} \setminus \partial\mathbb{A}$ также является многообразием, но уже без края.

Дадим эквивалентное определение многообразия с краем. Обозначим полупространство евклидова пространства $\mathbb{R}^n$, определяемое уравнением $x^n > 0$, через $\mathbb{R}_+^n$. Будем считать, что на замыкании $\overline{\mathbb{R}_+^n} \subset \mathbb{R}^n$ задана индуцированная топология. При этом открытые множества в $\overline{\mathbb{R}_+^n}$ могут как содержать, так и не содержать точки края, определяемого уравнением $x^n = 0$.

Определение. Топологическое хаусдорфово пространство $\mathbb{M}$ со счетной базой, каждая точка которого имеет окрестность, гомеоморфную открытому множеству в $\mathbb{R}_+^n$, называется n -мерным *топологическим многообразием с краем*. Точки, которые имеют окрестность, гомеоморфную $\mathbb{R}^n$, называются *внутренними*. Остальные точки называются *краевыми*. $\square$

Дифференцируемая структура на многообразии с краем вводится так же как и на многообразии без края. Если не оговорено противное, то мы рассматриваем гладкие дифференцируемые структуры класса C^∞ .

Теорема 2.3.1. Пусть $\mathbb{M}$, $\dim \mathbb{M} = n$, – многообразие с краем. Если край $\partial\mathbb{M}$ не является пустым, то $\partial\mathbb{M}$ представляет собой многообразие размерности $n - 1$ и без края $\partial(\partial\mathbb{M}) = \emptyset$.

Доказательство. См., например, [44], глава III, утверждение 1.3. $\square$

Определение. Многообразие с краем $\mathbb{M}$ называется *ориентируемым*, если ориентируемо соответствующее ему многообразие без края $\mathbb{M} \setminus \partial\mathbb{M}$. $\square$

Теорема 2.3.2. Если многообразие с краем $\mathbb{M}$ ориентируемо, то его край $\partial\mathbb{M}$ также является ориентируемым многообразием. Ориентация на $\mathbb{M}$ индуцирует каноническую ориентацию края $\partial\mathbb{M}$.

Доказательство. Пусть на $M \setminus \partial M$ задана какая-либо ориентация, которую назовем положительной. Рассмотрим естественное вложение края $h : \partial M \rightarrow M$. Тогда дифференциал отображения действует на касательные векторы $h_* : T_x(\partial M) \rightarrow T_x(M)$. Для каждой точки края $x \in \partial M$ выберем первый базисный вектор $e_1 \in T_x(M)$, $e_1 \notin T_x(\partial M)$ таким образом, чтобы он был ориентирован наружу. Это значит, что для любой дифференцируемой функции, удовлетворяющей условиям $f(\partial M) = 0$ и $f(M \setminus \partial M) \leq 0$, производная вдоль e_1 неотрицательна $e_1 f \geq 0$. Дополним этот вектор базисными векторами края $\{e_2, \dots, e_n\}$ таким образом, чтобы базис $\{e_1, \dots, e_n\}$ имел положительную ориентацию в M . Тогда ориентация $\{e_2, \dots, e_n\}$ задает *каноническую ориентацию* края. $\square$

На языке дифференциальных форм задание согласованной ориентации многообразия M и его края ∂M означает следующее. Пусть $\{e^a\}$, $a = 1, \dots, n$, – набор 1-форм, дуальных к базису e_a , построенному в доказательстве теоремы 2.3.2: $e^a(e_b) = \delta_b^a$. Поскольку 1-формы линейно независимы, то n -форма $e^1 \wedge e^2 \wedge \dots \wedge e^n$ отлична от нуля и задает ориентацию на M . Формы $e^2, \dots, e^n$, по построению, линейно независимы на крае ∂M и задают каноническую ориентацию края. В дальнейшем мы всегда предполагаем, что ориентация края ∂M индуцирована ориентацией самого многообразия M .

Из определения края следует выражение для края прямого произведения двух многообразий (правило Лейбница)

$$\partial(M_1 \times M_2) = (\partial M_1 \times M_2) \cup (M_1 \times \partial M_2).$$

Определение. В моделях гравитации принято называть вселенную *замкнутой*, если она представляет собой компактное многообразие без края. *Открытая вселенная* является некомпактным многообразием без края. В общем случае будем называть компактное многообразие без края *замкнутым*. Такие многообразия являются, конечно, замкнутыми множествами в топологическом смысле. Однако термин замкнутый в данном определении и в определении замкнутого множества в топологии имеют разный смысл. $\square$

Пример 2.3.4. Прямая представляет собой замкнутое (и одновременно открытое) множество в естественной топологии. В то же время она не является замкнутым многообразием, поскольку некомпактна. $\square$

2.4 Расслоения

В разделе 2.1 было определено прямое произведение $M \times F$ двух дифференцируемых многообразий M и F , которое также является дифференцируемым многообразием. При этом дифференцируемая структура на прямом произведении многообразий определяется дифференцируемыми структурами сомножителей. В настоящем разделе мы обобщим понятие прямого произведения многообразий.

Мы предполагаем, что все многообразия являются достаточно гладкими.

Определение. *Расслоением* называется четверка $E(M, \pi, F)$, где E, M, F – многообразия, а $\pi : E \rightarrow M$ – дифференцируемое отображение, которые удовлетворяют следующим условиям:

- 1) каждая точка $x \in M$ имеет окрестность U_x такую, что существует диффеоморфизм $\chi : \pi^{-1}(U_x) \rightarrow U_x \times F$ (локальная тривиальность);

2) композиция отображений $\pi \circ \chi^{-1} : \mathbb{U}_x \times \mathbb{F} \rightarrow \mathbb{M}$ есть проекция на первый сомножитель: $(y, v) \mapsto y$ для всех $y \in \mathbb{U}_x$ и $v \in \mathbb{F}$. Многообразие $\mathbb{E}$ называется *пространством расслоения*, $\mathbb{M}$ – *базой* расслоения, $\mathbb{F}$ – *типичным слоем* и π – *проекцией*. $\square$

Поскольку отображение χ в условии 1) является диффеоморфизмом, то

$$\dim \mathbb{E} = \dim \mathbb{M} + \dim \mathbb{F}.$$

Предполагается, что база многообразия является связным многообразием. В противном случае расслоения можно рассматривать над каждой связной компонентой. Расслоение также называется *расслоенным пространством* и обозначается $\mathbb{E} \xrightarrow{\pi} \mathbb{M}$. Мы считаем, что проекция π является дифференцируемым сюръективным отображением на базу $\mathbb{M}$. В противном случае можно рассмотреть расслоение с базой $\pi(\mathbb{E}) \subset \mathbb{M}$. Поскольку отображение π непрерывно, то согласно теореме 1.4.3 прообраз $\pi^{-1}(x)$ является замкнутым подмногообразием в $\mathbb{E}$, которое называется *слоем* над точкой базы $x \in \mathbb{M}$.

В дальнейшем, для краткости, мы иногда будем обозначать расслоение одной буквой $\mathbb{E}$. На рис. 2.5 схематично показана структура расслоения.

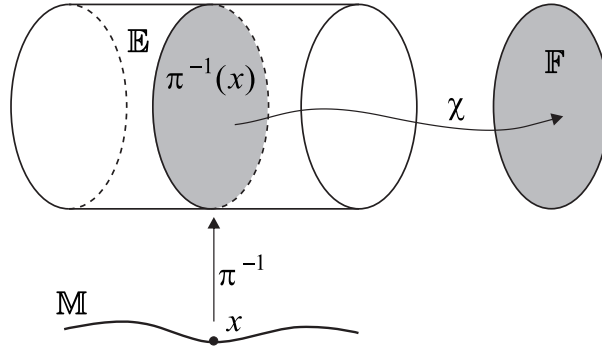


Рис. 2.5: Схематичное изображение структуры расслоения. $\mathbb{E}$ – пространство расслоения, $\mathbb{M}$ – база, $\mathbb{F}$ – типичный слой, $\pi : \mathbb{E} \rightarrow \mathbb{M}$ – проекция, $\chi : \pi^{-1}(x) \rightarrow \mathbb{F}$ – диффеоморфизм.

Замечание. Атласы, заданные на базе и типичном слое, определяют атлас на пространстве расслоения в силу локальной тривиальности расслоения. Тем самым первое условие в определении расслоения является достаточным для того, чтобы дифференцируемые структуры на трех многообразиях $\mathbb{E}$, $\mathbb{M}$ и $\mathbb{F}$ были согласованы между собой. Второе требование является условием коммутативности диаграммы

$$\begin{array}{ccc} \mathbb{E} \supset \pi^{-1}(\mathbb{U}_x) & \xrightarrow{\chi} & \mathbb{U}_x \times \mathbb{F} \\ & \searrow \pi & \downarrow \text{pr} = \pi \circ \chi^{-1} \\ & & \mathbb{U}_x \end{array}$$

где pr – естественная проекция прямого произведения $\mathbb{U}_x \times \mathbb{F} \xrightarrow{\text{pr}} \mathbb{U}_x$ на первый сомножитель. В общем случае это могло бы быть не так. В приведенной диаграмме окрестность $\mathbb{U}_x$ нельзя заменить на все многообразие $\mathbb{M}$, т.к. отображение χ определено локально. $\square$

Пример 2.4.1. Прямое произведение двух многообразий $\mathbb{E} = \mathbb{M} \times \mathbb{F}$ с проекцией на первый сомножитель $\mathbb{M} \times \mathbb{F} \xrightarrow{\pi} \mathbb{M}$ является расслоением с базой $\mathbb{M}$ и типичным слоем $\mathbb{F}$, которое называется *тривиальным*. С равным успехом четверка $\mathbb{M} \times \mathbb{F} \xrightarrow{\pi'} \mathbb{F}$ с проекцией на второй сомножитель является тривиальным расслоением с базой $\mathbb{F}$ и типичным слоем $\mathbb{M}$. $\square$

Многообразия $\mathbb{M}, \mathbb{F}, \mathbb{E}$ могут быть как с краем, так и без края. Мы допускаем также в качестве базы или типичного слоя 0-мерные многообразия, т.е. конечные или счетные наборы точек с дискретной топологией.

По построению, пространство расслоения $\mathbb{E}$ представляет собой объединение несчетного числа слоев, $\mathbb{E} = \bigcup_{x \in \mathbb{M}} \pi^{-1}(x)$, каждый из которых диффеоморфен типичному слою $\mathbb{F}$ и “нумеруется” точкой базы. Это – частный случай слоений, рассмотренных в разделе 2.12.

Как было отмечено, дифференцируемые структуры на базе и слое согласованы с дифференцируемой структурой на пространстве расслоения $\mathbb{E}$. Опишем это более подробно. Пусть $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$ – расслоение, $\dim \mathbb{M} = n$, $\dim \mathbb{F} = k$. По определению, отображение χ действует на каждую точку $p \in \mathbb{E}$: $\chi(p) = (x = \pi(p), v) \in \mathbb{M} \times \mathbb{F}$. Поскольку база и типичный слой – многообразия, то существуют карты:

$$\begin{aligned} (\mathbb{U}_x, \varphi) : \quad \mathbb{M} \supset \quad \mathbb{U}_x &\xrightarrow{\varphi} \varphi(\mathbb{U}_x) \subset \mathbb{R}^n, \\ (\mathbb{V}_v, \phi) : \quad \mathbb{F} \supset \quad \mathbb{V}_v &\xrightarrow{\phi} \phi(\mathbb{V}_v) \subset \mathbb{R}^k. \end{aligned}$$

Локально определено отображение из пространства расслоения $\mathbb{E}$ в евклидово пространство $\mathbb{R}^{n+k}$:

$$\rho : \quad \chi^{-1}(\mathbb{U}_x \times \mathbb{V}_v) \rightarrow \varphi(\mathbb{U}_x) \times \phi(\mathbb{V}_v) \in \mathbb{R}^{n+k}.$$

Таким образом, для каждой точки расслоения $p \in \mathbb{E}$ определена карта $(\chi^{-1}(\mathbb{U}_x \times \mathbb{V}_v), \rho)$. Определено также отображение, действующее из $\mathbb{R}^n \times \mathbb{R}^k$ в $\mathbb{R}^n$:

$$\tilde{\pi} = \varphi \circ \pi \circ \rho^{-1} : \quad \varphi(\mathbb{U}_x) \times \phi(\mathbb{V}_v) \ni (y, w) \mapsto y \in \varphi(\mathbb{U}_x),$$

где $y \in \varphi(\mathbb{U}_x) \subset \mathbb{R}^n$ и $w \in \phi(\mathbb{V}_v) \subset \mathbb{R}^k$.

Определение. Дифференцируемое отображение $\mathbb{M} \xrightarrow{\sigma} \mathbb{E}$ называется *сечением* или *глобальным сечением* расслоения $\mathbb{E} \xrightarrow{\pi} \mathbb{M}$, если $\pi \circ \sigma = \text{id}(\mathbb{M})$. Аналогичным образом, дифференцируемое отображение $\mathbb{U} \xrightarrow{\sigma} \mathbb{E}$, где $\mathbb{U}$ – открытое подмножество базы $\mathbb{M}$, называется *локальным сечением* расслоения $\mathbb{E} \xrightarrow{\pi} \mathbb{M}$, если $\pi \circ \sigma = \text{id}(\mathbb{U})$. $\square$

Локальные сечения существуют у любого расслоения – это функции со значениями в пространстве расслоения $\mathbb{E}$, которые определены в областях $\mathbb{U} \subset \mathbb{M}$. Глобальные сечения расслоений, как мы увидим в дальнейшем, существуют далеко не всегда.

Пример 2.4.2. Пусть $\mathbb{E} = \mathbb{M} \times \mathbb{R}$ – тривиальное расслоение, типичным слоем которого является поле вещественных чисел. Тогда множество всех гладких сечений

$$f : \quad \mathbb{M} \ni \quad x \mapsto (x, f(x)) \in \mathbb{E} = \mathbb{M} \times \mathbb{R}$$

совпадает с множеством графиков всех вещественнозначных бесконечно дифференцируемых функций $C^\infty(\mathbb{M})$. $\square$

Пример 2.4.3 (Цилиндр). Цилиндр единичной высоты – это прямое произведение окружности на единичный отрезок $\mathbb{S}^1 \times [0, 1]$. Его можно рассматривать как расслоение с базой $\mathbb{M} = \mathbb{S}^1$ и типичным слоем $[0, 1]$. У этого расслоения существуют глобальные сечения – гладкие функции на окружности со значениями в единичном отрезке $[0, 1]$. Например, $f(\varphi) = \sin \varphi$, где $\varphi \in [0, 2\pi]$ – координата на окружности. $\square$

Пример 2.4.4 (Лист Мебиуса). Лист Мебиуса, изображенный на рис.2.4, *b*, является расслоением с базой $\mathbb{S}^1$, типичным слоем которого, как и у цилиндра, является единичный отрезок $[0, 1]$. Это расслоение нетривиально, т.к. не имеет вида прямого произведения. $\square$

Пример 2.4.5 (Бутылка Клейна). Построение бутылки Клейна изображено на рис.2.6. Мы берем цилиндр конечной высоты и отождествляем точки граничных окружностей, предварительно отобразив точки окружности с одной стороны цилиндра относительно произвольного диаметра. Эту поверхность нельзя вложить в трехмерное евклидово пространство $\mathbb{R}^3$ и поэтому трудно представить. Бутылку Клейна можно рассматривать, как расслоение с базой $\mathbb{M} = \mathbb{S}^1$ и типичным слоем $\mathbb{F} = \mathbb{S}^1$. Это расслоение нетривиально. $\square$

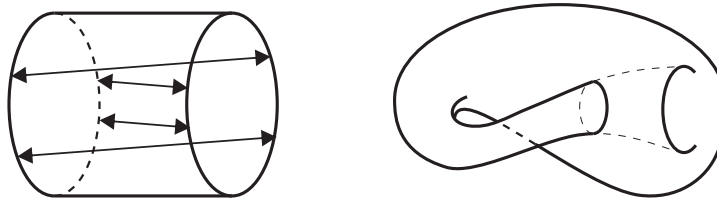


Рис. 2.6: Бутылка Клейна

Пример 2.4.6. Пусть задана группа Ли $\mathbb{G}$ и ее нормальная подгруппа $\mathbb{H} \subset \mathbb{G}$. Рассмотрим отображение группового многообразия $\mathbb{G}$ на пространство правых (или левых) смежных классов $\pi : \mathbb{G} \rightarrow \mathbb{G}/\mathbb{H}$, определенное правилом $\pi(g) = \mathbb{H}g$, где $g \in \mathbb{G}$. Тогда $\mathbb{G} \xrightarrow{\pi} \mathbb{G}/\mathbb{H}$ – расслоение. Пространством расслоения является группа Ли $\mathbb{G}$, базой – факторгруппа $\mathbb{G}/\mathbb{H}$ и типичным слоем – нормальная подгруппа $\mathbb{H}$. Дифференцируемая структура на факторпространстве $\mathbb{G}/\mathbb{H}$ будет определена позже в теореме 9.1.1. Если размерность группы равна размерности нормальной подгруппы, $\dim \mathbb{G} = \dim \mathbb{H}$, то базой является конечный или счетный набор точек, т.е. 0-мерное многообразие. $\square$

Замечание. В определении расслоения общего вида мы не предполагаем наличия каких-либо структур в типичном слое, кроме структуры дифференцируемого многообразия. В дальнейшем мы рассмотрим частные случаи расслоенных пространств, когда типичным слоем является векторное пространство (векторное расслоение) или группа Ли (главное расслоение). $\square$

2.5 Скалярные поля и плотности

В моделях математической физики, как правило, постулируется, что пространство и пространство-время, в котором мы живем, являются многообразиями. Само по себе это очень глубокое предположение. Однако для построения физических моделей

его недостаточно. Для описания движения и взаимодействия различных физических объектов в пространстве-времени необходимо задание дополнительных структур на многообразии. Такими структурами являются скалярные и векторные поля, r -формы, тензорные поля, метрика и связность, которые, в частности, характеризуются различными трансформационными свойствами при преобразовании координат. Начнем с простейшего объекта – скалярного поля (функции).

Рассмотрим вещественнозначную функцию f на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, т.е. отображение

$$f : \mathbb{M} \rightarrow \mathbb{R}. \quad (2.4)$$

Это отображение часто, особенно в физических приложениях, называют *скалярным полем* на $\mathbb{M}$. По определению отображения (1.65) скалярное поле должно быть однозначным. Функция называется *дифференцируемой* класса $\mathcal{C}^k$, если отображение (2.4), заданное в координатах,

$$f \circ \varphi_i^{-1} : \mathbb{R}^n \supset \varphi_i(\mathbb{U}_i) \ni x \mapsto f(x) \in \mathbb{R}$$

k раз непрерывно дифференцируемо в каждой карте атласа $\{\mathbb{U}_i, \varphi_i\}$. Конечно, не имеет смысла говорить о степени гладкости функции, которая превышает степень гладкости дифференцируемой структуры многообразия. Поэтому мы предполагаем, что степень гладкости функции меньше или равна степени гладкости многообразия.

В двух областях $\mathbb{U}_i$ и $\mathbb{U}_j$ скалярное поле задается, соответственно, двумя функциями $f(x)$ и $f'(x')$ от n переменных $x = \{x^\alpha\}$ и $x' = \{x^{\alpha'}\}$, $\alpha, \alpha' = 1, \dots, n$. Если области пересекаются, то в области пересечения согласно, (1.66), справедливо равенство

$$f'(x'(x)) = f(x), \quad (2.5)$$

поскольку в каждой точке функция имеет только одно значение. Формулу (2.5) можно интерпретировать, как правило преобразования функции при замене координат $x^\alpha \mapsto x^{\alpha'}(x)$. Другими словами, значение функции после преобразования в точке x' равно ее прежнему значению в точке x .

Замечание. В дальнейшем мы будем позволять себе некоторую вольность в обозначениях. Запись $f(x)$ в зависимости от контекста будет пониматься двояко. Во-первых, $f(x)$ обозначает значение функции f в произвольной точке многообразия $x \in \mathbb{M}$ безотносительно какой-либо карты. Это не есть функция n вещественных переменных. Во-вторых, запись $f(x)$ обозначает также значение функции $f \circ \varphi^{-1}$ в точке $\{x^\alpha\} \in \mathbb{R}^n$. Это – обычная функция от n вещественных переменных (координат). Строго говоря, частная производная от функции на многообразии $\partial_\alpha f$ не определена, т.к. мы не знаем, что такое разность двух точек $x_1 - x_2$ на многообразии. Тем не менее мы будем употреблять запись $\partial_\alpha f$, принимая

$$\partial_\alpha f := \partial_\alpha (f \circ \varphi^{-1}).$$

Фактически это означает, что в каждой отдельно взятой карте мы отождествляем точки многообразия с точками евклидова пространства: $x \sim \varphi(x) = \{x^\alpha\} \in \mathbb{R}^n$, и функции: $f \sim f \circ \varphi^{-1}$. При проведении вычислений в одной карте это не приводит к какой-либо путанице. $\square$

Множество всех функций класса $\mathcal{C}^k$ на многообразии $\mathbb{M}$ обозначим $\mathcal{C}^k(\mathbb{M})$. Множество гладких (бесконечно дифференцируемых) функций обозначим $\mathcal{C}^\infty(\mathbb{M})$. Скалярное поле называется *тривиальным*, если оно равно нулю на $\mathbb{M}$.

На множестве функций $\mathcal{C}^k(\mathbb{M})$ определим две поточечные операции: сложение и умножение:

$$(f + g)(x) := f(x) + g(x), \quad (fg)(x) := f(x)g(x), \quad \forall x \in \mathbb{M}.$$

То есть значения суммы и произведения двух функций в данной точке равно, соответственно, сумме и произведению значений этих функций в той же точке. Очевидно, что сумма и произведение двух функций снова дает функцию. По отношению к этим операциям функции образуют коммутативное кольцо. Кроме этого функции можно умножать на действительные числа. Умножение на числа вместе с операцией сложения превращает множество функций в векторное пространство над полем вещественных чисел. Если на множестве функций рассматривать все три операции (умножение на числа, сложение и умножение функций), то оно образует коммутативную ассоциативную алгебру с единицей над полем вещественных чисел, которую так же будем обозначать $\mathcal{C}^k(\mathbb{M})$. Эта алгебра является бесконечномерной.

В разделе 1.5 было показано, что координаты точки евклидова пространства сами можно рассматривать, как набор функций. Для каждой карты $(\mathbb{U}, \varphi)$ многообразия $\mathbb{M}$ определен набор функций $x^\alpha(x)$, $\alpha = 1, \dots, n$ от точки многообразия $x \in \mathbb{U}$,

$$\varphi(x) = \{x^\alpha(x)\} \in \mathbb{R}^n, \quad (2.6)$$

которые называются *координатными функциями*. Эти функции свои для каждой карты. Во многих случаях координатные функции, определенные на $\mathbb{U} \subset \mathbb{M}$, упрощают запись, позволяя опускать символ отображения $\varphi(x)$.

Определим новый геометрический объект – скалярную плотность h . С этой целью рассмотрим две пересекающиеся карты $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$. При преобразовании координат можно умножить функцию h_i , заданную на карте $\mathbb{U}_i$, на якобиан преобразования (1.69) в степени p :

$$h_j(x'(x)) := J_{ji}^p(x) h_i(x), \quad (2.7)$$

Пусть $\{\mathbb{U}_i, \varphi_i\}$ – некоторый атлас на многообразии $\mathbb{M}$. В каждой карте $(\mathbb{U}_i, \varphi_i)$ зададим отображение

$$h_i : \mathbb{R}^n \supset \varphi_i(\mathbb{U}_i) \ni x = \{x^\alpha\} \mapsto h_i(x) \in \mathbb{R}$$

таким образом, что в области пересечения карт $\mathbb{U}_i \cap \mathbb{U}_j$ выполнен закон преобразования (2.7) для всех $x \in \mathbb{U}_i \cap \mathbb{U}_j$. Это определение корректно, т.к. в области пересечения трех карт $\mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k$ с координатами $x = \{x^\alpha\}$, $x' = \{x^{\alpha'}\}$ и $x'' = \{x^{\alpha''}\}$ выполнено равенство

$$h_k(x''(x'(x))) = J_{kj}^p h_j(x'(x)) = J_{kj}^p J_{ji}^p h_i(x) = J_{ki}^p h_i(x),$$

где мы воспользовались равенством $J_{kj} J_{ji} = J_{ki}$ для якобианов преобразования координат $x \mapsto x'(x)$, $x' \mapsto x''(x')$ и $x \mapsto x''(x)$.

Преобразования (2.7) образуют группу. Действительно, якобианы преобразования координат, по определению, отличны от нуля, якобиан двух последовательных преобразований равен произведению якобианов, а якобиан обратного преобразования равен J^{-1} .

Определение. Геометрический объект $h = \{h_i\}$, заданный в некотором атласе $\{\mathbb{U}_i, \varphi_i\}$ многообразия $\mathbb{M}$ с правилом преобразования (2.7) в области пересечения любых двух карт называется скалярной *плотностью степени p* , и мы будем писать $\deg h = p$. $\square$

Строго говоря, скалярная плотность не является функцией в смысле определения (2.4), и мы не можем писать $h(x)$, где x – точка многообразия. Имеет смысл лишь запись $h(x^\alpha)$ для функции $h \circ \varphi^{-1}$, которая задана в координатном евклидовом пространстве. Несмотря на это, использование скалярных плотностей, например, при интегрировании, бывает удобным. Кроме того, производить вычисления с плотностями часто бывает проще, чем с тензорами, как, например, в общей теории относительности.

Забегая вперед, отметим, что поскольку определитель репера $\det e_\alpha^a = \sqrt{|g|} \neq 0$ является скалярной плотностью веса -1 , то произвольную скалярную плотность степени p можно представить в виде

$$h = \sqrt{|g|}^{-p} f,$$

где $f = \sqrt{|g|}^p h$ – скалярное поле (функция).

Замечание. Множество скалярных плотностей фиксированной степени алгебры не образует, т.к. произведение двух плотностей степеней p_1 и p_2 дает скалярную плотность степени $p_1 + p_2$. $\square$

2.6 Векторные поля и 1-формы

2.6.1 Локальное определение

Начнем с локального определения векторных полей и 1-форм, которое является более наглядным. Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$. Ограничим наше рассмотрение двумя пересекающимися картами $(\mathbb{U}_i, \varphi_i)$ и $(\mathbb{U}_j, \varphi_j)$ с координатами x^α и $x^{\alpha'}$, соответственно. В области пересечения этих карт (или при преобразовании координат $x^\alpha \rightarrow x^{\alpha'}(x)$) дифференциалы умножаются на матрицу Якоби, а частные производные – на ее обратную:

$$dx^{\alpha'} = dx^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha}, \quad (2.8)$$

$$\partial_{\alpha'} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} \partial_\alpha. \quad (2.9)$$

Матрицы преобразования дифференциалов и частных производных являются взаимно обратными по правилу дифференцирования сложных функций:

$$\frac{\partial x^\alpha}{\partial x^\beta} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} \frac{\partial x^{\alpha'}}{\partial x^\beta} = \delta_\beta^\alpha. \quad (2.10)$$

Векторные поля и 1-формы на многообразии определяются, исходя из правила преобразования дифференциалов и частных производных. А именно, n достаточно гладких функций $X^\alpha(x)$, заданных на карте $(\mathbb{U}_i, \varphi_i)$ и преобразующихся по правилу (2.8),

$$X^{\alpha'} := X^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha}, \quad (2.11)$$

при преобразовании координат, называются *компонентами векторного* или *контравариантного векторного* поля. Аналогично, n достаточно гладких функций $A_\alpha(x)$, преобразующихся по правилу (2.9),

$$A_{\alpha'} := \frac{\partial x^\alpha}{\partial x^{\alpha'}} A_\alpha. \quad (2.12)$$

называются *компонентами ковекторного*, или *ковариантного векторного* поля, или *1-формы*. 1-формы называются также *формами Пфаффа*. В общем случае (ко)векторное поле имеет n независимых компонент. Каждая из функций X^α или A_α является компонентой векторного или ковекторного поля относительно координатных базисов $e_\alpha := \partial_\alpha$ и $e^\alpha := dx^\alpha$. Смысл обозначения координатных базисов частными производными и дифференциалами не случаен и будет ясен из дальнейшего.

Для того, чтобы задать компоненты (ко-)векторного поля на всем многообразии, их необходимо задать в каждой карте некоторого атласа $\{\mathbb{U}_i, \varphi_i\}$ таким образом, чтобы во всех областях пересечения карт $\mathbb{U}_i \cap \mathbb{U}_j$ они были связаны преобразованием (2.11). Это определение непротиворечиво. Действительно, если точка x лежит в пересечении трех карт $(\mathbb{U}_i, \varphi_i)$, $(\mathbb{U}_j, \varphi_j)$ и $(\mathbb{U}_k, \varphi_k)$ с координатами x^α , $x^{\alpha'}$ и $x^{\alpha''}$ соответственно, то компоненты векторов преобразуются по правилу:

$$X^{\alpha''} = X^{\alpha'} \frac{\partial x^{\alpha''}}{\partial x^{\alpha'}} = X^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha} \frac{\partial x^{\alpha''}}{\partial x^{\alpha'}} = X^\alpha \frac{\partial x^{\alpha''}}{\partial x^\alpha}.$$

Данное равенство является следствием правила дифференцирования сложных функций и означает коммутативность следующей диаграммы

$$\begin{array}{ccc} \mathbb{U}_i & \xrightarrow{\varphi_j^{-1} \circ \varphi_i} & \mathbb{U}_j \\ & \searrow \varphi_k^{-1} \circ \varphi_i & \downarrow \varphi_k^{-1} \circ \varphi_j \\ & & \mathbb{U}_k \end{array}$$

для всех точек $x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k$.

Само векторное поле X на многообразии $\mathbb{M}$ является инвариантным геометрическим объектом и не зависит от выбора системы координат. Выше мы определили *компоненты* векторного поля на многообразии $\mathbb{M}$ путем их задания в некотором атласе. Далее мы должны показать, что таким образом определенный геометрический объект – векторное поле – не зависит от выбора атласа. Это также обеспечено правилом преобразования (2.11). Таким образом, для задания векторного поля X на многообразии $\mathbb{M}$ достаточно задать его компоненты в некотором атласе и указать правило преобразования (2.11) или (2.12). Аналогично дается глобальное определение ковекторного поля.

Пример 2.6.1. Если векторное поле имеет нулевые компоненты в одной системе координат, то они равны нулю и во всех других системах. Нулевое векторное поле, компоненты которого равны нулю во всех картах, называется *тривиальным*. Это единственное векторное поле, компоненты которого инвариантны относительно преобразований координат. Аналогично определяется нулевая 1-форма. $\square$

Пример 2.6.2. Дифференциалы dx^α , рассматриваемые как функции от точки $x \in \mathbb{U}$, являются компонентами гладкого векторного поля. Это векторное поле определено в произвольной карте, а в областях пересечения карт справедливо равенство (2.8). Сами координатные функции $x^\alpha(x)$, хотя и имеют векторный индекс, векторного поля не образуют. Это – набор скалярных полей. $\square$

Пример 2.6.3. Частные производные от произвольной функции $f \circ \varphi^{-1}$ на образе $\varphi(\mathbb{U}) \subset \mathbb{R}^n$ являются компонентами ковариантного векторного поля $\{\partial_\alpha f\}$, которое называется градиентом функции. Это ковекторное поле определено в произвольной карте и имеет правильный закон преобразования (2.12). $\square$

Пример 2.6.4. Примером векторного поля на кривой $\gamma = \{x^\alpha(t)\} \in \mathbb{M}$ является вектор скорости (1.9). Действительно, при преобразовании координат компоненты вектора скорости $\dot{x}^\alpha$ преобразуются, как дифференциалы. При этом вектор скорости рассматривается в точке кривой $\gamma \in \mathbb{M}$. В то же время сами координатные функции $x^\alpha(t)$ определены на отрезке $t \in [0, 1]$, а не на многообразии и векторного поля не образуют. Векторное поле скорости называется также касательным векторным полем к кривой γ . $\square$

Преобразования векторных полей (2.11) и 1-форм (2.12) различны, поэтому контравариантные и ковариантные индексы необходимо различать и они всегда будут писаться, соответственно, сверху и снизу.

Преобразования векторов (2.11) и 1-форм (2.12) линейны и однородны, причем матрица $\partial x^\alpha / \partial x^{\alpha'}$ в каждой точке многообразия невырождена и поэтому принадлежит группе невырожденных матриц $\mathrm{GL}(n, \mathbb{R})$. То есть каждому преобразованию координат соответствует $\mathrm{GL}(n, \mathbb{R})$ преобразование компонент векторного поля в касательном пространстве. Поскольку элементы группы $\mathrm{GL}(n, \mathbb{R})$ зависят от точки многообразия, то такие преобразования называются локальными. Обратное утверждение, вообще говоря, неверно. Не каждому локальному $\mathrm{GL}(n, \mathbb{R})$ преобразованию компонент векторного поля можно сопоставить некоторое преобразование координат. Это видно из подсчета параметров: локальное $\mathrm{GL}(n, \mathbb{R})$ преобразование параметризуется n^2 функциями по числу элементов $n \times n$ матрицы, в то время как преобразования координат параметризуются n функциями.

Из закона преобразования частных производных (2.9) и дифференциалов (2.8) следует, что суммы

$$X = X^\alpha \partial_\alpha = X^\alpha e_\alpha, \quad (2.13)$$

$$A = dx^\alpha A_\alpha = e^\alpha A_\alpha \quad (2.14)$$

инвариантны относительно преобразований координат. Эти формулы представляют собой разложения векторов и 1-форм по координатному базису, определенному далее в разделе 2.6.4.

Замечание. Скажем несколько слов об обозначениях. Там, где это возможно, мы будем записывать индексы суммирования по правилу “с десяти до четырех” (имеется в виду циферблат часов), т.е. сначала будем писать верхний индекс, а затем – нижний. Это правило является следствием записи *оператора внешнего дифференцирования* в виде

$$d = dx^\alpha \partial_\alpha. \quad (2.15)$$

В обратном порядке запись $d = \partial_\alpha dx^\alpha$ выглядит чрезвычайно неуклюже. В дифференциальной геометрии, когда все координаты являются вещественными числами, это правило не играет существенной роли, т.к. компоненты векторов и 1-форм можно менять местами. Однако, если часть координат антикоммутирует, то порядок индексов существенен, и необходимо придерживаться какого-либо фиксированного правила. Правило “с десяти до четырех” не является единственным, однако оно широко используется при построении моделей супергравитации в суперпространстве, где часть координат является антикоммутирующей. $\square$

Дифференцируемое векторное поле, заданное в какой-нибудь одной карте $\mathbb{U} \subset \mathbb{M}$, не всегда может быть продолжено до дифференцируемого векторного поля на всем многообразии.

Пример 2.6.5. Рассмотрим сферу $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ единичного радиуса, вложенную в евклидово пространство (рис.2.7):

$$x^2 + y^2 + z^2 = 1, \quad (x, y, z) \in \mathbb{R}^3.$$

Рассмотрим открытую верхнюю полусферу:

$$\mathbb{U} = \{(x, y, z) \in \mathbb{S}^2 : z > 0\}.$$

Зададим на $\mathbb{U}$ систему координат, спроектировав точки полусферы на плоскость x, y :

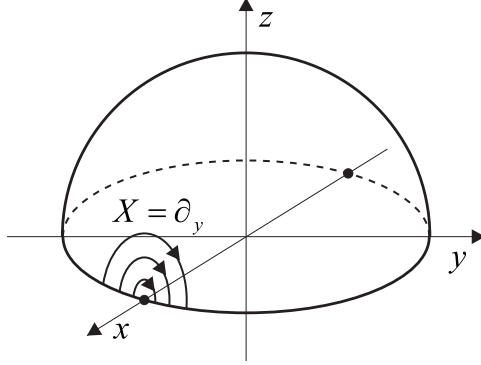


Рис. 2.7: Полусфера, вложенная в трехмерное евклидово пространство. Точки $(1, 0, 0)$ и $(-1, 0, 0)$ являются особыми для векторного поля $X := \partial_y$.

$(x, y, z) \mapsto (x, y)$. Таким образом мы построили карту $(\mathbb{U}, \varphi)$. Зададим в этой карте векторное поле $X = \partial_y$. В координатном базисе оно имеет компоненты $(0, 1)$ и поэтому является гладким векторным полем на $\mathbb{U}$. Точки сферы $(1, 0, 0)$ и $(-1, 0, 0)$ являются предельными для $\mathbb{U}$ и существенно особыми точками для векторного поля X , поскольку предел зависит от пути, по которому мы стремимся к данным точкам. Пока речь идет об открытом подмножестве $\mathbb{U}$ проблем не возникает, т.к. указанные точки не принадлежат $\mathbb{U}$. Однако любое продолжение векторного поля на окрестность, содержащую любую из точек $(1, 0, 0)$ и $(-1, 0, 0)$ приведет к векторному полю с особенностью. В данном случае особенность означает не обращение компонент векторного поля в бесконечность, а то, что в указанных точках векторное поле не будет однозначно определено. $\square$

2.6.2 Глобальное определение векторных и ковекторных полей

Дадим глобальное определение векторных полей и 1-форм на многообразии $\mathbb{M}$, как это обычно делается в современных курсах дифференциальной геометрии. Рассмотрим дифференцируемую кривую $\gamma_1 : [-1, 1] \ni t \rightarrow \mathbb{M}$, проходящую через некоторую точку $p \in \mathbb{M}$. Пусть $(\mathbb{U}, \varphi)$ – карта, содержащая точку $p \in \mathbb{U} \subset \mathbb{M}$. Тогда в координатах кривая задается набором функций:

$$\varphi \circ \gamma_1 = \{x_1^\alpha(t)\}, \quad p = \gamma_1(0).$$

Для определенности мы выбрали такую параметризацию кривой, что точке p соответствует значение $t = 0$. Рассмотрим другую дифференцируемую кривую γ_2 , также проходящую через точку p ,

$$\varphi \circ \gamma_2 = \{x_2^\alpha(t)\}, \quad p = \gamma_2(0).$$

Мы говорим, что две кривые касаются друг друга в точке p , если векторы скорости кривых в этой точке совпадают:

$$\dot{x}_1^\alpha|_{t=0} = \dot{x}_2^\alpha|_{t=0}, \quad \alpha = 1, \dots, n.$$

Поскольку векторы скорости при преобразованиях координат преобразуются одинаково, то данное определение не зависит от карты, покрывающей точку $p \in \mathbb{M}$. Касание кривых в точке p является отношением эквивалентности в классе всех кривых, проходящих через эту точку. Обозначим класс эквивалентности кривых, проходящих через точку p , который соответствует некоторому представителю γ_p , квадратными скобками $[\gamma_p]$. Каждый класс эквивалентности в координатах взаимно однозначно характеризуется набором чисел $\{\dot{x}^\alpha(0)\}$.

Определение. Касательным вектором $X_{(\gamma)}(p)$ к многообразию $\mathbb{M}$ в точке $p \in \mathbb{M}$ называется класс эквивалентности кривых $[\gamma_p]$, проходящих через эту точку. Множество всех касательных векторов в точке p называется *касательным пространством* к многообразию в точке p и обозначается $\mathbb{T}_p(\mathbb{M})$. Объединение всех касательных пространств

$$\mathbb{T}(\mathbb{M}) := \bigcup_{p \in \mathbb{M}} \mathbb{T}_p(\mathbb{M}). \quad (2.16)$$

называется *касательным расслоением* с базой $\mathbb{M}$ и естественной проекцией $\pi : \mathbb{T}(\mathbb{M}) \rightarrow \mathbb{M}$, которая задана отображением $(p, [\gamma_p]) \mapsto p$. Слоем касательного расслоения в точке p является касательное пространство $\pi^{-1}(p) = \mathbb{T}_p(\mathbb{M})$. Векторным полем $X(x)$ на многообразии $\mathbb{M}$ называется сечение касательного расслоения $\mathbb{T}(\mathbb{M})$, т.е. отображение

$$X : \quad \mathbb{M} \ni \quad x \mapsto X(x) \quad \in \mathbb{T}(\mathbb{M})$$

такое, что $\pi(X(x)) = x$. □

То, что каждый слой является многообразием и диффеоморфен n -мерному векторному пространству, мы покажем чуть ниже.

Пусть f – функция на многообразии $\mathbb{M}$. Тогда каждому вектору $X_{(\gamma)}$ в точке p мы ставим в соответствие производную функции вдоль вектора

$$X_{(\gamma)}f|_p = \left. \frac{df(\gamma(t))}{dt} \right|_{t=0} = X^\alpha(p) \partial_\alpha (f \circ \varphi^{-1})|_{t=0}, \quad (2.17)$$

где $X^\alpha(p) := \dot{x}^\alpha(0)$ – компоненты вектора $X(p)$ в точке p в некоторой карте. Поскольку векторы скорости для всех кривых из одного класса эквивалентности совпадают, то вектор $X(p)$ в точке p взаимно однозначно характеризуется своими компонентами $\{X^\alpha(p)\}$. Отсюда следует, что векторное поле в произвольной карте $(\mathbb{U}, \varphi)$ взаимно однозначно задается своими компонентами $X(x) = \{X^\alpha(x)\}$, $x \in \mathbb{U}$. При этом компоненты векторного поля при переходе от одной системы координат к другой преобразуются по правилу (2.11), что является следствием правила дифференцирования сложных функций. Таким образом, локальное определение векторного поля, данное ранее, вытекает из глобального определения настоящего раздела.

В каждой точке многообразия $p \in \mathbb{M}$ множество векторов $\mathbb{T}_p(\mathbb{M})$ обладает естественной структурой векторного пространства.

Теорема 2.6.1. Касательное пространство $\mathbb{T}_p(\mathbb{M})$ имеет естественную структуру вещественного векторного пространства той же размерности, что и само многообразие, $\dim \mathbb{T}_p(\mathbb{M}) = \dim \mathbb{M} = n$.

Доказательство. Рассмотрим два вектора $X_1 = [\gamma_1] \in \mathbb{T}_p(\mathbb{M})$ и $X_2 = [\gamma_2] \in \mathbb{T}_p(\mathbb{M})$ в произвольной точке $p \in \mathbb{M}$. Пусть $(\mathbb{U}, \varphi)$ – координатная окрестность точки p . Пусть $\gamma_1 \in [\gamma_1]$ и $\gamma_2 \in [\gamma_2]$ – две произвольные кривые из классов эквивалентности $[\gamma_1]$ и $[\gamma_2]$. Тогда в евклидовом пространстве определены две кривые $\varphi \circ \gamma_1$ и $\varphi \circ \gamma_2$, которые задаются набором функций $\varphi \circ \gamma_1 = \{x_1^\alpha(t)\}$ и $\varphi \circ \gamma_2 = \{x_2^\alpha(t)\}$. Поскольку координаты являются вещественными числами, то их можно складывать и умножать. Определим сумму двух векторов и умножение на число $a \in \mathbb{R}$ как следующие классы эквивалентности

$$\begin{aligned} X_1 + X_2 &:= [\varphi^{-1} \circ (\varphi \circ \gamma_1 + \varphi \circ \gamma_2)], \\ aX &:= [\varphi^{-1} \circ (a\varphi \circ \gamma)]. \end{aligned}$$

Касательные векторы к кривым $\varphi \circ \gamma_1 + \varphi \circ \gamma_2$ и $a\varphi \circ \gamma$ в точке p имеют, соответственно, компоненты: $\{\dot{x}_1^\alpha(0) + \dot{x}_2^\alpha(0)\}$ и $\{a\dot{x}^\alpha(0)\}$. Это определение суммы векторов и умножения на числа не зависит от выбора карты $(\mathbb{U}, \varphi)$ и представителя $\gamma \in [\gamma]$. Тем самым касательное пространство $\mathbb{T}_p(\mathbb{M})$ снабжается структурой векторного пространства. Поскольку каждый вектор $X(p)$ в карте взаимно однозначно задается набором n чисел $\{X^\alpha(p)\}$, то размерность касательного пространства совпадает с размерностью самого многообразия, $\dim \mathbb{T}_p(\mathbb{M}) = \dim \mathbb{M} = n$. $\square$

Таким образом, мы установили, что касательное пространство $\mathbb{T}_p(\mathbb{M})$ в каждой точке многообразия имеет естественную структуру векторного пространства $\mathbb{R}^n$. Определим в этом векторном пространстве евклидову топологию, которая является единственной топологией, согласованной с линейной структурой (см., например, [6], утверждение 4.7, следствие 1). Тем самым касательное пространство $\mathbb{T}_p(\mathbb{M})$ так же является многообразием. В дальнейшем мы всегда предполагаем, что касательное пространство $\mathbb{T}_p(\mathbb{M})$ снабжено естественной структурой векторного пространства $\mathbb{R}^n$ и евклидовой топологией.

Типичным слоем касательного расслоения $\mathbb{T}(\mathbb{M}) \xrightarrow{\pi} \mathbb{M}$ является евклидово пространство $\mathbb{R}^n$, на котором введена структура векторного пространства. Это частный случай векторных расслоений, рассмотренных далее в разделе 5.1.

Касательное расслоение является многообразием размерности $2n$. При этом, если на базе $\mathbb{M}$ задана дифференцируемая структура класса $\mathcal{C}^k$, $k \geq 1$, то касательное расслоение $\mathbb{T}(\mathbb{M})$ является дифференцируемым многообразием класса $\mathcal{C}^{k-1}$. Понижение класса дифференцируемости связано с тем, что, если функции преобразования координат принадлежат классу $\mathcal{C}^k$, то матрица Якоби преобразования координат, которая действует в касательном пространстве, принадлежит классу $\mathcal{C}^{k-1}$.

Пусть $\mathbb{M}$ – гладкое многообразие. Векторное поле $X(x)$ называется гладким, если в любой карте атласа $\{\mathbb{U}_i, \varphi_i\}$ компоненты $X^\alpha(x)$, задающие векторное поле, являются гладкими функциями. Множество всех гладких векторных полей на многообразии $\mathbb{M}$ обозначим $\mathcal{X}(\mathbb{M})$. Это множество, так же как и множество всех векторов в фиксированной точке $\mathbb{T}_p(\mathbb{M})$, обладает структурой вещественного линейного пространства с поточечным определением сложения и умножения на числа. Более того, вместо умножения на числа можно рассматривать умножение на гладкие функции $f \in \mathcal{C}^\infty(\mathbb{M})$. Легко проверить, что, если $X(x)$ – векторное поле, то $fX(x)$ также является векторным полем. Таким образом, множество векторных полей $\mathcal{X}(\mathbb{M})$ является модулем над алгеброй гладких функций $\mathcal{C}^\infty(\mathbb{M})$. Как линейное пространство множество векторных полей является бесконечномерным.

Теперь нетрудно дать глобальное определение ковариантных векторных полей.

Определение. Множество линейных функционалов на касательном пространстве $T_p(\mathbb{M})$ в точке $p \in \mathbb{M}$ называется *кокасательным векторным пространством* и обозначается $T_p^*(\mathbb{M})$. Объединение всех кокасательных пространств

$$T^*(\mathbb{M}) := \bigcup_{p \in \mathbb{M}} T_p^*(\mathbb{M}). \quad (2.18)$$

называется *кокасательным расслоением* с базой $\mathbb{M}$ и естественной проекцией $\pi : T^*(\mathbb{M}) \rightarrow \mathbb{M}$. Слоем кокасательного расслоения в точке p является кокасательное пространство $\pi^{-1}(p) = T_p^*(\mathbb{M})$. *Кокасательным векторным полем* или *1-формой* $A(x)$ на многообразии $\mathbb{M}$ называется сечение кокасательного расслоения $T^*(\mathbb{M})$

$$A : \quad \mathbb{M} \ni \quad x \mapsto A(x) \quad \in T^*(\mathbb{M}), \quad \pi(A(x)) = x,$$

или линейное отображение множества векторных полей

$$A : \quad \mathcal{X}(\mathbb{M}) \ni \quad X \mapsto A(X) \quad \in \mathcal{C}^\infty(\mathbb{M}). \quad \square$$

В координатах ковекторное поле задается набором n функций $A_\alpha(x)$, которые при преобразовании координат преобразуются по правилу (2.12). Тогда линейное отображение задается простым суммированием компонент:

$$A(X) = X^\alpha A_\alpha.$$

Кокасательное векторное поле $A(x)$ называется *гладким*, если его компоненты $\{A_\alpha(x)\}$ являются гладкими функциями во всех картах. Кокасательное пространство $T_p^*(\mathbb{M})$ в точке $p \in \mathbb{M}$ снабжается естественной структурой векторного пространства $\mathbb{R}^n$ и евклидовой топологией. Множество всех кокасательных векторных полей, которое обозначим $\Lambda_1(\mathbb{M})$, так же как и множество векторных полей $\mathcal{X}(\mathbb{M})$, образует модуль над алгеброй гладких функций $\mathcal{C}^\infty(\mathbb{M})$.

В дальнейшем нам понадобится

Предложение 2.6.1. *Касательное пространство к прямому произведению двух многообразий $\mathbb{M} \times \mathbb{N}$ в точке $(p, q) \in \mathbb{M} \times \mathbb{N}$ естественно изоморфно прямой сумме касательных пространств:*

$$T_{(p,q)}(\mathbb{M} \times \mathbb{N}) \ni \quad X \simeq (\pi_{1*}(X), \pi_{2*}(X)) \quad \in T_p(\mathbb{M}) \oplus T_q(\mathbb{N}),$$

где введены проекции на первый и второй сомножитель

$$\begin{aligned} \pi_1 : \quad \mathbb{M} \times \mathbb{N} \ni \quad (p, q) &\mapsto p \quad \in \mathbb{M}, \\ \pi_2 : \quad \mathbb{M} \times \mathbb{N} \ni \quad (p, q) &\mapsto q \quad \in \mathbb{N}, \end{aligned}$$

и π_{1*} , π_{2*} – дифференциалы соответствующих отображений (см. следующий раздел).

Доказательство. Достаточно спроектировать кривую в прямом произведении $\mathbb{M} \times \mathbb{N}$ на каждый из сомножителей. $\square$

2.6.3 Кокасательные векторные поля и ростки

Дадим также независимое определение кокасательного пространства 1-форм $T_p^*(\mathbb{M})$ в точке $p \in \mathbb{M}$, без обращения к понятию дуального пространства как это было сделано в предыдущем разделе. Рассмотрим алгебру гладких функций $C^\infty(\mathbb{M})$. Зафиксируем точку $p \in \mathbb{M}$. Будем считать две функции $f, g \in C^\infty(\mathbb{M})$ эквивалентными, $f \sim g$, если существует окрестность $\mathbb{U} \ni x$ такая, что $f|_{\mathbb{U}} = g|_{\mathbb{U}}$. Очевидно, что отношение $\sim$ является отношением эквивалентности в алгебре функций $C^\infty(\mathbb{M})$. Класс эквивалентности функции f обозначается $[f_p]$ и называется C^∞ *ростком* в точке $p \in \mathbb{M}$. Обозначим множество всех ростков в точке p через

$$\mathcal{C}_p := C^\infty(\mathbb{M}) / \sim = \{[f_p] : f \in C^\infty(\mathbb{M})\}. \quad (2.19)$$

Это множество естественным образом снабжается структурой векторного пространства, которая переносится из алгебры функций,

$$[f_p] + [g_p] := [f_p + g_p], \quad a[f_p] := [af_p], \quad a \in \mathbb{R}.$$

Таким образом, множество ростков $\mathcal{C}_p$ в точке $p \in \mathbb{M}$ превращается в бесконечномерное вещественное векторное пространство.

Обозначим множество гладких кривых γ , которые задаются набором функций $\{x^\alpha(t)\}$, проходящих через точку $\gamma(0) = p$, символом Γ_p . Тогда производная функции вдоль кривой в точке p имеет вид (2.17). Выражение $\partial_\alpha(f \circ \varphi^{-1})$ в правой части равенства зависит только от ростка $[f] \in \mathcal{C}_p$, но не от представителя $f \in [f_p]$. Поэтому будем писать

$$X_{(\gamma)}[f_p] : \mathcal{C}_p \rightarrow \mathbb{R}.$$

Это отображение линейно по росткам:

$$X_{(\gamma)}([f_p] + [g_p]) = X_{(\gamma)}[f_p] + X_{(\gamma)}[g_p], \quad X_{(\gamma)}(a[f_p]) = aX_{(\gamma)}[f_p], \quad a \in \mathbb{R}.$$

Введем обозначение для тех ростков, производные которых вдоль всех кривых $\gamma \in \Gamma_p$, проходящих через точку p , равны нулю

$$\mathcal{H}_p = \{[f_p] \in \mathcal{C}_p : X_{(\gamma)}[f_p] = 0\}.$$

Множество $\mathcal{H}_p$ является линейным подпространством в $\mathcal{C}_p$. В координатах принадлежность $[f_p] \in \mathcal{H}_p$ задается равенством $\partial_\alpha(f \circ \varphi^{-1}) = 0$. То есть подпространство $\mathcal{H}_p$ состоит в точности из тех ростков, у которых все частные производные первого порядка равны нулю.

Определение. Кокасательным пространством в точке $p \in \mathbb{M}$ называется фактор пространство $T_p^*(\mathbb{M}) = \mathcal{C}_p / \mathcal{H}_p$. $\square$

Из данного определения следует, что кокасательное пространство состоит из нулевого вектора, который соответствует $\mathcal{H}_p$, и тех ростков $[f_p]$, для которых хотя бы одна частная производная отлична от нуля. В частности, если две функции отличаются на постоянную, то они принадлежат разным росткам, но одному классу эквивалентности ростков, т.к. все производные от постоянной функции равны нулю. Поскольку две функции принадлежат одному классу эквивалентности ростков $[f_p] \sim [g_p]$ тогда и только тогда, когда совпадают все их частные производные первого порядка, то каждый росток взаимно однозначно определяется градиентом функции. Это доказывает эквивалентность независимого определения кокасательного пространства

определению, данному ранее, т.к. градиент функции есть 1-форма в смысле прежнего определения. В основе данного определения, так же, как и в определении вектора, лежит понятие кривой на многообразии и вектора скорости.

В дальнейшем, в целях упрощения обозначений, мы почти всегда будем писать

$$\partial_\alpha f := \frac{\partial f}{\partial x^\alpha} = \frac{\partial(f \circ \varphi^{-1})}{\partial x^\alpha},$$

несмотря на то, что функция $f(x)$ определена на точках многообразия $x \in \mathbb{M}$, а не в евклидовом пространстве. Это является общепринятой записью и не приводит к путанице. В тех местах, где нужно подчеркнуть различие точки многообразия и точки евклидова пространства, мы будем использовать полную запись.

2.6.4 Векторные поля и дифференцирования

Дадим второе, теперь уже алгебраическое, глобальное определение векторных полей на многообразии $\mathbb{M}$. Зафиксируем произвольную точку $p \in \mathbb{M}$ и рассмотрим некоторую координатную окрестность этой точки $(\mathbb{U}, \varphi)$. С каждым векторным полем $X \in \mathcal{X}(\mathbb{U})$ естественным образом связывается оператор дифференцирования (2.13) в алгебре гладких функций $\mathcal{C}^\infty(\mathbb{U})$. Его действие на функцию в произвольной карте

$$\mathcal{C}^\infty(\mathbb{U}) \ni f \mapsto Xf := X^\alpha \partial_\alpha f \in \mathcal{C}^\infty(\mathbb{U}) \quad (2.20)$$

представляет собой *дифференцирование вдоль векторного поля*. Это дифференцирование не зависит от выбора карты, т.к. запись (2.20) инвариантна относительно преобразований координат, и удовлетворяет свойствам:

- 1) $X(af + bg) = aXf + bXg$ — линейность,
- 2) $X(fg) = (Xf)g + f(Xg)$ — правило Лейбница,

для всех $a, b \in \mathbb{R}$ и $f, g \in \mathcal{C}^\infty(\mathbb{U})$.

Определение. Непрерывное линейное отображение алгебры функций $\mathcal{C}^\infty(\mathbb{U})$ в $\mathbb{R}$:

$$X_p : \mathcal{C}^\infty(\mathbb{U}) \ni f \mapsto X_p f \in \mathbb{R},$$

удовлетворяющее правилу Лейбница 2), называется *дифференцированием* в точке $p \in \mathbb{M}$. Множество всех дифференцирований в данной точке обозначим $\mathbb{D}_p(\mathbb{M})$. $\square$

Замечание. Каждое векторное поле на $\mathbb{U} \subset \mathbb{M}$ отображает алгебру функций $\mathcal{C}^\infty(\mathbb{U})$ в себя. Обратим внимание, что в определении дифференцирования в точке p стоит отображение не в алгебру функций $\mathcal{C}^\infty(\mathbb{U})$, а в вещественную прямую $\mathbb{R}$. Ясно, что если два векторных поля касаются друг друга в точке p , то они порождают одно и то же дифференцирование. $\square$

Множество дифференцирований $\mathbb{D}_p(\mathbb{M})$ снабжается естественной структурой вещественного векторного пространства:

$$\begin{aligned} (X_1 + X_2)f &:= X_1f + X_2f, \\ (aX)f &:= a(Xf), \end{aligned} \quad (2.21)$$

где $X_1, X_2, X \in \mathcal{X}(\mathbb{M})$ и $a \in \mathbb{R}$.

Теорема 2.6.2. *Касательное пространство $\mathbb{T}_p(\mathbb{M})$ и пространство дифференцирований $\mathbb{D}_p(\mathbb{M})$ изоморфны как векторные пространства.*

Доказательство. То, что каждому вектору $X \in \mathbb{T}_p(\mathbb{M})$ однозначно ставится в соответствие дифференцирование было показано выше. Нетрудно проверить, что это отображение сохраняет линейную структуру.

Докажем обратное утверждение. С этой целью рассмотрим важный пример отображений $\mathcal{C}^\infty(\mathbb{U}) \rightarrow \mathbb{R}$:

$$\left(\frac{\partial}{\partial x^\alpha} \right)_p f(p) := \partial_\alpha (f \circ \varphi^{-1})|_{\varphi(p)}, \quad \alpha = 1, \dots, n, \quad (2.22)$$

Эти отображения линейны, удовлетворяют правилу Лейбница и, следовательно, являются дифференцированиями. Подчеркнем, что символ $\left(\frac{\partial}{\partial x^\alpha} \right)_p$ не является частной производной, т.к. определен в точке многообразия $p \in \mathbb{M}$, а не евклидова пространства $\mathbb{R}^n$. В правой же части равенства (2.22) стоит частная производная от функции, определенной в евклидовом пространстве. Теперь докажем два утверждения.

Лемма 2.6.1. *Пусть X – дифференцирование в $\mathcal{C}^\infty(\mathbb{U})$ и $f_c(x) = c = \text{const}$ – постоянная функция на $\mathbb{M}$. Тогда $Xc = 0$.*

Доказательство. Представим постоянную функцию на $\mathbb{M}$ в виде $f_c = cf_1 = cf_1 \cdot f_1$, где $f_1(x) \equiv 1$ – функция, равная единице на всем многообразии $\mathbb{M}$. Используем линейность дифференцирования и правило Лейбница:

$$Xc = cX(1 \cdot 1) = cX(1) + cX(1) = 2cX(1) = 2Xc,$$

что возможно только при $Xc = 0$. Это – нетривиальное использование, казалось бы, тривиального тождества: $1 \cdot 1 = 1$. $\square$

Лемма 2.6.2. *Для любой гладкой функции $f \in \mathcal{C}^\infty(\mathbb{U})$ существует такой набор функций $f_\alpha \in \mathcal{C}^\infty(\mathbb{U})$, $\alpha = 1, \dots, n$, что для любой точки x в некоторой окрестности точки $p \in \mathbb{U}$ выполнены равенства:*

$$f_\alpha(p) = \left(\frac{\partial}{\partial x^\alpha} \right)_p f(p), \quad (2.23)$$

$$f(x) = f(p) + x^\alpha(x)f_\alpha(x), \quad (2.24)$$

где $\{x^\alpha(x)\} : \mathbb{U} \rightarrow \mathbb{R}^n$ – координатные функции.

Доказательство. Для определенности будем считать, что образ точки $p \in \mathbb{U}$ совпадает с началом координат евклидова пространства, $\varphi(p) = (0, \dots, 0) \in \mathbb{R}^n$. Пусть $F(x^1, \dots, x^n) = f \circ \varphi^{-1}$ – координатное представление функции f в некоторой окрестности точки p . Тогда справедливо тождество

$$\begin{aligned} F(x^1, \dots, x^n) = & F(x^1, \dots, x^n) - F(x^1, \dots, x^{n-1}, 0) + \\ & + F(x^1, \dots, x^{n-1}, 0) - F(x^1, \dots, x^{n-2}, 0, 0) + \\ & + \dots + \\ & + F(x^1, 0, \dots, 0) - F(0, \dots, 0) + \\ & + F(0, \dots, 0). \end{aligned}$$

Это тождество перепишем в виде

$$\begin{aligned}
 F(x^1, \dots, x^n) &= F(0, \dots, 0) + \sum_{\alpha=1}^n F(x^1, \dots, tx^\alpha, 0, \dots, 0) \Big|_{t=0}^{t=1} = \\
 &= F(0, \dots, 0) + \sum_{\alpha=1}^n \int_0^1 dt \frac{\partial F}{\partial (tx^\alpha)}(x^1, \dots, x^{\alpha-1}, tx^\alpha, 0, \dots, 0) x^\alpha = \\
 &= F(0, \dots, 0) + x^\alpha F_\alpha(x^1, \dots, x^n),
 \end{aligned} \tag{2.25}$$

где

$$F_\alpha(x^1, \dots, x^n) := \int_0^1 dt \frac{\partial F}{\partial (tx^\alpha)}(x^1, \dots, x^{\alpha-1}, tx^\alpha, 0, \dots, 0)$$

– набор гладких функций в некоторой окрестности начала координат евклидова пространства. Теперь вернемся на многообразие и определим набор функций $f_\alpha := F_\alpha \circ \varphi$. Тогда из последнего равенства (2.25) следует равенство (2.24).

Теперь надо определить вид функций $f_\alpha(x)$. С этой целью применим дифференцирование $\left(\frac{\partial}{\partial x^\alpha}\right)_x$ к равенству (2.24)

$$\left(\frac{\partial}{\partial x^\alpha}\right)_x f(x) = \left(\frac{\partial}{\partial x^\alpha}\right)_x f(p) + \left(\frac{\partial}{\partial x^\alpha}\right)_x x^\beta(x) f_\beta(x) + x^\beta(x) \left(\frac{\partial}{\partial x^\alpha}\right)_x f_\beta(x).$$

Первое слагаемое равно нулю, как следствие леммы 2.6.1. Поскольку $\left(\frac{\partial}{\partial x^\alpha}\right)_x x^\beta(x) = \delta_\alpha^\beta$, то в точке p имеем равенство (2.23), т.к. $x^\alpha(p) = 0$. $\square$

Следствие. Если $X_p \in \mathbb{D}_p(\mathbb{U})$ – дифференцирование в точке $p \in \mathbb{M}$, то

$$X_p = X_p x^\alpha(p) \left(\frac{\partial}{\partial x^\alpha}\right)_p = X_p^\alpha \left(\frac{\partial}{\partial x^\alpha}\right)_p. \tag{2.26}$$

Доказательство. Пусть $(\mathbb{U}, \varphi)$ – произвольная карта в окрестности точки $p \in \mathbb{M}$. Тогда, возможно, $x^\alpha(p) \neq 0$. В этом случае сдвинем начало координат в евклидовом пространстве: $y^\alpha := x^\alpha - x^\alpha(p)$. Тогда из леммы 2.6.2 следует представление

$$f(x) = f(p) + (x^\alpha(x) - x^\alpha(p)) f_\alpha(x).$$

Применяя дифференцирование X к этому равенству и переходя в точку p , получим равенство (2.26). $\square$

Таким образом, множество всех дифференцирований $\mathbb{D}_p(\mathbb{U})$ в произвольной точке $p \in \mathbb{M}$ представляет собой конечномерное векторное пространство, $\dim \mathbb{D}_p(\mathbb{U}) = n$, с базисом $\left(\frac{\partial}{\partial x^\alpha}\right)_p$. Это пространство изоморфно касательному пространству $\mathbb{T}_p(\mathbb{M})$, и теорема доказана. $\square$

Определение. Координатным базисом векторных полей $\mathcal{X}(\mathbb{U})$ на карте $(\mathbb{U}, \varphi)$ многообразия $\mathbb{M}$ называется набор гладких векторных полей

$$\left\{ e_\alpha(x) := \left(\frac{\partial}{\partial x^\alpha}\right)_x \right\} \in \mathcal{X}(\mathbb{U}), \quad \alpha = 1, \dots, n,$$

определенных формулой (2.22). Дуальный базис $\{e^\alpha(x)\} \in \Lambda_1(\mathbb{U})$, $e^\alpha(e_\beta) = \delta_\beta^\alpha$, называется координатным базисом ковекторных полей (1-форм). Координатный базис для ковекторных полей обозначается dx_p^α или просто dx^α . Координатные базисы называют также голономными. $\square$

Замечание. Подчеркнем, что координатный базис – это не набор частных производных, а векторные поля на многообразии. Их действие, как дифференцирований, определено только для достаточно гладких функций $\mathcal{C}^k(\mathbb{U})$. Действие векторных полей $e_\alpha(x)$ на тензоры более высокого ранга не определено. $\square$

Замечание. Обозначение координатных базисов векторных и ковекторных полей через ∂_α и dx^α оправдано простой формулой из математического анализа

$$\frac{\partial x^\alpha}{\partial x^\beta} = \delta_\beta^\alpha.$$

Для тривиальных многообразий $\mathbb{M} \approx \mathbb{R}^n$, покрытых одной картой, векторы e_α можно отождествить с операторами частных производных ∂_α , которые действуют на дифференцируемые функции. Тогда дуальный базис e^α ковекторных полей естественным образом отождествляется с дифференциалами координатных функций dx^α . $\square$

Из разложения по базису (2.26) следует, что компонента векторного поля X в точке $p \in \mathbb{M}$ – это результат действия векторного поля на координатную функцию $X^\alpha(p) = Xx^\alpha(x)|_{x=p}$.

В дальнейшем мы будем писать сокращенно $X = X^\alpha(x)\partial_\alpha$, имея в виду, что на функции $f \circ \varphi^{-1}$ в евклидовом пространстве координатный базис действительно действует, как частная производная. Для 1-форм в координатном базисе мы часто будем использовать общепринятую запись $A = dx^\alpha A_\alpha$, $e^\alpha := dx^\alpha$.

Пусть $D_x(\mathbb{U}) \in \mathbb{D}_x(\mathbb{M})$ – некоторое дифференцирование в точке $x \in \mathbb{M}$. Рассмотрим объединение $D(\mathbb{M}) := \bigcup_{x \in \mathbb{M}} D_x(\mathbb{U})$ по всем точкам многообразия, которое соответствует некоторому векторному полю $X \in \mathcal{X}(\mathbb{M})$. Оно задает отображение

$$D(\mathbb{M}) : \mathcal{C}^\infty(\mathbb{M}) \rightarrow \mathcal{C}^\infty(\mathbb{M}), \quad (2.27)$$

которое называется дифференцированием в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$, т.е. непрерывный линейный эндоморфизм в $\mathcal{C}^\infty(\mathbb{M})$, удовлетворяющий правилу Лейбница. Поэтому каждому векторному полю $X \in \mathcal{X}(\mathbb{M})$ ставится в соответствие некоторое дифференцирование $D(\mathbb{M}) \in \mathcal{D}(\mathbb{M})$, где $\mathcal{D}(\mathbb{M})$ – множество всех дифференцирований в алгебре функций. Верно также и обратное утверждение: любому дифференцированию в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$ соответствует единственное векторное поле. Мы доказали аналогичное утверждение в фиксированной точке многообразия, где пространство $\mathbb{D}_x(\mathbb{U})$ является конечномерным. Доказательство в рассматриваемом бесконечномерном случае сложнее и приведено в [8]. Линейная структура на $\mathcal{D}(\mathbb{M})$ вводится так же, как и в точке (2.21). Таким образом, множество векторных полей $\mathcal{X}(\mathbb{M})$ биективно отображается на множество дифференцирований $\mathcal{D}(\mathbb{M})$, при этом как векторные пространства эти множества изоморфны. Эта биекция позволяет дать эквивалентное алгебраическое определение векторного поля.

Определение. Векторным полем $X \in \mathcal{X}(\mathbb{M})$ на многообразии $\mathbb{M}$ называется дифференцирование в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$. $\square$

2.6.5 Векторные поля и интегральные кривые

Начнем с локального описания. Пусть x^α , $\alpha = 1, \dots, n$, – локальные координаты в окрестности точки $x \in \mathbb{U}_x \subset \mathbb{M}$. Рассмотрим векторное поле $X = X^\alpha \partial_\alpha$, все компоненты которого отличны от нуля в окрестности $\mathbb{U}_x$. Бесконечно малые перемещения

dx^α точки x вдоль этого векторного поля должны быть пропорциональны компонентам X^α и поэтому удовлетворять системе уравнений

$$\frac{dx^1}{X^1} = \frac{dx^2}{X^2} = \dots = \frac{dx^n}{X^n}. \quad (2.28)$$

Эту систему уравнений можно переписать в виде равенства нулю 1-форм:

$$A^M := \frac{dx^M}{X^M} - \frac{dx^1}{X^1} = 0, \quad M = 2, \dots, n.$$

В каждой точке $x \in \mathbb{M}$ векторное поле X задает одномерное подпространство в касательном пространстве $\mathbb{T}_x(\mathbb{M})$, а совокупность 1-форм $\{A^M\}$ – $(n-1)$ -мерное ортогональное дополнение в сопряженном пространстве $\mathbb{T}_x^*(\mathbb{M})$, поскольку $A^M(X) = 0$.

Согласно теории дифференциальных уравнений, уравнения (2.28) допускают $n-1$ функционально независимых решений

$$\varphi^M(x^\alpha) = c^M = \text{const}, \quad M = 2, \dots, n, \quad (2.29)$$

для которых $A^M = d\varphi^M = 0$. При этом прямоугольная $n \times (n-1)$ матрица, составленная из производных $\partial_\alpha \varphi^M$, имеет ранг $n-1$, и каждая из функций φ^M является решением уравнения в частных производных

$$X^\alpha \partial_\alpha \varphi^M = 0. \quad (2.30)$$

Совершим преобразование координат $x^\alpha \mapsto y^\alpha(x)$, выбрав в качестве последних $n-1$ координат функции (2.29) $(y^2, \dots, y^n) := (\varphi^2, \dots, \varphi^n)$, а координату x^1 оставим без изменения. Якобиан этого преобразования отличен от нуля в силу функциональной независимости функций φ^M . Тогда из закона преобразования векторного поля (2.11) и уравнения (2.30) следует, что в новой системе координат все компоненты векторного поля, кроме первой, равны нулю:

$$X^1 \neq 0, \quad X^2 = \dots = X^n = 0.$$

Заменим теперь координату x^1 на функцию $y^1(x)$, которая удовлетворяет дифференциальному уравнению

$$\frac{dy^1}{dx^1} X^1 = 1.$$

Это уравнение локально разрешимо, и, значит, в новой системе координат $X^1 = 1$.

Если у векторного поля X^α часть компонент равнялась нулю до преобразования координат, то все, сказанное выше, можно повторить для ненулевых компонент. Отсюда следует частный случай теоремы Фробениуса, которая будет сформулирована в разделе 2.11.

Теорема 2.6.3. *Для произвольного отличного от нуля векторного поля $X \in \mathcal{X}(\mathbb{M})$ в некоторой окрестности $\mathbb{U}_x$ произвольной точки $x \in \mathbb{M}$ существует такая система координат, в которой все компоненты X , кроме одной (например, первой) обращаются в нуль. Координатную функцию $x^1(x)$, $x \in \mathbb{U}_x$ можно подобрать таким образом, чтобы первая компонента векторного поля X^1 была равна единице в этой окрестности, т.е. $X = \partial_1$.*

Замечание. Эта теорема показывает, что векторное поле в окрестности точки, в которой оно отлично от нуля, устроено довольно просто. Если в некоторой точке векторное поле обращается в нуль, то в окрестности этой точки оно может быть устроено очень сложно [45]. Нули гладкого касательного поля к многообразию связаны с топологическими свойствами. Например, на четномерной сфере не существует векторного поля, нигде не обращающегося в нуль (см. теорему 10.2.1). В то же время такое поле всегда можно задать на торе. $\square$

Если в уравнения (2.29) подставить координаты некоторой фиксированной точки $p = \{x_p^\alpha\} \in \mathbb{M}$, то определяются значения постоянных c^M . При этих значениях постоянных система $n - 1$ трансцендентных уравнений относительно x^α (2.29) определяет кривую, проходящую через точку p . Если кривая γ параметризуется параметром t , $\gamma = \{x^\alpha(t)\}$, то уравнения (2.28) эквивалентны системе обыкновенных дифференциальных уравнений

$$\frac{dx^\alpha}{dt} = X^\alpha(x), \quad (2.31)$$

с начальными условиями

$$x^\alpha|_{t=t_p} = x_p^\alpha, \quad (2.32)$$

где t_p – значение параметра вдоль кривой, при котором она проходит через точку p .

Из существования и единственности решения задачи Коши для системы дифференциальных уравнений (2.31) (см. теорему 2.6.5 ниже) вытекает

Теорема 2.6.4. *Если векторное поле X на многообразии $\mathbb{M}$ дифференцируемо, то через каждую точку $p \in \mathbb{M}$ проходит одна и только одна интегральная кривая этого векторного поля.*

Общее решение системы уравнений (2.31) зависит от n постоянных интегрирования. Одна из постоянных соответствует сдвигу параметра вдоль кривой $t \mapsto t + \text{const}$, а оставшиеся $n - 1$ постоянных определяются положением точки p на гиперповерхности (2.29), проходящей через точку p .

Если решение системы уравнений (2.31) представимо в виде ряда, то вблизи точки p оно выглядит очень просто

$$x^\alpha(t) = x_p^\alpha + X_p^\alpha(t - t_p) + \dots, \quad |t - t_p| \ll 1.$$

То есть компоненты векторного поля определяют главную линейную часть интегральной кривой.

Замечание. Интегральные кривые векторного поля являются ни чем иным, как *характеристиками* (линиями уровня) для решений дифференциального уравнения в частных производных первого порядка (2.30). $\square$

Условие дифференцируемости векторного поля в теореме 2.6.4 можно ослабить.

Определение. Функция (отображение)

$$f : \mathbb{R}^n \supset \mathbb{U} \ni x \mapsto f(x) \in \mathbb{R}^m$$

удовлетворяет *условию Липшица* в области $\mathbb{U}$, если существует такая положительная постоянная C , что выполнено неравенство

$$|f(x_2) - f(x_1)| \leq C|x_2 - x_1|,$$

для всех $x_{1,2} \in \mathbb{U}$. $\square$

В условии Липшица знак модуля обозначает обычный модуль вектора в евклидовом пространстве.

Ясно, что из условия Липшица вытекает непрерывность функции f .

Теорема 2.6.5. Пусть правая часть системы уравнений (2.31) удовлетворяет условию Липшица в области $\mathbb{U}$. Тогда через каждую точку $\mathbb{U}$ проходит одна и только одна интегральная кривая системы уравнений (2.31).

Доказательство. См., например, [46], § 31. □

Если правая часть системы уравнений (2.31) только непрерывна, то и тогда через каждую точку $p \in \mathbb{U}$ проходит хотя бы одна интегральная кривая. Однако единственность может быть нарушена.

Пример 2.6.6. Рассмотрим обыкновенное дифференциальное уравнение

$$\frac{2}{3}\dot{x} = x^{1/3}.$$

Его решения имеют вид

$$x = 0 \quad \text{и} \quad x = \pm(x + c)^{3/2}.$$

Отсюда следует, что через каждую точку $x \neq 0$ проходят две интегральные кривые. Через точку $x = 0$ проходит даже три интегральные кривые: $x = 0$, $x = \pm t^{3/2}$. Если положить в условии Липшица $x_1 = 0$, то оно примет вид

$$|x_2|^{1/3} \leq C|x_2|.$$

Ясно, что такой постоянной C не существует, т.к. правая часть неравенства при $x_2 \rightarrow 0$ стремится к нулю быстрее. □

В дальнейшем мы всегда предполагаем, что векторное поле по крайней мере дифференцируемо.

Предложение 2.6.2. Если дифференцируемое векторное поле X обращается в нуль в некоторой точке многообразия $p \in \mathbb{M}$, то эта точка является неподвижной относительно потока векторного поля, т.е. $x(t) = p$ для всех $t \in \mathbb{R}$. Обратно. Если точка $p \in \mathbb{M}$ является неподвижной на интегральной кривой некоторого векторного поля, то в этой точке векторное поле обращается в нуль, $X(p) = 0$.

Доказательство. Постоянные функции $x^\alpha(t) = x_p^\alpha$ удовлетворяют системе уравнений (2.31), если $X_p^\alpha = 0$. Обратное утверждение очевидно. □

Если векторное поле умножить на произвольную достаточно гладкую отличную от нуля функцию: $X \mapsto fX$, то уравнение для интегральной кривой $x(\tau)$ примет вид

$$\frac{dx^\alpha}{d\tau} = fX^\alpha.$$

Введем новый параметр $t(\tau)$ вдоль кривой, который является решением уравнения

$$\frac{dt}{d\tau} = f.$$

Решение этого уравнения существует и является монотонным, т.к. $f \neq 0$. В новой параметризации уравнение для интегральной кривой принимает прежний вид (2.31). Таким образом, два векторных поля X и fX , отличающиеся умножением на отличную от нуля функцию, имеют интегральные кривые, которые совпадают, как подмножества в $\mathbb{M}$. Отличие сводится только к различным параметризациям кривых.

Перейдем к глобальному описанию.

Определение. *Интегральной кривой* векторного поля $X \in \mathcal{X}(\mathbb{M})$, проходящей через точку $p \in \mathbb{M}$ называется кривая $\gamma : (a, b) \ni t \mapsto x(t) \in \mathbb{M}$ такая, что $\gamma(t_p) = p$ и

$$\gamma_* \partial_t = X(x(t)), \quad (2.33)$$

где γ_* – дифференциал отображения некоторого открытого интервала $(a, b) \subset \mathbb{R}$, содержащего точку t_p , и ∂_t – касательный вектор к интервалу в точке $t \in (a, b)$. Вектор $X(x(t))$ называется *касательным вектором* к кривой γ в точке $x(t)$ (вектором скорости). $\square$

Нетрудно проверить, что в каждой карте уравнение (2.33) записывается в виде системы обыкновенных дифференциальных уравнений (2.31).

Вообще говоря, параметр t в общем случае определен лишь на некотором конечном или полубесконечном интервале $(a, b) \subset \mathbb{R}$ даже для гладких векторных полей, заданных на всем многообразии.

Определение. Векторное поле $X \in \mathcal{X}(\mathbb{M})$ называется *полным*, если все интегральные кривые этого поля определены при всех значениях $t \in \mathbb{R}$. $\square$

Теорема 2.6.6. *На компактном многообразии $\mathbb{M}$ любое гладкое векторное поле $X \in \mathcal{X}(\mathbb{M})$, не обращающееся в нуль, является полным.*

Доказательство. См., например, [1], глава 5, §35, теорема 1. $\square$

Замечание. Для многообразий понятие компакта и компактного пространства совпадают, т.к. многообразие, по определению, является хаусдорфовым пространством.

На некомпактном многообразии векторное поле может быть либо полным, либо неполным.

Пример 2.6.7. Векторному полю ∂_x на вещественной прямой $\mathbb{R}$ соответствуют интегральные кривые $x = t + \text{const}$. Они полны на всей прямой $\mathbb{R}$. Однако они неполны на полупрямой $\mathbb{R}_+$ или любом конечном открытом интервале $(a, b) \subset \mathbb{R}$. $\square$

Пример 2.6.8. Рассмотрим гладкое векторное поле $X = x^2 \partial_x$ на вещественной прямой $\mathbb{R}$. Общее решение уравнения интегральной кривой $\dot{x} = x^2$ имеет вид

$$x = -\frac{1}{t - c}, \quad c = \text{const}.$$

Таким образом, для каждого значения постоянной c , имеются две никак не связанные между собой интегральные кривые:

$$\begin{aligned} \gamma_1 : (-\infty, c) &\rightarrow (0, \infty), \\ \gamma_2 : (c, \infty) &\rightarrow (-\infty, 0). \end{aligned}$$

При этом точка $x = 0$ соответствует бесконечному значению параметра вдоль интегральной кривой $t = \pm\infty$, а бесконечно удаленные точки $x = \pm\infty$ – конечному значению параметра $t = c$. Поэтому векторное поле $X = x^2 \partial_x$ неполно. $\square$

Определение. Точка $p \in \mathbb{M}$, в которой векторное поле обращается в нуль, $X(p) = 0$, называется *особой точкой векторного поля*. Если компоненты векторного поля разлагаются в ряд Тейлора, то система уравнений (2.31) в координатной окрестности особой точки в линейном приближении имеет вид

$$\frac{dx^\alpha}{dt} = x^\beta B_\beta^\alpha, \quad B_\beta^\alpha := \partial_\beta X^\alpha|_{x=p}. \quad (2.34)$$

Если $\det B \neq 0$, то особая точка называется *невырожденной*. $\square$

При преобразовании координат $x^\alpha \mapsto x^{\alpha'}(x)$ уравнение (2.34) сохраняет свой вид. При этом компоненты матрицы B преобразуются по правилу

$$B_\beta^\alpha \mapsto B_{\beta'}^{\alpha'} = \frac{\partial x^\beta}{\partial x^{\beta'}} B_\beta^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha},$$

т.е. подвергаются преобразованию подобия.

Напомним, что любое дифференциальное уравнение n -того порядка можно записать в виде эквивалентной ей системы уравнений первого порядка, состоящей из n уравнений. Следовательно, теория обыкновенных дифференциальных уравнений произвольного порядка, разрешенных относительно старшей производной, сводится к нахождению и исследованию свойств интегральных кривых векторных полей. При этом поведение интегральных кривых в окрестностях особых точек представляет исключительный интерес, т.к. позволяет понять качественное поведение решений.

В общем случае поведение интегральных кривых в окрестности особой точки довольно сложно. Особые точки можно классифицировать, приведя матрицу B к какому-либо каноническому виду с помощью преобразования подобия, что означает переход в новую систему координат.

Пример 2.6.9. Рассмотрим двумерное многообразие (поверхность) $\mathbb{M}$, на котором задано дифференцируемое векторное поле $X \in \mathcal{X}(\mathbb{M})$. Пусть $p \in \mathbb{M}$ – невырожденная особая точка векторного поля X . Выберем систему координат в окрестности особой точки так, чтобы она находилась в начале координат. Обозначим собственные числа матрицы B через λ_1 и λ_2 . В общем случае они комплексны. Из невырожденности следует, что $\lambda_1 \lambda_2 \neq 0$. Невырожденные особые точки в рассматриваемом случае делятся на шесть классов.

а) Седло. $\lambda_1 \lambda_2 < 0$. Пусть собственные числа вещественны и разных знаков. Тогда существует система координат в которой матрица B диагональна и уравнения для интегральных кривых (2.34) примут вид

$$\dot{x}^1 = x^1 \lambda_1, \quad \dot{x}^2 = x^2 \lambda_2. \quad (2.35)$$

Они легко интегрируются:

$$x^1 = C_1 e^{\lambda_1 t}, \quad x^2 = C_2 e^{\lambda_2 t}, \quad t \in \mathbb{R}, \quad (2.36)$$

где $C_{1,2}$ – постоянные интегрирования. Касательный вектор к интегральной кривой имеет угол наклона

$$\frac{dx^2}{dx^1} = \frac{C_2 \lambda_2}{C_1 \lambda_1} e^{(\lambda_2 - \lambda_1)t}.$$

Соответствующие интегральные кривые для $\lambda_1 < 0 < \lambda_2$ показаны на рис.2.8а. Стрелки указывают направление возрастания параметра t . Ни одна из интегральных кривых не проходит через начало координат, которое является неподвижной точкой.

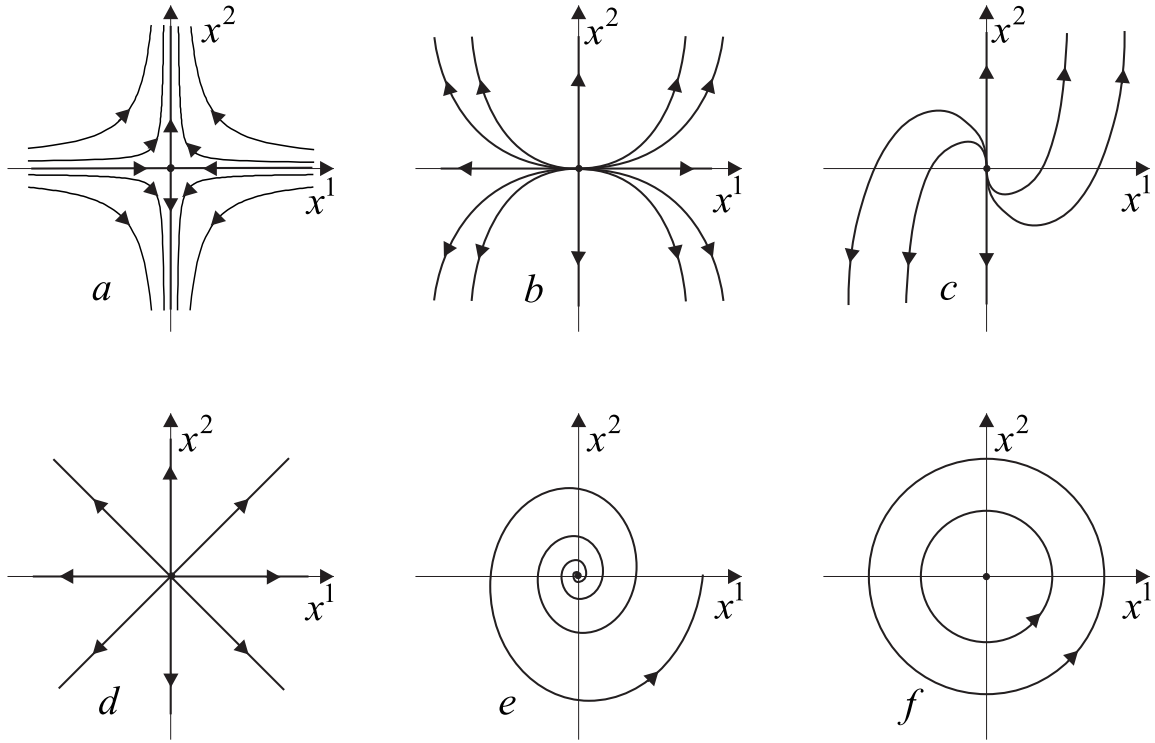


Рис. 2.8: Седло (a), узел (b), жорданов узел (c), дикритический узел (d), фокус (e), центр (f).

b) *Узел.* $\lambda_1 \lambda_2 > 0$, $\lambda_1 \neq \lambda_2$. Если собственные числа матрицы B вещественны одного знака и различны, то ее также можно диагонализировать. При этом уравнения для интегральных кривых и решений имеют прежний вид (2.35), (2.36). Меняется только знак одного из собственных чисел. На рис.2.8b показаны интегральные кривые для $0 < \lambda_1 < \lambda_2$.

c) *Жорданов узел.* $\lambda_1 = \lambda_2$, матрица B недиагонализуема. Пусть собственные числа матрицы B вещественны и равны, и матрицу B нельзя диагонализировать преобразованием подобия. Тогда матрицу B можно привести к жордановой клетке (см. дополнение 37.1)

$$B = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}.$$

Соответствующие уравнения интегральных кривых

$$\dot{x}^1 = x^1 \lambda_1, \quad \dot{x}^2 = x^1 + x^2 \lambda_1,$$

легко интегрируются:

$$x^1 = C_1 e^{\lambda_1 t}, \quad x^2 = (C_2 + C_1 t) e^{\lambda_1 t}, \quad t \in \mathbb{R},$$

Касательная к интегральной кривой имеет угол наклона

$$\frac{dx^2}{dx^1} = \frac{C_2}{C_1} + \frac{1}{\lambda_1} + t.$$

На рис.2.8c показаны интегральные кривые для $0 < \lambda_1$.

d) *Дикритический узел.* $\lambda_1 = \lambda_2$, матрица B кратна единичной. Пусть собственные числа матрицы B вещественны и равны, и матрица B кратна единичной,

$$B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}.$$

При преобразовании координат матрица B не меняется, т.к. пропорциональна единичной матрице. В этом случае уравнения для интегральных кривых,

$$\dot{x}^1 = x^1 \lambda_1, \quad \dot{x}^2 = x^2 \lambda_1,$$

легко интегрируются:

$$x^1 = C_1 e^{\lambda_1 t}, \quad x^2 = C_2 e^{\lambda_1 t}, \quad t \in \mathbb{R},$$

Касательная к интегральной кривой имеет постоянный угол наклона

$$\frac{dx^2}{dx^1} = \frac{C_2}{C_1}.$$

На рис.2.8d показаны интегральные кривые для $0 < \lambda_1$.

e) *Фокус.* $\lambda_{1,2}$ комплексны, $\operatorname{re} \lambda_{1,2} \neq 0$. Поскольку комплексные собственные значения могут встречаться только комплексно сопряженными парами, то $\lambda_1 = \mu + i\tilde{\omega}$ и $\lambda_2 = \mu - i\tilde{\omega}$, где $\mu, \tilde{\omega} \in \mathbb{R}$ и $\mu \neq 0$. В этом случае матрицу B можно привести к виду

$$B = \begin{pmatrix} 0 & -\omega^2 \\ 1 & 2\mu \end{pmatrix},$$

где $\omega^2 := \tilde{\omega}^2 + \mu^2$. Соответствующие уравнения для интегральных кривых принимают вид

$$\begin{aligned} \dot{x}^1 &= x^2, \\ \dot{x}^2 &= -\omega^2 x^1 + 2\mu x^2. \end{aligned}$$

Дифференцирование первого уравнения приводит к уравнению для осциллятора с трением

$$\ddot{x}^1 - 2\mu \dot{x}^1 + \omega^2 x^1 = 0.$$

Отсюда следует, что интегральные кривые имеют вид

$$\begin{aligned} x^1 &= C_1 e^{\mu t} \cos(\tilde{\omega} t + C_2), \\ x^2 &= C_1 e^{\mu t} [\mu \cos(\tilde{\omega} t + C_2) - \tilde{\omega} \sin(\tilde{\omega} t + C_2)]. \end{aligned}$$

На рис.2.8e показаны интегральные кривые для $\tilde{\omega} > 0$, $\mu > 0$.

f) *Центр.* $\lambda_{1,2}$ комплексны, $\operatorname{re} \lambda_{1,2} = 0$. Пусть собственные числа чисто мнимые: $\lambda_1 = -\lambda_2 = i\omega$, $\omega \in \mathbb{R}$. В этом случае матрицу B можно привести к виду

$$B = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$

Уравнения для интегральных кривых,

$$\dot{x}^1 = \omega x^2, \quad \dot{x}^2 = -\omega x^1,$$

соответствуют гармоническому осциллятору

$$\ddot{x}^1 + \omega^2 x^1 = 0$$

и легко интегрируются:

$$x^1 = C_1 \cos(\omega t + C_2), \quad x^2 = -C_1 \sin(\omega t + C_2), \quad t \in \mathbb{R}.$$

Интегральные кривые являются окружностями,

$$(x^1)^2 + (x^2)^2 = C_1^2.$$

Они изображены на рис.2.8f при $\omega > 0$.

Во всех случаях интегральные кривые определены при всех значениях параметра t . Это означает, что векторные поля полны. Для узлов и фокуса интегральные кривые стремятся к началу координат при $t \rightarrow -\infty$. Начало координат в соответствии с предложением 2.6.2 является неподвижной точкой. $\square$

Множество всех полных векторных полей на некомпактном многообразии $\mathbb{M}$ является подмножеством в $\mathcal{X}(\mathbb{M})$. В отличие от всего множества $\mathcal{X}(\mathbb{M})$ это подмножество не образует абелеву группу (модуль) по отношению к сложению.

Пример 2.6.10. Рассмотрим два векторных поля $X = y^2 \partial_x$ и $Y = x^2 \partial_y$ на евклидовой плоскости $\mathbb{R}^2$. Оба поля являются полными, однако их сумма $X + Y$ неполна. Действительно, векторное поле X определяет интегральные кривые:

$$\begin{aligned} \dot{x} = y^2 &\Rightarrow x = y^2 t + C_0, & C_0, C_1 = \text{const}, \\ \dot{y} = 0 &\Rightarrow y = C_1. \end{aligned}$$

Эти интегральные кривые параллельны оси x и проходят через все точки y , кроме $y = 0$, см. рис.2.9,a (стрелки показывают возрастание параметра t). Поскольку все интегральные прямые определены при всех $t \in (-\infty, \infty)$, то векторное поле X полно. Все точки оси абсцисс $(x, 0) \in \mathbb{R}^2$ для векторного поля X являются вырожденными особыми точками и неподвижны. Через них не проходит ни одна интегральная кривая.

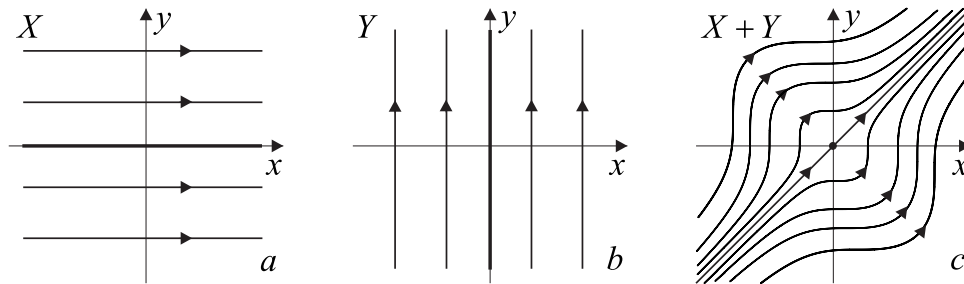


Рис. 2.9: Интегральные кривые для векторного поля X (a), Y (b) и $X + Y$ (c).

Аналогично, векторное поле Y определяет интегральные кривые:

$$\begin{aligned} \dot{x} = 0 &\Rightarrow x = C_2, & C_0, C_2 = \text{const}, \\ \dot{y} = x^2 &\Rightarrow y = x^2 t + C_0. \end{aligned}$$

Интегральные кривые параллельны оси y и проходят через все точки x , кроме $x = 0$, см. рис.2.9,b. Соответствующее векторное поле Y полно. Все точки оси ординат $(0, y) \in \mathbb{R}^2$ являются вырожденными особыми точками и неподвижны.

Уравнения для интегральных кривых, определяемых суммой векторных полей, имеют вид:

$$\dot{x} = y^2, \quad (2.37)$$

$$\dot{y} = x^2. \quad (2.38)$$

При $x \neq 0$ и $y \neq 0$ отсюда следует дифференциальное уравнение для формы интегральной кривой:

$$x^2 dx = y^2 dy \Rightarrow y^3 = x^3 + c, \quad c = \text{const.}$$

Подставляя это решение в уравнение (2.37), получаем равенство

$$\dot{x} = (x^3 + c)^{2/3} \Leftrightarrow t = \int \frac{dx}{(x^3 + c)^{2/3}}. \quad (2.39)$$

При $c = 0$ это уравнение легко интегрируется

$$x = -\frac{1}{t - c_0}, \quad c_0 = \text{const.} \quad (2.40)$$

После подстановки в (2.38) получаем уравнение

$$y = -\frac{1}{t - c_0} + c_1, \quad c_1 = \text{const.} \quad (2.41)$$

Таким образом, мы получили интегральные кривые для векторного поля $X + Y$, которые неполны, т.к. уходят в бесконечность $x = \pm\infty$, $y = \pm\infty$ при конечном значении параметра $t = c_0$. Остальные интегральные кривые можно не исследовать, поскольку, по определению, векторное поле полно, если все интегральные кривые определены при всех $t \in \mathbb{R}$.

Интегральные кривые для суммы векторных полей показаны на рис.2.9,c. Все интегральные кривые неполны в бесконечности. Это легко видеть, т.к. при $x \rightarrow \pm\infty$ постоянной c в интеграле (2.39) можно пренебречь, и, следовательно, справедлива асимптотика (2.40). Матрица B (2.34), определяющая линейное приближение, для суммы векторных полей $X + Y$ имеет вид

$$B = \begin{pmatrix} 0 & 2x \\ 2y & 0 \end{pmatrix}.$$

Она вырождена только в начале координат. Поэтому точка $(0, 0)$ является вырожденной особой точкой. К ней подходят интегральные кривые (2.40), (2.41) при бесконечном значении параметра t . $\square$

Ранее мы показали, что всюду отличное от нуля дифференцируемое векторное поле, заданное на многообразии $\mathbb{M}$, определяет семейство интегральных кривых, проходящих через каждую точку $\mathbb{M}$, причем через каждую точку проходит единственная кривая.

Определение. Пусть $s(t, p)$ – интегральная кривая векторного поля $X \in \mathcal{X}(\mathbb{M})$, проходящая через точку p , $s(0, p) = p$. Будем считать, что векторное поле полно. Тогда отображение

$$s : \mathbb{R} \times \mathbb{M} \ni t, p \mapsto s(t, p) \in \mathbb{M}, \quad (2.42)$$

генерируемое векторным полем X , называется *поток векторного поля*. $\square$

Предложение 2.6.3. *Отображение (2.42) удовлетворяет тождеству*

$$s(t_1, s(t_2, p)) = s(t_1 + t_2, p), \quad (2.43)$$

для всех значений $t_1, t_2 \in \mathbb{R}$, для которых формула (2.43) имеет смысл.

Доказательство. Предложение следует из единственности решения системы дифференциальных уравнений. Действительно, в произвольной карте выполнено равенство

$$\begin{aligned} \frac{d}{dt_1} s^\alpha(t_1, s(t_2, p)) &= X^\alpha(s(t_2, p)) \\ s(0, s(t_2, p)) &= s(t_2, p). \end{aligned}$$

С другой стороны

$$\begin{aligned} \frac{d}{dt_1} s^\alpha(t_1 + t_2, p) &= \frac{d}{d(t_1 + t_2)} s^\alpha(t_1 + t_2, p) = X^\alpha(s(t_1 + t_2, p)), \\ s(0 + t_2, p) &= s(t_2, p). \end{aligned}$$

Тем самым и правая, и левая часть равенства (2.43) удовлетворяют одной и той же системе уравнений с одинаковыми начальными условиями. $\square$

Замечание. Поток векторного поля можно представлять себе, как стационарный поток жидкости. В этом случае параметр t является временем, а X – векторным полем скорости частиц жидкости. $\square$

Пример 2.6.11. Рассмотрим гладкое векторное поле $K = -y\partial_x + x\partial_y$ на евклидовой плоскости $\mathbb{R}^2$. Нетрудно проверить, что поток этого векторного поля имеет вид

$$s : \mathbb{R} \times \mathbb{R}^2 \ni t \times (x, y) \mapsto (x \cos t - y \sin t, x \sin t + y \cos t) \in \mathbb{R}^2.$$

Интегральная кривая, проходящая через точку (x, y) , представляет собой окружность с центром в начале координат. В начале координат векторное поле обращается в нуль, и интегральная кривая вырождается в точку. Матрица (2.34) имеет вид

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Ее собственные числа являются мнимыми $\lambda_{1,2} = \pm i$. Поэтому начало координат является для векторного поля K невырожденной особой точкой – центром.

Векторное поле K является ничем иным, как векторным полем Киллинга двумерных вращений евклидовой плоскости и принимает особо простой вид в полярных координатах, $K = \partial_\varphi$. Интегральные кривые поля K являются в данном случае траекториями Киллинга. Поток векторного поля на $\mathbb{R}^2$ определяется независимо от наличия метрики. Однако интерпретация векторного поля K , как поля Киллинга уже связана с наличием на плоскости евклидовой метрики. $\square$

Пример 2.6.12. Рассмотрим гладкое векторное поле $K = y\partial_x + x\partial_y$ на плоскости $\mathbb{R}^2$. Поток этого векторного поля имеет вид

$$s : \mathbb{R} \times \mathbb{R}^2 \ni t \times (x, y) \mapsto (x \operatorname{ch} t + y \operatorname{sh} t, x \operatorname{sh} t + y \operatorname{ch} t) \in \mathbb{R}^2.$$

Интегральная кривая, проходящая через точку (x, y) , является ветвью гиперболы $x^2 - y^2 = \text{const}$ с центром в начале координат. При $y = \pm x$ гиперболы вырождаются в прямые линии, проходящие через начало координат под углом $\pm\pi/4$. Матрица (2.34) имеет вид

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Она невырождена и имеет различные вещественные собственные значения $\lambda_{1,2} = \pm 1$. Поэтому начало координат является невырожденной особой точкой – седлом.

Векторное поле K является векторным полем Киллинга для метрики Лоренца, заданной на плоскости $\mathbb{R}^2$, а интегральные кривые – траекториями Киллинга. $\square$

Определение. При фиксированном значении параметра t поток $s(t, x)$ представляет собой диффеоморфизм, обозначаемый также

$$s_t : \mathbb{M} \rightarrow \mathbb{M}.$$

Из предложения 2.6.3 следует, что он представляет собой абелеву группу:

- 1) $s_{t_1} s_{t_2} = s_{t_1+t_2}$;
- 2) s_0 – единичный элемент;
- 3) $s_t^{-1} = s_{-t}$ – обратный элемент.

Эта группа называется *однопараметрической группой преобразований*, генерируемой векторным полем X . Действительно, из системы уравнений (2.31) следует, что при малых значениях параметра поток имеет вид

$$s_\epsilon : x^\alpha \mapsto x^\alpha + \epsilon X^\alpha.$$

То есть векторное поле X является генератором бесконечно малых преобразований многообразия $\mathbb{M}$. $\square$

Замечание. Псевдогруппа гладких преобразований координат $\operatorname{diff} \mathbb{M}$ на многообразии $\mathbb{M}$, бесконечномерна. С соответствующими оговорками множество гладких векторных полей $\mathcal{X}(\mathbb{M})$ можно рассматривать, как бесконечномерную алгебру Ли для $\operatorname{diff} \mathbb{M}$. $\square$

Сам поток часто обозначают

$$s^\alpha(t, x) = \exp(tX)x^\alpha$$

и называют *экспоненциальным отображением*. Это обозначение оправдано следующим образом. Разложим функцию $s^\alpha(t, x)$ в ряд Тейлора по t :

$$\begin{aligned} s^\alpha(t, x) &= x^\alpha + t \left. \frac{\partial s^\alpha(\tau, x)}{\partial \tau} \right|_{\tau=0} + \frac{t^2}{2!} \left. \frac{\partial^2 s^\alpha(\tau, x)}{\partial \tau^2} \right|_{\tau=0} + \dots = \\ &= \left[1 + t \frac{\partial}{\partial \tau} + \frac{t^2}{2!} \frac{\partial^2}{\partial \tau^2} + \dots \right] s^\alpha(\tau, x) \Big|_{\tau=0} = \\ &= \exp \left(t \frac{\partial}{\partial \tau} \right) s^\alpha(\tau, x) \Big|_{\tau=0} = \exp(tX)x^\alpha. \end{aligned}$$

При этом выполнены формальные свойства экспоненты:

$$\begin{aligned}\exp(0X)x &= x, \\ \exp(t_1X)\exp(t_2X)x &= \exp[(t_1 + t_2)X]x, \\ \frac{d}{dt}\exp(tX)x &= X\exp(tX)x.\end{aligned}$$

Выше было показано, что любое полное векторное поле генерирует единственную однопараметрическую группу преобразований. Верно и обратное утверждение: любая однопараметрическая группа $s^\alpha(t, x)$ определяет векторное поле. Для этого достаточно положить

$$X^\alpha := \left. \frac{ds^\alpha}{dt} \right|_{t=0}.$$

Если векторное поле $X(\mathbb{M})$ является неполным, то понятие потока и однопараметрической группы преобразований можно ввести только локально (см., например, [47].)

2.6.6 1-формы и гиперповерхности

Рассмотрим отличную от нуля в каждой точке 1-форму $dx^\alpha A_\alpha(x)$ в некоторой карте на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$. Тогда линейное алгебраическое уравнение

$$dx^\alpha A_\alpha = 0 \tag{2.44}$$

относительно дифференциалов dx^α имеет $n - 1$ линейно независимых решений в каждой точке $x \in \mathbb{M}$. При этом любое решение уравнения (2.44) является линейной комбинацией данных решений. Отсюда следует, что совокупность векторов dx^α , удовлетворяющих соотношению (2.44), задает $(n - 1)$ -мерное подпространство в касательном пространстве $\mathbb{T}_x(\mathbb{M})$. Таким образом, отличная от нуля 1-форма задает распределение $(n - 1)$ -мерных подпространств в касательном расслоении $\mathbb{T}(\mathbb{M})$.

Казалось бы, что существуют такие $(n - 1)$ -мерные подмногообразия в $\mathbb{M}$, что касательные векторы к ним образуют $(n - 1)$ -мерное распределение, задаваемое 1-формой. Однако в общем случае это не так. Критерий существования таких подмногообразий дает теорема Фробениуса 2.11.4, которая будет рассмотрена несколько позже. Сейчас мы остановимся на простейшем случае.

Пусть 1-форма (2.44) является точной, т.е. имеет вид,

$$dx^\alpha A_\alpha = dx^\alpha \partial_\alpha f, \tag{2.45}$$

для некоторой функции $f(x) \in \mathcal{C}^1(\mathbb{M})$. Тогда уравнение (2.44) можно рассмотреть, как дифференциальное уравнение на f , любое решение которого имеет вид

$$f(x) = \text{const.} \tag{2.46}$$

Уравнение (2.46) при разных значениях константы определяет семейство $(n - 1)$ -мерных подмногообразий в $\mathbb{M}$, которые называются *гиперповерхностями*. При $n = 2$ эти подмногообразия называются *линиями уровня*. В этом случае касательные векторы к гиперповерхностям определяют те же $(n - 1)$ -мерные подпространства в касательном пространстве, что и 1-форма. Говорят также, что гиперповерхность имеет коразмерность один.

Не ограничивая общности, для непостоянной функции можно считать, что только $\partial_1 f \neq 0$ в некоторой области. Поэтому функцию f можно выбрать в качестве первой координаты $f := x^1$. Тогда семейство гиперповерхностей будет задано уравнением $x^1 = \text{const}$, а направления вдоль остальных координатных осей будут определять касательные к гиперповерхности направления.

2.6.7 Алгебра Ли векторных полей

Ранее мы дали два эквивалентных определения векторных полей $\mathcal{X}(\mathbb{M})$ как сечений касательного расслоения $\mathbb{T}(\mathbb{M})$ и как дифференцирований в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$. Это – разные определения, одно из которых может иметь определенные преимущества в той или иной ситуации. При рассмотрении алгебраических вопросов, как правило, удобнее использовать алгебраическое определение векторных полей через дифференцирования. Используя это определение, мы введем на множестве векторных полей $\mathcal{X}(\mathbb{M})$ структуру алгебры Ли.

Определение. Последовательное применение двух дифференцирований (векторных полей) X и Y к некоторой функции f снова дает функцию из $\mathcal{C}^\infty(\mathbb{M})$. Определим композицию двух дифференцирований $X \circ Y$ формулой $(X \circ Y)f := X(Yf)$. Отображение $(X \circ Y)$ является линейным, однако правило Лейбница для него не выполнено:

$$(X \circ Y)fg = XfYg + fX(Yg) + XgYf + gX(Yf) \neq f(X \circ Y)g + g(X \circ Y)f.$$

Это означает, что композиция векторных полей $X \circ Y$ не является векторным полем. По другому, отображение $X \circ Y$ в координатах содержит не только первые, но и вторые производные. Рассмотрим композицию этих дифференцирований в другом порядке,

$$(Y \circ X)fg = YfXg + fY(Xg) + YgXf + gY(Xf).$$

Нетрудно проверить, что разность $Y \circ X - X \circ Y$ удовлетворяет правилу Лейбница

$$(X \circ Y - Y \circ X)fg = f(X \circ Y - Y \circ X)g + g(X \circ Y - Y \circ X)f,$$

т.е. является векторным полем. Эта разность называется *коммутатором векторных полей* или *скобкой Ли* и обозначается

$$[X, Y] := X \circ Y - Y \circ X. \quad \square$$

Из определения следует, что коммутатор двух векторных полей антисимметричен,

$$[X, Y] = -[Y, X], \quad (2.47)$$

и коммутаторы трех произвольных векторных полей удовлетворяют *тождеству Якоби*

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0, \quad (2.48)$$

где слагаемые отличаются циклической перестановкой.

Замечание. Антисимметрия коммутатора (2.47) эквивалентна условию $[X, X] = 0$. Действительно, то, что это условие вытекает из (2.47) очевидно. Для доказательства обратного утверждения достаточно рассмотреть уравнение $[X + Y, X + Y] = 0$. $\square$

Алгебра Ли является неассоциативной алгеброй, при этом условие ассоциативности заменяется на тождества Якоби.

Рассмотрим векторные поля в произвольной карте $X = X^\alpha \partial_\alpha$, $Y = Y^\alpha \partial_\alpha$. Тогда коммутатор дает новое векторное поле

$$Z := [X, Y] = (X^\beta \partial_\beta Y^\alpha - Y^\beta \partial_\beta X^\alpha) \partial_\alpha. \quad (2.49)$$

Используя закон преобразования компонент векторных полей (2.11), нетрудно проверить, что выражение в правой части инвариантно относительно преобразования координат.

Определение. Множество векторных полей с операцией сложения и коммутирования, которое удовлетворяет условиям (2.47) и (2.48) образует *кольцо Ли*. Умножение векторного поля на числа снова дает векторное поле, при этом коммутатор (2.49) билинеен

$$\begin{aligned} [aX + bY, Z] &= a[X, Z] + b[Y, Z], \\ [X, aY + bZ] &= a[X, Y] + b[X, Z], \end{aligned}$$

где $a, b \in \mathbb{R}$. Множество векторных полей с операциями умножения на вещественные числа, сложения и коммутирования образует *алгебру Ли* над полем вещественных чисел. Эта алгебра бесконечномерна и также обозначается $\mathcal{X}(\mathbb{M})$. $\square$

Алгебры Ли образуют не только векторные поля на многообразиях. Структуру алгебры Ли можно также ввести на абстрактном векторном пространстве.

Пример 2.6.13. Двумерное векторное пространство с базисом e_1 и e_2 становится алгеброй Ли, если положить

$$[e_1, e_1] = [e_2, e_2] = 0, \quad [e_1, e_2] = e_2$$

и продолжить эту операцию по линейности. Это – алгебра Ли группы Ли аффинных преобразований прямой, которая рассмотрена в разделе 8.7. $\square$

Пример 2.6.14. Трехмерное векторное пространство $\mathbb{R}^3$ с ортонормальным базисом e_i , $i = 1, 2, 3$, является алгеброй Ли, если в качестве коммутатора двух векторов $X, Y \in \mathbb{R}^3$ выбрать их векторное произведение

$$[X, Y]^i := -\varepsilon^{ijk} X_j Y_k,$$

где ε^{ijk} – полностью антисимметричный тензор третьего ранга и опускание индексов производится с помощью евклидовой метрики δ_{ij} . Эта алгебра Ли совпадает с алгеброй Ли группы трехмерных вращений $\mathbb{SO}(3)$ (1.130). $\square$

Пример 2.6.15. Векторное пространство $\mathfrak{gl}(n, \mathbb{R})$ всех вещественных $n \times n$ -матриц образует конечномерную алгебру Ли, если положить

$$[A, B] := AB - BA, \quad A, B \in \mathfrak{gl}(n, \mathbb{R}),$$

где AB – обычное произведение матриц. При этом $\dim \mathfrak{gl}(n, \mathbb{R}) = n^2$. $\square$

Забегаая вперед (см. раздел 6.2), заметим, что коммутатор векторных полей (2.49) можно записать в эквивалентном виде, используя ковариантную производную и тензор кручения

$$[X, Y] = (X^\beta \nabla_\beta Y^\alpha - Y^\beta \nabla_\beta X^\alpha - X^\beta Y^\gamma T_{\beta\gamma}^\alpha) \partial_\alpha.$$

В таком виде правая часть этого равенства явно ковариантна.

2.7 Тензорные поля

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$. В каждой точке $x \in \mathbb{M}$ у нас есть два n -мерных векторных пространства: касательное $\mathbb{T}_x(\mathbb{M})$ и кокасательное $\mathbb{T}_x^*(\mathbb{M})$ пространства. Рассмотрим их тензорное произведение

$$\mathbb{T}_{s,x}^r(\mathbb{M}) := \underbrace{\mathbb{T}_x(\mathbb{M}) \otimes \dots \otimes \mathbb{T}_x(\mathbb{M})}_r \otimes \underbrace{\mathbb{T}_x^*(\mathbb{M}) \otimes \dots \otimes \mathbb{T}_x^*(\mathbb{M})}_s, \quad (2.50)$$

где мы взяли r экземпляров касательного и s экземпляров кокасательного пространства. Для определенности мы фиксировали порядок сомножителей. Таким образом в каждой точке многообразия мы построили векторное пространство размерности $\dim \mathbb{T}_{s,x}^r(\mathbb{M}) = n^{r+s}$.

Определение. Объединение

$$\mathbb{T}_s^r(\mathbb{M}) := \bigcup_{x \in \mathbb{M}} \mathbb{T}_{s,x}^r(\mathbb{M}),$$

взятое по всем точкам многообразия, называется *расслоением тензоров типа (r, s)* на многообразии $\mathbb{M}$. Сечение этого расслоения $T_s^r(x)$ называется *тензорным полем типа (r, s)* или r раз *контравариантным* и s раз *ковариантным* тензорным полем на многообразии $\mathbb{M}$. Число $r + s$ называется *рангом* тензорного поля. $\square$

Базой этого расслоения является многообразие $\mathbb{M}$, типичным слоем – векторное пространство

$$\underbrace{\mathbb{R}^n \otimes \dots \otimes \mathbb{R}^n}_r \otimes \underbrace{\mathbb{R}^n \otimes \dots \otimes \mathbb{R}^n}_s,$$

где $\mathbb{R}^n$ – типичный слой касательного расслоения. Слой над $x \in \mathbb{M}$ является векторное пространство (2.50) (тем самым мы определили проекцию). Дифференцируемая структура на расслоении тензоров задается дифференцируемыми структурами на базе и в типичном слое аналогично тому, как она была построена для касательного расслоения.

Координатные базисы в касательном и кокасательном пространствах, $e_\alpha = \partial_\alpha$ и $e^\alpha = dx^\alpha$, индуцируют координатный базис в тензорном произведении, который мы обозначим

$$e_{\alpha_1} \otimes \dots \otimes e_{\alpha_r} \otimes e^{\beta_1} \otimes \dots \otimes e^{\beta_s}. \quad (2.51)$$

Напомним, что тензорное произведение векторов не является коммутативным,

$$e_\alpha \otimes e^\beta \neq e^\beta \otimes e_\alpha,$$

поэтому порядок следования базисных векторов в произведении (2.51) фиксирован: сначала мы пишем базисные векторы касательного, а затем кокасательного пространства.

Рассмотрим произвольную карту $(\mathbb{U}, \varphi)$ на многообразии. Тогда тензорное поле типа (r, s) в координатах имеет вид

$$T_s^r(x) = T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}(x) e_{\alpha_1} \otimes \dots \otimes e_{\alpha_r} \otimes e^{\beta_1} \otimes \dots \otimes e^{\beta_s}. \quad (2.52)$$

Нижние и верхние индексы называют, соответственно, *ковариантными* и *контравариантными*. Общее число индексов $r + s$ равно рангу тензорного поля.

Замечание. Для определенности, у компонент $T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}(x)$ мы сначала выписали все ковариантные индексы, а затем – все контравариантные. Порядок индексов зафиксирован порядком сомножителей в правой части (2.50) и принятом нами соглашением для записи тензорного поля типа $(1, 1)$:

$$dx^\beta T_\beta^\alpha \partial_\alpha.$$

Контравариантные индексы, так же, как и ковариантные, упорядочены между собой. В разделе 4 будет введена операция опускания и подъема индексов с помощью метрики. Она будет неоднозначной, если порядок контравариантных и ковариантных индексов не фиксирован.

Ниже мы построим тензорную алгебру для тензорных полей вида (2.52). Аналогично можно построить тензорную алгебру для произвольного расположения сомножителей в правой части (2.50), когда касательные и кокасательные пространства чередуются в произвольном порядке. Мы будем предполагать, что все индексы упорядочены определенным образом. В этом случае обозначение векторного поля $T_s^r(x)$ является грубым, т.к. учитывает только общее число ковариантных и контравариантных индексов, а не их последовательность. $\square$

Набор функций $X_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}(x)$ с r верхними и s нижними индексами называется *компонентами* тензорного поля типа (r, s) в карте (U, φ) . Тензорное поле называется гладким, если все компоненты – гладкие функции. При преобразованиях координат каждый контравариантный индекс умножается на матрицу Якоби (1.68), так же, как и компоненты вектора, а каждый ковариантный индекс – на обратную матрицу Якоби, так же, как и 1-форма.

Пример 2.7.1. Компоненты тензорного поля типа $(1, 1)$ при преобразовании координат $x^\alpha \mapsto x^{\alpha'}(x)$ преобразуются по правилу:

$$X_{\beta'}^{\alpha'} = \frac{\partial x^\beta}{\partial x^{\beta'}} X_\beta^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha}. \quad (2.53)$$

Аналогично преобразуются компоненты тензорных полей произвольного типа. $\square$

В дальнейшем, для краткости, тензорные поля мы часто будем называть просто тензорами.

Очевидно, что, если все компоненты тензорного поля равны нулю в какой-то одной системе координат, то они равны нулю во всех остальных системах отсчета. У нетривиальных тензоров хотя бы одна компонента должна быть отлична от нуля. В общем случае у тензора типа (r, s) на многообразии размерности n имеется n^{r+s} независимых компонент в каждой точке.

Обозначим множество гладких тензорных полей типа (r, s) символом $\mathcal{T}_s^r(\mathbb{M})$. При этом $\mathcal{T}_0^0(\mathbb{M}) = \mathcal{C}^\infty(\mathbb{M})$, $\mathcal{T}_0^1 = \mathcal{X}(\mathbb{M})$ и $\mathcal{T}_1^0 = \Lambda_1(\mathbb{M})$. В дальнейшем индекс 0 у множеств тензорных полей, имеющих только контравариантные или ковариантные индексы, писаться не будет: $\mathcal{T}_0^r(\mathbb{M}) = \mathcal{T}^r(\mathbb{M})$ и $\mathcal{T}_s^0(\mathbb{M}) = \mathcal{T}_s(\mathbb{M})$.

Тензорные поля фиксированного типа в каждой точке можно складывать и умножать на числа, т.е. они образуют (бесконечномерное) векторное пространство над полем вещественных чисел. Кроме того, тензорное поле произвольного типа можно поточечно умножать на произвольные функции, при этом получится новое тензорное поле того же типа. Таким образом, они образуют модуль над алгеброй гладких функций $\mathcal{C}^\infty(\mathbb{M})$.

Введем обозначение для прямой суммы тензорных полей

$$\mathcal{T}(\mathbb{M}) := \bigoplus_{r,s=0}^{\infty} \mathcal{T}_s^r(\mathbb{M}).$$

На множестве $\mathcal{T}(\mathbb{M})$ можно ввести поточечное тензорное умножение, которое двум тензорам типа (r, s) и (p, q) ставит в соответствие тензорное поле типа $r + p, s + q$. А именно, зафиксируем точку $x \in \mathbb{M}$. Из универсального факторизационного свойства тензорного произведения следует, что существует единственное билинейное отображение из $\mathbb{T}_{s,x}^r \times \mathbb{T}_{q,x}^p$ в $\mathbb{T}_{s+q,x}^{r+p}$, которое отображает пару тензоров

$$\begin{aligned} (X_1 \otimes \cdots \otimes X_r \otimes A_1 \otimes \cdots \otimes A_s) &\in \mathbb{T}_{s,x}^r, \\ (Y_1 \otimes \cdots \otimes Y_p \otimes B_1 \otimes \cdots \otimes B_q) &\in \mathbb{T}_{q,x}^p \end{aligned}$$

в тензор типа $r + p, s + q$:

$$X_1 \otimes \cdots \otimes X_r \otimes Y_1 \otimes \cdots \otimes Y_p \otimes A_1 \otimes \cdots \otimes A_s \otimes B_1 \otimes \cdots \otimes B_q \in \mathbb{T}_{s+q,x}^{r+p}. \quad (2.54)$$

Это отображение называется тензорным произведением тензоров в данной точке.

Замечание. Мы зафиксировали порядок сомножителей в произведении (2.54), который соответствует тензорному произведению (2.50).

Определение. Тензорным произведением тензорных полей $\mathcal{T}_s^r(\mathbb{M})$ и $\mathcal{T}_q^p(\mathbb{M})$ называется тензорное поле типа $\mathcal{T}_{s+q}^{r+p}(\mathbb{M})$, полученное поточечным тензорным произведением (2.54). $\square$

Чтобы получить выражение для компонент тензорного произведения в определенной карте, достаточно в качестве векторных X, Y и ковекторных A, B полей в определении тензорного произведения выбрать координатный базис. Пусть в некоторой карте задано два тензорных поля:

$$\begin{aligned} K &= K_{\beta_1 \cdots \beta_s}^{\alpha_1 \cdots \alpha_r} e_{\alpha_1} \otimes e_{\alpha_r} \otimes e^{\beta_1} \cdots \otimes e^{\beta_s} \in \mathcal{T}_s^r(\mathbb{M}), \\ L &= L_{\delta_1 \cdots \delta_q}^{\gamma_1 \cdots \gamma_p} e_{\gamma_1} \otimes e_{\gamma_p} \otimes e^{\delta_1} \cdots \otimes e^{\delta_q} \in \mathcal{T}_q^p(\mathbb{M}). \end{aligned}$$

Тогда компоненты их тензорного произведения

$$(K \otimes L)_{\beta_1 \cdots \beta_s \delta_1 \cdots \delta_q}^{\alpha_1 \cdots \alpha_r \gamma_1 \cdots \gamma_p} = K_{\beta_1 \cdots \beta_s}^{\alpha_1 \cdots \alpha_r} L_{\delta_1 \cdots \delta_q}^{\gamma_1 \cdots \gamma_p}$$

просто равны произведению компонент каждого сомножителя, как чисел.

Пример 2.7.2. Произведение двух векторных полей $X \otimes Y$ дает контравариантный тензор второго ранга с компонентами

$$T(x) = T^{\alpha\beta}(x) e_{\alpha} \otimes e_{\beta} = X(x) \otimes Y(x) = X^{\alpha}(x) Y^{\beta}(x) e_{\alpha} \otimes e_{\beta},$$

Эта операция является некоммутативной, поскольку первый индекс поля $T(x)$ относится к векторному полю X , а не Y . $\square$

Вместе с тензорным умножением, множество тензорных полей $\mathcal{T}(\mathbb{M})$ образует некоммутативную ассоциативную *тензорную алгебру* над полем вещественных чисел. Эта алгебра бесконечномерна, поскольку векторное пространство тензоров фиксированного типа бесконечномерно само по себе и, вдобавок, ранг тензоров неограничен. Алгебра тензоров имеет естественную градуировку, как прямая сумма тензоров фиксированного типа. Образующими тензорной алгебры являются векторные поля и 1-формы.

Определение. Каждой паре индексов (i, j) таких, что $1 \leq i \leq r$ и $1 \leq j \leq s$, мы ставим в соответствие линейное отображение $C^{ij} : \mathcal{T}_s^r(\mathbb{M}) \rightarrow \mathcal{T}_{s-1}^{r-1}(\mathbb{M})$ с помощью следующей формулы

$$\begin{aligned} X_1 \otimes \dots \otimes X_r \otimes A_1 \otimes \dots \otimes A_s &\mapsto \\ \mapsto (A_j, X_i) X_1 \otimes \dots \otimes X_{i-1} \otimes X_{i+1} \otimes \dots \otimes X_r \otimes A_1 \otimes \dots \otimes A_{j-1} \otimes \dots \otimes A_{j+1} \otimes A_s, \end{aligned}$$

где $(A_j, X_i) = A_j(X_i)$ – значение 1-формы A_j на векторе X_i . Это отображение называется *сверткой* и обозначается C^{ij} . $\square$

Компоненты свернутого тензора $K \in \mathcal{T}_s^r(\mathbb{M})$ имеют вид

$$K_{\beta_1 \dots \beta_{j-1} \gamma \beta_{j+1} \dots \beta_s}^{\alpha_1 \dots \alpha_{i-1} \gamma \alpha_{i+1} \dots \alpha_r} = \delta_{\alpha_i}^{\beta_j} K_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r},$$

где произведена свертка по одному верхнему и одному нижнему индексу.

Пример 2.7.3. Тензору типа $(1, 1)$ ставится в соответствие скалярное поле $\text{tr } K = K_\alpha^\alpha$, которое называется *следом* тензора $K = K_\beta^\alpha e_\alpha \otimes e^\beta$. $\square$

Пример 2.7.4. Значением 1-формы $A = dx^\alpha A_\alpha$ на векторном поле $X = X^\alpha \partial_\alpha$ является свертка тензорного произведения $A \otimes X$: $(A, X) := A(X) = X^\alpha A_\alpha$. $\square$

Определение. Тензорное поле $K \in \mathcal{T}_s^r(\mathbb{M})$ называется *разложимым*, если его можно представить в виде

$$K = X_1 \otimes \dots \otimes X_r \otimes A_1 \otimes \dots \otimes A_s,$$

для некоторых векторов: $X_i \in \mathcal{T}^1(\mathbb{M}) = \mathcal{X}(\mathbb{M})$, $i = 1, \dots, r$ и 1-форм: $A_j \in \mathcal{T}_1(\mathbb{M}) = \Lambda_1(\mathbb{M})$, $j = 1, \dots, s$. $\square$

Пример 2.7.5. Сумма двух разложимых контравариантных тензоров второго ранга

$$X_1 \otimes Y_1 + X_2 \otimes Y_2 = (X_1^\alpha Y_1^\beta + X_2^\alpha Y_2^\beta) e_\alpha \otimes e_\beta$$

может не быть разложимым тензором. $\square$

Если тензор имеет два или более индексов одного типа, то с помощью симметризации или антисимметризации по верхним или нижним индексам можно строить новые тензорные поля. Поскольку преобразование координат действует одинаково на каждый ковариантный и контравариантный индекс, то симметризация и антисимметризация индексов является инвариантной операцией и свойство симметрии по индексам сохраняется при преобразовании координат.

Определение. Тензорное поле называется *неприводимым*, если нельзя найти такие линейные комбинации его компонент с постоянными коэффициентами, которые сами образовывали бы тензор. $\square$

Пример 2.7.6. Скалярные, векторные поля и 1-формы являются неприводимыми тензорными полями. $\square$

Пример 2.7.7. Ковариантные или контравариантные тензорные поля второго ранга приводимы, т.к. их компоненты можно разложить на симметричную и антисимметричную неприводимую части:

$$X_{\alpha\beta} = X_{(\alpha\beta)} + X_{[\alpha\beta]}, \quad (2.55)$$

где

$$X_{(\alpha\beta)} := \frac{1}{2}(X_{\alpha\beta} + X_{\beta\alpha}), \quad X_{[\alpha\beta]} := \frac{1}{2}(X_{\alpha\beta} - X_{\beta\alpha}). \quad \square$$

Замечание. Выделение следа у тензоров со всеми ковариантными или контравариантными индексами невозможно без наличия метрики. $\square$

Пример 2.7.8. Тензорное поле X_α^β типа (1,1) также приводимо, поскольку у него можно выделить след $\text{tr } X$ и бесследовую часть Y_α^β ($Y_\alpha^\alpha = 0$):

$$X_\alpha^\beta = Y_\alpha^\beta + \frac{1}{n} \delta_\alpha^\beta \text{tr } X, \quad (2.56)$$

где

$$\text{tr } X := X_\alpha^\alpha, \quad Y_\alpha^\beta := X_\alpha^\beta - \frac{1}{n} \delta_\alpha^\beta \text{tr } X. \quad \square$$

Пример 2.7.9. На четырехмерном ориентируемом многообразии существует дополнительная возможность. Если задана метрика, то существует полностью антисимметричный тензор четвертого ранга $\varepsilon_{\alpha\beta\gamma\delta}$ (см. следующий раздел), и антисимметричный тензор второго ранга можно представить в виде суммы двух компонент, неприводимых относительно преобразований координат с положительным Якобианом (сохраняющих ориентацию),

$$X_{\alpha\beta} = Y_{\alpha\beta} + Z_{\alpha\beta},$$

где первое слагаемое удовлетворяет условию *самодуальности* $Y_{\alpha\beta} = \varepsilon_{\alpha\beta\gamma\delta} Y^{\gamma\delta}$, а второе – *антисамодуальности* $Z_{\alpha\beta} = -\varepsilon_{\alpha\beta\gamma\delta} Z^{\gamma\delta}$. Это разложение взаимно однозначно:

$$Y_{\alpha\beta} = \frac{X_{\alpha\beta} + \varepsilon_{\alpha\beta\gamma\delta} X^{\gamma\delta}}{2}, \quad Z_{\alpha\beta} = \frac{X_{\alpha\beta} - \varepsilon_{\alpha\beta\gamma\delta} X^{\gamma\delta}}{2} \quad \square$$

Тензорные поля ранга три и выше, в общем случае, приводимы. Если на многообразии не задано никаких других объектов, кроме тензорного поля, то разложение на неприводимые компоненты может осуществляться только с помощью взятия следа, симметризации или антисимметризации по индексам.

Замечание. При проведении вычислений с тензорными полями важно иметь ввиду следующее обстоятельство. Если тензорное поле приводимо, то его разложение на неприводимые компоненты в большинстве случаев упрощает вычисления и проясняет математическую структуру модели. $\square$

Пример 2.7.10. *Символ Кронекера*, компоненты которого в каждой карте многообразия $\mathbb{M}$ составляют n -мерную единичную матрицу,

$$\delta_\beta^\alpha = \begin{cases} 1, & \alpha = \beta, \\ 0, & \alpha \neq \beta, \end{cases} \quad (2.57)$$

и имеют один верхний и один нижний индекс, определяет тензорное поле типа (1,1). Он инвариантен относительно преобразований координат

$$\delta_{\beta'}^{\alpha'} = \frac{\partial x^\beta}{\partial x^{\beta'}} \delta_\beta^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha} = \frac{\partial x^\beta}{\partial x^{\beta'}} \frac{\partial x^{\alpha'}}{\partial x^\beta} = \begin{cases} 1, & \alpha' = \beta', \\ 0, & \alpha' \neq \beta', \end{cases}$$

т.к. верхний и нижний индексы преобразуется с помощью взаимно обратных матриц. Символ Кронекера представляет собой исключение в двух отношениях. Во-первых, он инвариантен относительно преобразований координат и, во-вторых, его индексы можно писать один под другим, поскольку подъем и опускание индексов с помощью метрики приводит к симметричным тензорам. $\square$

Тензорное поле $T_\alpha^\beta \in \mathcal{T}_1^1(\mathbb{M})$ типа $(1, 1)$ можно рассматривать, как линейный оператор (эндоморфизм), действующий в пространстве векторов $\mathcal{T}^1(\mathbb{M}) = \mathcal{X}(\mathbb{M})$ и 1-форм $\mathcal{T}_1(\mathbb{M}) = \Lambda_1(\mathbb{M})$. В компонентах действие оператора задается правилами:

$$A'_\alpha = T_\alpha^\beta A_\beta, \quad X'^\alpha = X^\beta T_\beta^\alpha.$$

Действие оператора T_α^β естественным образом распространяется на тензоры произвольного типа.

Пример 2.7.11. Символ Кронекера представляет собой тождественный оператор. $\square$

Пример 2.7.12. Проекционные операторы (4.25) являются тензорными полями типа $(1, 1)$. $\square$

На многообразии $\mathbb{M}$ можно также определить *тензорные плотности* степени $p \in \mathbb{Z}$ и ранга (r, s) , если при преобразовании координат все их компоненты умножить на якобиан преобразования в степени p , как и в случае скалярных полей (2.7). Например, тензорная плотность типа $(1, 1)$ и степени p преобразуется по правилу

$$X_{\alpha'}^{\beta'} = J^p \frac{\partial x^\alpha}{\partial x^{\alpha'}} X_\alpha^\beta \frac{\partial x^{\beta'}}{\partial x^\beta}. \quad (2.58)$$

В каждой точке многообразия тензорные плотности фиксированного типа и степени образуют векторное пространство над полем вещественных чисел и модуль над алгеброй гладких функций. По аналогии с тензорным произведением тензоров можно ввести тензорное произведение плотностей, которое двум плотностям типа (r_1, s_1) , (r_2, s_2) и степеней p_1 и p_2 ставит в соответствие тензорную плотность типа $(r_1 + r_2, s_1 + s_2)$ и степени $p_1 + p_2$. Множество тензорных плотностей и всех их линейных комбинаций в фиксированной точке образует некоммутативную ассоциативную алгебру над полем вещественных чисел. Эта алгебра имеет естественную градуировку, как прямая сумма тензорных плотностей фиксированного типа и степени.

Так же, как и в случае скалярных плотностей тензорные плотности $X_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}$ степени p можно представить в виде

$$X_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r} = \sqrt{|g|}^{-p} Y_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r},$$

где $\sqrt{|g|} = \det e_\alpha^a$ — определитель репера, а $Y_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}$ — тензорное поле того же типа, что и исходная тензорная плотность.

Множество тензорных полей можно интерпретировать, как множество полилинейных отображений.

Теорема 2.7.1. *Множество ковариантных тензорных полей $\mathcal{T}_s(\mathbb{M})$ можно рассматривать, как s линейное отображение ($\mathcal{C}(\mathbb{M})$ -модуль) из $\underbrace{\mathcal{X} \times \dots \times \mathcal{X}}_s$ в алгебру*

непрерывных функций $\mathcal{C}(\mathbb{M})$ такое, что

$$K(f_1 X_1, \dots, f_s X_s) = f_1 \dots f_s K(X_1, \dots, X_s), \quad K(x) \in \mathcal{T}_s(\mathbb{M}),$$

для всех $f_i \in \mathcal{C}(\mathbb{M})$ и $X_i \in \mathcal{X}(\mathbb{M})$. Обратно, каждое такое отображение можно рассматривать, как тензорное поле типа $(0, s)$.

Доказательство. См., например, [48], предложение 3.1. $\square$

Аналогично можно интерпретировать тензоры произвольного типа (r, s) , как множество всех $r + s$ линейных отображений:

$$\mathcal{T}_s^r(\mathbb{M}) : \underbrace{\mathcal{X} \times \dots \times \mathcal{X}}_s \times \underbrace{\Lambda_1 \times \dots \times \Lambda_1}_r \rightarrow \mathcal{C}(\mathbb{M}). \quad (2.59)$$

Поскольку между тензорными полями и полилинейными отображениями существует взаимно однозначное соответствие, то некоторые авторы принимают эти отображения в качестве глобального определения тензорных полей.

2.8 Полностью антисимметричные тензоры

В настоящем разделе мы рассмотрим полностью антисимметричные тензоры, которые играют очень важную роль в различных приложениях дифференциальной геометрии.

Определение. Рассмотрим тензорные поля типа $(r, 0)$ или $(0, r)$ при $r \leq n$ на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, компоненты которых антисимметричны относительно перестановки любой пары индексов. В инвариантном виде условие антисимметричности для ковариантных тензоров записывается в виде:

$$T(X_1, \dots, X_i, \dots, X_j, \dots, X_r) = -T(X_1, \dots, X_j, \dots, X_i, \dots, X_r), \quad 1 \leq i < j \leq r,$$

где $T \in \mathcal{T}_r(\mathbb{M})$, $X_1, \dots, X_r \in \mathcal{X}(\mathbb{M})$. Эти тензоры неприводимы и называются *полностью антисимметричными* ковариантными тензорами ранга r . $\square$

Компонента полностью антисимметричного тензорного поля может быть отлична от нуля только в том случае, если все индексы различны, поскольку при совпадении двух или более индексов соответствующая компонента равна нулю. На многообразии размерности n не существует полностью антисимметричного тензора ранга большего, чем размерность многообразия, т.к. в этом случае по крайней мере два индекса будут совпадать.

Очевидно, что число независимых компонент полностью антисимметричного тензора ранга r равно числу выборов r различных индексов из n : $C_n^r = n!/r!(n-r)!$. В частности, полностью антисимметричный тензор $X_{\alpha_1 \dots \alpha_n}$ максимального ранга n имеет только одну независимую компоненту. Нетрудно проверить, что при преобразовании координат $x \mapsto y(x)$ полностью антисимметричный тензор типа $(0, n)$ преобразуется по закону

$$X'_{\alpha_1 \dots \alpha_n} = \frac{\partial x^{\beta_1}}{\partial y^{\alpha_1}} \dots \frac{\partial x^{\beta_n}}{\partial y^{\alpha_n}} X_{\beta_1 \dots \beta_n} = X_{\alpha_1 \dots \alpha_n} J^{-1}, \quad (2.60)$$

где $J := \det(\partial y^\alpha / \partial x^\beta)$ – якобиан преобразования координат. То есть каждая компонента полностью антисимметричного ковариантного тензора максимального ранга умножается на якобиан преобразования в минус первой степени, и ее фиксированную компоненту можно рассматривать как скалярную плотность степени -1 .

В каждой карте $(\mathbb{U}, \varphi)$ можно построить полностью антисимметричный объект, компоненты которого равны по модулю единице

$$\hat{\varepsilon}_{\alpha_1 \dots \alpha_n} := \operatorname{sgn} \sigma(\alpha_1 \dots \alpha_n), \quad \hat{\varepsilon}_{1 \dots n} = 1, \quad (2.61)$$

где $\text{sgn } \sigma$ – знак перестановки σ , который равен $+1$ или -1 , если для получения последовательности индексов $\alpha_1, \dots, \alpha_n$ из последовательности натуральных чисел $1, \dots, n$ необходимо переставить, соответственно, четное и нечетное число пар индексов. Объект $\hat{\varepsilon}_{\alpha_1 \dots \alpha_n}$ не может быть тензором, т.к. в общем случае якобиан преобразования J отличен от единицы. Из закона преобразования тензорных плотностей (2.7) следует, что каждую фиксированную компоненту $\hat{\varepsilon}_{\alpha_1 \dots \alpha_n}$ можно рассматривать, как скалярную тензорную плотность степени -1 . Поскольку компоненты антисимметричной тензорной плотности постоянны, то в произвольной системе координат справедливо равенство

$$\partial_\alpha \hat{\varepsilon}_{\alpha_1 \dots \alpha_n} = 0. \quad (2.62)$$

В (псевдо)римановом пространстве при наличии метрики $g_{\alpha\beta}$ (см. раздел 4.1) можно построить полностью антисимметричный тензор

$$\varepsilon_{\alpha_1 \dots \alpha_n} = \sqrt{|g|} \hat{\varepsilon}_{\alpha_1 \dots \alpha_n}, \quad (2.63)$$

где введено сокращенное обозначение для определителя метрики, которое будет часто использоваться в дальнейшем

$$g := \det g_{\alpha\beta}. \quad (2.64)$$

Полностью антисимметричный тензор (2.63) преобразуется по стандартному закону (2.60).

Замечание. Здесь и в дальнейшем мы примем следующее обозначение: шляпка над символом означает, что рассматривается тензорная плотность, а не тензор. $\square$

Полностью антисимметричные тензоры типа $(k, 0)$ со всеми контравариантными индексами называются *поливекторами*. Поливектор максимального ранга n имеет одну независимую компоненту, которая при преобразовании координат преобразуется по правилу

$$X'^{\alpha_1 \dots \alpha_n} = X^{\alpha_1 \dots \alpha_n} J, \quad (2.65)$$

В пространстве контравариантных тензоров можно ввести полностью антисимметричную тензорную плотность степени 1 с компонентами, равными по модулю единице, аналогично тому, как это было сделано для ковариантных тензоров. Мы положим

$$\hat{\varepsilon}^{\alpha_1 \dots \alpha_n} := \hat{\varepsilon}_{\alpha_1 \dots \alpha_n} \text{sgn}, \quad (2.66)$$

где множитель sgn зависит от того задана ли на многообразии метрика или нет

$$\text{sgn} := \begin{cases} 1, & \text{если метрика не задана,} \\ \text{sgn}(\det g_{\alpha\beta}), & \text{если метрика задана.} \end{cases} \quad (2.67)$$

Компоненты этой плотности также равны по модулю единице и постоянны: $\partial_\alpha \hat{\varepsilon}^{\alpha_1 \dots \alpha_n} = 0$. При наличии метрики можно ввести полностью антисимметричный контравариантный тензор

$$\varepsilon^{\alpha_1 \dots \alpha_n} = \frac{1}{\sqrt{|g|}} \hat{\varepsilon}^{\alpha_1 \dots \alpha_n}. \quad (2.68)$$

Замечание. Отметим, что равенство (2.66) имеет смысл, несмотря на то, что индексы слева контравариантны, а справа ковариантны, т.к. компоненты тензорных плотностей $\hat{\varepsilon}$ не зависят от выбора системы координат. Появление множителя (2.67)

в определении (2.66) объясняется тем, что на (псевдо)римановом многообразии мы требуем, чтобы тензор с контравариантными индексами можно было бы получить из тензора с ковариантными индексами простым подъемом индексов:

$$\varepsilon^{\alpha_1 \dots \alpha_n} := g^{\alpha_1 \beta_1} \dots g^{\alpha_n \beta_n} \varepsilon_{\beta_1 \dots \beta_n}.$$

Отсюда следует равенство (2.66). $\square$

Наличие тензора $\varepsilon_{\alpha_1 \dots \alpha_n}$ позволяет представить компоненты произвольного ковариантного антисимметричного тензора максимального ранга в виде

$$X_{\alpha_1 \dots \alpha_n} = X^*(x) \varepsilon_{\alpha_1 \dots \alpha_n}, \quad X^* := \varepsilon^{\alpha_1 \dots \alpha_n} X_{\alpha_1 \dots \alpha_n} \operatorname{sgn}, \quad (2.69)$$

где $X^*(x)$ – некоторое (псевдо)скалярное поле, которое меняет знак при отражении одной из координатных осей.

Аналогичное представление имеет место для произвольного контравариантного тензора максимального ранга. Это означает, что полностью антисимметричные тензоры максимального ранга имеют ровно одну независимую компоненту.

2.9 Отображения многообразий

Рассмотрим отображение h многообразия $\mathbb{M}$, $\dim \mathbb{M} = m$, в многообразие $\mathbb{N}$, $\dim \mathbb{N} = n$,

$$h : \mathbb{M} \ni x \mapsto y \in \mathbb{N}. \quad (2.70)$$

Многообразие $\mathbb{N}$ мы будем называть *пространством-мишенью*. Пусть при этом отображении карта $(\mathbb{U}, \varphi)$ многообразия $\mathbb{M}$ отображается в некоторую карту $(\mathbb{V}, \phi)$ многообразия $\mathbb{N}$, $h(\mathbb{U}) \subset \mathbb{V}$. Обозначим координаты на $\mathbb{U}$ и $\mathbb{V}$ через x^α , $\alpha = 1, \dots, m$, и y^μ , $\mu = 1, \dots, n$. Тогда отображение

$$\phi \circ h \circ \varphi^{-1} : \mathbb{R}^m \supset \varphi(\mathbb{U}) \rightarrow \phi(\mathbb{V}) \subset \mathbb{R}^n \quad (2.71)$$

двух областей евклидова пространства $\varphi(\mathbb{U}) \subset \mathbb{R}^m$ и $\phi(\mathbb{V}) \subset \mathbb{R}^n$ задается n функциями от m переменных $y^\mu(x)$. При этом размерность многообразия $\mathbb{M}$ может быть меньше, равна или больше размерности $\mathbb{N}$.

Определение. Отображение гладких многообразий называется *гладким (дифференцируемым)*, если задается гладкими (дифференцируемыми) функциями (2.71) в полных атласах на $\mathbb{M}$ и $\mathbb{N}$. $\square$

В дальнейшем, если не оговорено противное, мы будем рассматривать гладкие отображения.

Определение. Дифференцируемое отображение (2.70) индуцирует линейное отображение касательных пространств

$$h_* : \mathbb{T}_x(\mathbb{M}) \ni X = X^\alpha \partial_\alpha \mapsto Y = h_* X = Y^\mu \partial_\mu \in \mathbb{T}_{h(x)}(\mathbb{N}) \quad (2.72)$$

следующим образом. Рассмотрим кривую $\gamma \subset \mathbb{M}$, проходящую через произвольную точку $p \in \mathbb{M}$ в направлении произвольного вектора $X_{(\gamma)}(p)$. Эта кривая отображается в некоторую кривую $h(\gamma)$ на $\mathbb{N}$. По определению, вектор $X_{(\gamma)}(p) \in \mathbb{T}_p(\mathbb{M})$ отображается в тот вектор $Y_{(h \circ \gamma)} \in \mathbb{T}_{h(p)}(\mathbb{N})$, который касается кривой $h(\gamma)$ в точке $h(p)$. Поскольку

вектор в точке p – это класс эквивалентности кривых, то это условие записывается в виде

$$h_*(X_{(\gamma)}(p)) = [h \circ \gamma] = Y_{(h \circ \gamma)}(h(p)), \quad X_{(\gamma)}(p) = [\gamma].$$

Теперь мы упростим обозначения, опустив индекс кривой у вектора и обозначение точки p . Поскольку по правилу дифференцирования сложной функции

$$\dot{y}^\mu(\phi \circ h \circ \gamma(t)) = \frac{\partial y^\mu}{\partial x^\alpha} \dot{x}^\alpha(\gamma(t)),$$

то вектор Y единственен и не зависит от представителя класса эквивалентности кривых, определяющих вектор X в точке $x \in \mathbb{M}$. Это отображение касательных пространств (2.72) называется *дифференциалом отображения*. $\square$

В компонентах дифференциал отображения задается матрицей Якоби:

$$Y^\mu = X^\alpha \frac{\partial y^\mu}{\partial x^\alpha}. \quad (2.73)$$

Дифференциал отображения h_* является линейным отображением:

$$\begin{aligned} h_*(X + Y) &= h_*(X) + h_*(Y), & X, Y &\in \mathcal{X}(\mathbb{M}), \\ h_*(aX) &= ah_*(X), & a &\in \mathbb{R}. \end{aligned}$$

множества всех векторных полей $\mathcal{X}(\mathbb{M})$ на $\mathbb{M}$ в множество векторных полей $\mathcal{X}(h(\mathbb{M}))$ на образе $h(\mathbb{M}) \subset \mathbb{N}$, который может не совпадать со всем $\mathbb{N}$.

Предложение 2.9.1. *Дифференциал отображения согласован со структурой алгебры Ли в пространствах векторных полей, т.е.*

$$h_*[X, Y] = [h_*X, h_*Y]. \quad (2.74)$$

Доказательство. Простая проверка. $\square$

Рассмотрим два отображения $\mathbb{M} \xrightarrow{h} \mathbb{N} \xrightarrow{g} \mathbb{P}$. Если обозначить координаты на многообразиях $\mathbb{M}$, $\mathbb{N}$ и $\mathbb{P}$, соответственно, через x , y и z , то по правилу дифференцирования сложной функции в соответствующих областях определения справедливо равенство

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \frac{\partial y}{\partial x}.$$

Здесь, для краткости, мы опустили индексы. Отсюда следует, что дифференциал произведения равен произведению дифференциалов каждого отображения

$$(g \circ h)_* = g_* \circ h_*. \quad (2.75)$$

В координатах мы имеем обычное произведение матриц Якоби. Это уравнение говорит о том, что прямое отображение $g \circ h : \mathbb{M} \rightarrow \mathbb{P}$ не зависит от выбора промежуточного многообразия $\mathbb{N}$: $\mathbb{M} \rightarrow \mathbb{N} \rightarrow \mathbb{P}$, т.е. диаграмма

$$\begin{array}{ccc} \mathbb{T}_x(\mathbb{M}) & \xrightarrow{h_*} & \mathbb{T}_{h(x)}(\mathbb{N}) \\ & \searrow g_* \circ h_* & \downarrow g_* \\ & & \mathbb{T}_{g(h(x))}(\mathbb{P}) \end{array}$$

коммутативна для всех точек $x \in \mathbb{M}$.

Замечание. Поскольку сами многообразия в общем случае не являются векторными пространствами, то дифференциал отображения не имеет смысла производной по Фреше отображения h . $\square$

Дифференциал отображения естественным образом обобщается на случай произвольных тензорных полей типа $(r, 0)$, имеющих только контравариантные индексы, и обозначается $(h_*)^r$. При этом каждый контравариантный индекс суммируется с матрицей Якоби отображения.

Пример 2.9.1. Для компонент контравариантных тензоров второго ранга имеем следующий закон преобразования

$$Y^{\mu\nu} = X^{\alpha\beta} \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial y^\nu}{\partial x^\beta},$$

где $X = X^{\alpha\beta}(x) \partial_\alpha \otimes \partial_\beta \in \mathcal{T}^2(\mathbb{M})$ и $Y = Y^{\mu\nu}(y) \partial_\mu \otimes \partial_\nu \in \mathcal{T}^2(\mathbb{N})$ $\square$

Замечание. Если векторное поле рассматривается, как дифференцирование в алгебре функций, то дифференциал отображения определяется следующей формулой

$$(h_*X)f = X(f \circ h), \quad \forall f \in \mathcal{C}^1(\mathbb{N}), \quad X \in \mathcal{X}(\mathbb{M}). \quad (2.76)$$

При таком определении дифференциала отображения, его линейность очевидна. Это еще раз говорит о том, что алгебраические определения удобнее в тех случаях, когда исследуются алгебраические свойства. $\square$

Определение. Инъективное отображение многообразий $h : \mathbb{M} \rightarrow \mathbb{N}$ индуцирует *возврат отображения* в кокасательных пространствах, который мы обозначим тем же символом, но со звездочкой вверх:

$$h^* : \mathcal{T}_{h(x)}^*(\mathbb{N}) \ni B \mapsto A = h^*B \in \mathcal{T}_x^*(\mathbb{M}). \quad (2.77)$$

Возврат отображения h^* дуален (сопряжен) к дифференциалу отображения h_* и определяется следующим равенством

$$A(X) = B(Y) \quad \text{или} \quad h^*B(X) = B(h_*X), \quad (2.78)$$

где $Y = h_*X$. $\square$

Замечание. Возврат отображения действует в сторону, обратную самому отображению h . $\square$

При определении возврата отображения h^* мы требуем, чтобы исходное отображение h было инъективным. В противном случае прообраз $h^{-1}(h(x))$ для некоторого $x \in \mathbb{M}$ состоит не из одного элемента, и возврат отображения не определен. Возврат отображения определен не на всех формах из $\mathcal{T}_1(\mathbb{N})$, а только на формах из образа инъективного отображения $\mathcal{T}_1(h(\mathbb{M}))$. Возврат отображения h^* будет определен на множестве всех форм $\mathcal{T}_1(\mathbb{N})$ тогда и только тогда, когда отображение h биективно.

В компонентах возврат отображения записывается в виде

$$A_\alpha = \frac{\partial y^\mu}{\partial x^\alpha} B_\mu, \quad (2.79)$$

т.е. так же, как и дифференциал отображения, определяется матрицей Якоби.

Возврат отображения естественным образом обобщается на тензорные поля из $\mathcal{T}_s(\mathbb{N})$ типа $(0, s)$, имеющие только ковариантные индексы, и обозначается $(h^*)^s$.

Пример 2.9.2. Для компонент ковариантных тензоров второго ранга имеем следующий закон преобразования

$$X_{\alpha\beta} = \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial y^\nu}{\partial x^\beta} Y_{\mu\nu},$$

где $X = X_{\alpha\beta}(x)dx^\alpha \otimes dx^\beta \in \mathcal{T}_2(\mathbb{M})$ и $Y = Y_{\mu\nu}(y)dy^\mu \otimes dy^\nu \in \mathcal{T}_2(\mathbb{N})$ □

Из правила дифференцирования сложных функций следует, что возврат произведения отображений равен произведению возвратов отображений, но взятых в обратном порядке

$$(g \circ f)^* = f^* \circ g^*. \quad (2.80)$$

Для тензоров смешанного типа в общем случае не существует индуцированного отображения, поскольку дифференциал и возврат отображения действуют в разные стороны.

Определение. Назовем *рангом отображения* $h : \mathbb{M} \rightarrow \mathbb{N}$ ранг соответствующей матрицы Якоби

$$\text{rank } h := \text{rank } \frac{\partial y^\mu}{\partial x^\alpha}.$$

Если $\text{rank } h = m$ во всех точках многообразия $\mathbb{M}$, то отображение h называется *невырожденным*. Для этого необходимо, чтобы $m \leq n$.

Рассмотрим отображение (2.70) двух многообразий одинаковой размерности $\dim \mathbb{M} = \dim \mathbb{N}$ и одного класса дифференцируемости $\mathcal{C}^k$. Если отображение h биективно, и оба отображения h и h^{-1} в координатах задаются функциями класса $\mathcal{C}^k$, то отображение h называется *диффеоморфизмом*. □

Пример 2.9.3. Пусть $(\mathbb{U}, \varphi)$ карта на дифференцируемом многообразии $\mathbb{M}$. Тогда отображение $\varphi : \mathbb{U} \rightarrow \varphi(\mathbb{U}) \subset \mathbb{R}^n$ является диффеоморфизмом. □

Пример 2.9.4. Преобразование координат $\mathbb{U}_0 \rightarrow \mathbb{U}'_0$ в теореме 1.5.1 является диффеоморфизмом. □

Композиция двух диффеоморфизмов снова будет диффеоморфизмом. Таким образом, диффеоморфизм является отношением эквивалентности в категории дифференцируемых многообразий. С точки зрения дифференциальной геометрии два диффеоморфных между собой многообразия можно рассматривать, как одно многообразие, заданное в различных координатах, поэтому мы пишем $\mathbb{M} \approx \mathbb{N}^1$. В этом случае индуцированные отображения касательных пространств представляют собой не что иное, как правила преобразования тензорных полей при преобразовании координат.

Замечание. Не имеет смысла рассматривать степень гладкости отображения, которая превышает гладкость самих многообразий. □

Замечание. Любой диффеоморфизм представляет собой гомеоморфизм многообразий, рассматриваемых, как топологические пространства, т.к. функции h и h^{-1} непрерывны. Обратное утверждение неверно. Как отмечено в разделе 2.1, на семи-мерной сфере можно задать несколько различных дифференцируемых структур. То есть гомеоморфные многообразия могут быть недиффеоморфны. □

¹Мы не используем обычный знак равенства, т.к. множества точек $\mathbb{M}$ и $\mathbb{N}$ могут отличаться по другим признакам, например, по наличию групповой структуры.

Множество невырожденных гладких отображений $h : \mathbb{M} \rightarrow \mathbb{M}$ многообразия $\mathbb{M}$ в себя образует группу преобразований многообразия, которая обозначается $\text{diff}(\mathbb{M})$. Если многообразие тривиально (т.е. покрывается одной картой), то это просто группа преобразований координат $\text{diff}(\mathbb{R}^n)$.

Определение. Пусть для отображения (2.70) $m < n$. Если отображение касательного пространства $T_x(\mathbb{M})$ для всех точек $x \in \mathbb{M}$ на его образ в касательном пространстве $T_{h(x)}(\mathbb{N})$ является взаимно однозначным, т.е. $\text{rank } h = \dim \mathbb{M}$, то отображение h называется *погружением*. По определению, любое невырожденное отображение задает погружение. При погружении само отображение h может не быть взаимно однозначным. Размерность многообразия $\mathbb{M}$ не может превосходить размерности многообразия $\mathbb{N}$, т.к. в этом случае не может быть взаимной однозначности дифференциала отображения. Если само отображение h на его образ $h(\mathbb{M})$ и его дифференциал являются взаимно однозначными, то отображение h называется *вложением*. В дальнейшем вложение многообразий мы будем обозначать специальным символом: $\mathbb{M} \hookrightarrow \mathbb{N}$. $\square$

Конечно, каждое вложение одновременно является и погружением.

Пример 2.9.5. Отображение окружности S^1 на плоскость $\mathbb{R}^2$ в виде восьмерки является погружением, но не вложением, см. рис. 2.10a. В то же время отображение окружности в гладкую замкнутую кривую без самопересечений представляет собой вложение, рис. 2.10b. $\square$

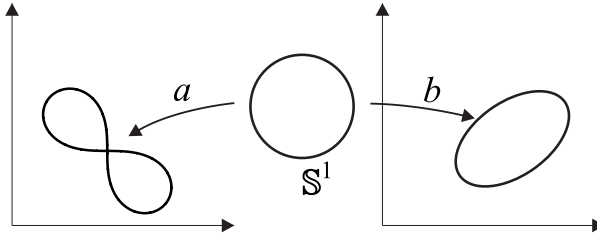


Рис. 2.10: Отображение окружности на плоскость в виде восьмерки является погружением, но не вложением (a). Вложение окружности в плоскость в виде гладкой замкнутой кривой без самопересечений (b).

Если $m = n$, отображение h биективно и дифференцируемо, а образ $h(\mathbb{M})$ совпадает со всем $\mathbb{N}$, то точки многообразий $\mathbb{M}$ и $\mathbb{N}$ можно отождествить и рассматривать вложение $\mathbb{M} \xrightarrow{h} \mathbb{N} \approx \mathbb{M}$ как диффеоморфизм. Нетривиальная ситуация может возникнуть, если многообразие $\mathbb{M}$ отображается только на часть $\mathbb{N}$. Тогда мы имеем диффеоморфизм между многообразием $\mathbb{M}$ и его образом $h(\mathbb{M})$.

2.10 Подмногообразия

Важным классом отображений многообразий являются подмногообразия.

Определение. Пусть $f : \mathbb{M} \hookrightarrow \mathbb{N}$ – вложение многообразия $\mathbb{M}$ в $\mathbb{N}$, размерностей m и n , при этом $m \leq n$, тогда пара $(f, \mathbb{M})$ называется *подмногообразием* в $\mathbb{N}$.

В дальнейшем, если не оговорено противное, под вложением мы понимаем гладкое вложение, когда отображение f определяется гладкими функциями.

Пример 2.10.1 (Лемниската Бернулли). Кривая четвертого порядка, заданная в декартовых координатах x, y на евклидовой плоскости $\mathbb{R}^2$ уравнением

$$(x^2 + y^2)^2 - 2a^2(x^2 - y^2) = 0, \quad a \in \mathbb{R}, \quad (2.81)$$

называется *лемнискатой Бернулли*, рис.2.11. Эта кривая обладает следующим свойством. Произведение расстояний от произвольной точки лемнискаты до двух заданных точек (фокусов) постоянно и равно квадрату половины расстояния между фокусами. Верно и обратное утверждение. На плоскости можно выбрать такую систему декартовых координат, что произвольная кривая с данным свойством будет задана уравнением (2.81).

В полярных координатах r, φ уравнение лемнискаты имеет вид:

$$r^2 = 2a^2 \cos(2\varphi).$$

Лемниската Бернулли не является одномерным многообразием, т.к. имеет точку самопересечения – начало координат. Точка самопересечения имеет касательные $y = \pm x$ и является точкой перегиба. Площадь каждой петли равна a^2 .

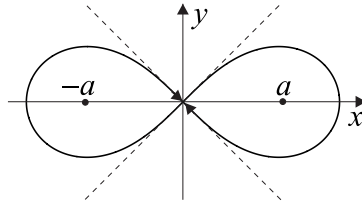


Рис. 2.11: Лемниската Бернулли.

Рассмотрим вложение прямой в евклидову плоскость

$$\gamma : \mathbb{R} \ni t \mapsto (x, y) \in \mathbb{R}^2,$$

гладко отобразив прямую в точки лемнискаты, как показано на рис.2.11. При этом мы считаем, что точка $t = 0$ отображается в начало координат $(0, 0) \in \mathbb{R}^2$. Стрелки показывают, что концы прямой $t = \pm\infty$ при вложении стремятся к началу координат $(0, 0) \in \mathbb{R}^2$. Это – гладкое вложение. Однако топология прямой не совпадает с топологией, которая индуцирована вложением. Действительно, любая последовательность $\{t_k\}$ такая, что $\lim_{k \rightarrow \infty} t_k = \pm\infty$ в естественной топологии на прямой, будет сходиться к точке $t = 0$ в индуцированной топологии. Конечно, вместо лемнискаты Бернулли можно было бы выбрать произвольную гладкую кривую в виде восьмерки. $\square$

Пример 2.10.2 (Всюду плотная обмотка тора). Двумерный тор $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ можно рассматривать, как двумерное многообразие, полученное путем отождествления противоположных сторон единичного квадрата на евклидовой плоскости. Тем самым точка тора задается упорядоченной парой чисел (x, y) , каждое из которых определено по модулю один, $x \sim x + 1$ и $y \sim y + 1$. Зафиксируем пару чисел a и $b \neq 0$ и рассмотрим отображение вещественной прямой в тор

$$f : \mathbb{R} \ni t \mapsto (at \bmod 1, bt \bmod 1) \in \mathbb{T}^2.$$

Если отношение a/b иррационально, то отображение f является вложением, причем образ прямой $f(\mathbb{R})$ всюду плотен в $\mathbb{T}^2$ (*всюду плотная обмотка тора*). Если отношение a/b рационально, то f представляет собой погружение, и его образ диффеоморфен окружности. Согласно данному выше определению всюду плотная обмотка тора

является одномерным подмногообразием на торе. Для всюду плотной обмотки тора индуцированная топология на образе прямой не совпадает с естественной топологией на $\mathbb{R}$. $\square$

При вложении дифференцируемая структура и топология, заданные на $\mathbb{M}$, естественным образом переносятся на образ $f(\mathbb{M})$, поскольку отображение f инъективно и дифференцируемо. Тем самым многообразие $\mathbb{M}$ диффеоморфно своему образу $f(\mathbb{M})$. С другой стороны, поскольку образ $f(\mathbb{M})$ является некоторым подмножеством в пространстве-мишени $\mathbb{N}$, то на нем определяется индуцированная из $\mathbb{N}$ топология. Рассмотренные выше примеры вложения прямой в плоскость и всюду плотной обмотки тора показывают, что топология на $f(\mathbb{M})$, наследуемая из $\mathbb{M}$ при вложении, совсем не обязательно совпадает с топологией, индуцированной из $\mathbb{N}$. В общем случае топология из $\mathbb{M}$ является более тонкой, чем топология, индуцированная из $\mathbb{N}$. Это наблюдение мотивирует следующее

Определение. Пусть вложение $f : \mathbb{M} \hookrightarrow \mathbb{N}$ является также и гомеоморфизмом. Тогда пара $(f, \mathbb{M})$ называется *регулярным подмногообразием*, а f – *регулярным вложением* $\mathbb{M}$ в $\mathbb{N}$. $\square$

Доказательство следующих трех теорем приведено в [49].

Теорема 2.10.1. Пусть $(f, \mathbb{M})$ является t -мерным подмногообразием в n -мерном многообразии $\mathbb{N}$, причем $t < n$. Для того, чтобы пара $(f, \mathbb{M})$ была регулярным подмногообразием в $\mathbb{N}$, необходимо и достаточно, чтобы оно было замкнутым подмногообразием некоторого открытого подмногообразия в $\mathbb{N}$.

Теорема 2.10.2. Пара $(f, \mathbb{M})$ является регулярным подмногообразием в $\mathbb{N}$ тогда и только тогда, когда для каждой точки $x \in \mathbb{M}$ существует система координат y^α на $\mathbb{U}_{f(x)} \subset \mathbb{N}$, $y^\alpha(f(x)) = 0$, такая, что пересечение $f(\mathbb{M}) \cap \mathbb{U}_{f(x)}$ определяется системой, состоящей из $n - t$ уравнений:

$$y^{m+1} = y^{m+2} = \dots = y^n = 0.$$

Теорема 2.10.3. Пусть пара $(f, \mathbb{M})$ является подмногообразием в $\mathbb{N}$. Если многообразие $\mathbb{M}$ компактно, тогда вложение $f : \mathbb{M} \rightarrow \mathbb{N}$ регулярно.

В дальнейшем под вложением мы будем понимать регулярное вложение, для краткости не оговаривая этого явно. В этом случае мы будем отождествлять подмногообразие $\mathbb{M}$ и его образ $f(\mathbb{M})$ и будем писать $\mathbb{M} \subset \mathbb{N}$.

Определение. Пусть $\mathbb{U}$ является открытым подмножеством $\mathbb{N}$. Если ограничить гладкую структуру $\mathbb{N}$ на $\mathbb{U}$, то $\mathbb{U}$ становится многообразием той же размерности $\dim \mathbb{U} = \dim \mathbb{N}$. Тогда вложение $f : \mathbb{U} \hookrightarrow \mathbb{N}$ осуществляется тождественным отображением, а подмногообразие $\mathbb{U} \subset \mathbb{N}$ называется *открытым подмногообразием* в $\mathbb{N}$.

Пусть $f(\mathbb{M})$ является подмногообразием в $\mathbb{N}$ таким, что выполнены два условия:

- 1) $f(\mathbb{M})$ – замкнутое подмножество в $\mathbb{N}$;
- 2) для каждой точки $y \in f(\mathbb{M})$ существует такая система координат y^α на $\mathbb{U}_y \subset \mathbb{N}$, что пересечение $f(\mathbb{M}) \cap \mathbb{U}_y$ задается уравнениями $y^{m+1} = y^{m+2} = \dots = y^n = 0$, где $m = \dim \mathbb{M}$.

Тогда подмногообразие $f(\mathbb{M})$ называется *замкнутым подмногообразием* в $\mathbb{N}$. Размерность замкнутого подмногообразия $f(\mathbb{M})$ всегда меньше размерности $\mathbb{N}$. $\square$

Замечание. Если подмногообразие (без края) $M \subset \mathbb{N}$ имеет ту же размерность, что и $\mathbb{N}$, то оно будет обязательно открытым подмногообразием, т.к. в противном случае оно имело бы край. Теорема 2.10.2 утверждает, что любое подмногообразие $M \subset \mathbb{N}$ меньшей размерности $1 < m < n$ является замкнутым подмногообразием. $\square$

Мы видим, что подмногообразия меньшей размерности $1 < m < n$ в евклидовом пространстве $M \subset \mathbb{R}^n$ можно задавать с помощью системы алгебраических уравнений на координаты, покрывающие все $\mathbb{R}^n$. Вернее, алгебраические уравнения задают только множества точек. Чтобы превратить это множество в многообразие мы предполагаем, что топология на M задается вложением $f : M \hookrightarrow \mathbb{R}^n$. Тогда подмногообразие определяется системой уравнений по крайней мере локально. Обратное утверждение неверно. Не каждая система алгебраических уравнений на координаты в евклидовом пространстве $\mathbb{R}^n$ определяет некоторое подмногообразие.

Пример 2.10.3. Рассмотрим уравнение $x^2 = y^2$ на евклидовой плоскости $\mathbb{R}^2$. Оно определяет две пересекающиеся прямые, которые не являются многообразием. $\square$

Приведем критерий, который часто используется в приложениях.

Теорема 2.10.4. Для того, чтобы подмножество $M \subset \mathbb{R}^n$, $\dim M = m \leq n$, было подмногообразием класса C^k , необходимо и достаточно, чтобы для каждой точки $x \in M$ существовала окрестность $U \subset M$ этой точки и система $n - m$ функций $f^M(x) \in C^k(U)$, $M = m + 1, \dots, n$, определенных на U , и обладающих следующими свойствами:

- 1) функции f^M функционально независимы;
- 2) пересечение $U \cap M$ в точности определяется системой уравнений

$$f^M = 0, \quad M = m + 1, \dots, n. \quad (2.82)$$

Доказательство. См., например, [22], теорема 33₂. $\square$

Заметим, что если $m = n$, то уравнений не будет, а M будет некоторым открытым подмножеством в M . При $m = 0$ число уравнений равно размерности многообразия. Тогда, если в точке $x \in \mathbb{R}^n$ система уравнений (2.82) удовлетворена, то в некоторой окрестности точки x других решений кроме самой точки не будет. В этом случае подмногообразие M имеет размерность 0 и представляет собой множество корней системы уравнений $f^M = 0$.

Определение. Подмногообразие (f, M) размерности $n - 1$ в $\mathbb{N}$ называется *гиперповерхностью* в $\mathbb{N}$. $\square$

Напомним, что функции f^M функционально независимы в области U тогда и только тогда, когда $n \times (n - m)$ матрица, составленная из частных производных

$$(\partial_\alpha f^M) = \begin{pmatrix} \partial_1 f^{m+1} & \partial_1 f^{m+2} & \dots & \partial_1 f^n \\ \partial_2 f^{m+1} & \partial_2 f^{m+2} & \dots & \partial_2 f^n \\ \vdots & \vdots & \vdots & \vdots \\ \partial_n f^{m+1} & \partial_n f^{m+2} & \dots & \partial_n f^n \end{pmatrix}, \quad (2.83)$$

имеет постоянный ранг $n - m$ во всех точках U .

Замечание. Забегая вперед, переформулируем условие 1) в теореме 2.10.4 в терминах 1-форм. Поопределению, 1-формы $df^M = dx^\alpha \partial_\alpha f^M$, заданные на области $\mathbb{U}$, называются *линейно независимыми* в области $\mathbb{U}$, если в каждой точке этой области ранг матрицы (2.83) равен $n - m$. Поэтому условие функциональной независимости функций можно заменить условием линейной независимости 1-форм df^M . $\square$

Если подмногообразие $\mathbb{M} \subset \mathbb{R}^n$ задано с помощью $n - m$ алгебраических уравнений (2.82), то множество функций f^M , в окрестности произвольной точки $x \in \mathbb{M}$ всегда можно дополнить m функциями $g^1(x), \dots, g^m(x)$ такими, что весь набор функций $g^1, \dots, g^m, f^{m+1}, \dots, f^n$ будет функционально независим. В этом случае можно перейти в новую систему координат

$$\begin{aligned} \{y^\alpha\} = \{y^A, y^M\} = \{g^A, f^M\} = (g^1, \dots, g^m, f^{m+1}, \dots, f^n), \\ A = 1, \dots, m, \quad M = m + 1, \dots, n, \end{aligned} \quad (2.84)$$

в которой подмногообразию $\mathbb{M}$ будет задаваться особенно просто $y^M := f^M = 0$. При этом функции $y^A = g^A$ образуют локальную систему координат на подмногообразии $\mathbb{M} \subset \mathbb{R}^n$. Конечно, выбор координат g^A на $\mathbb{M}$ является неоднозначным. Отметим, что $\{y^\alpha\}$ – это те координаты, которые фигурируют в теореме 2.10.2.

Пример 2.10.4. Двумерная сфера $\mathbb{S}_r^2$ радиуса r с центром в начале координат является замкнутым подмногообразием трехмерного евклидова пространства $\mathbb{R}^3$, которое задается одним уравнением $x^2 + y^2 + z^2 = r^2$. Она является гиперповерхностью в $\mathbb{R}^3$. В сферической системе координат уравнение сферы задается уравнением $r = \text{const}$, а угловые координаты образуют систему координат на сфере. $\square$

Функции $f^M(x)$ определяют $n - m$ точных 1-форм. В силу их линейной независимости, в каждой точке $x \in \mathbb{M}$ они определяют $(n - m)$ -мерное подпространство в кокасательном пространстве $\mathbb{T}_x^*(\mathbb{R}^n)$. Ортогональное дополнение этого подпространства является касательным к подмногообразию $\mathbb{M}$ и образовано всеми векторами $X^\alpha \partial_\alpha \in \mathbb{T}_x(\mathbb{R}^n)$, для которых выполнено условие $X^\alpha \partial_\alpha f^M = 0$. Это подпространство касательного пространства имеет размерность m , такую же, как и само подмногообразие. Нетрудно проверить, что, если два векторных поля X_1 и X_2 касательны к подмногообразию $\mathbb{M}$, то их коммутатор $[X_1, X_2]$ также касателен.

Ответ на обратный вопрос, в каком случае распределение векторных полей на многообразии определяет касательные подпространства к некоторому подмногообразию, дает теорема Фробениуса (см. следующий раздел).

В координатах $\{g^A, f^M\}$ дополнение касательного пространства к подмногообразию $\mathbb{M}$ до полного касательного пространства к $\mathbb{R}^n$ имеет простой геометрический смысл. Пусть в $\mathbb{R}^n$ задано векторное поле $X = X^\alpha \partial_\alpha$. Тогда в координатах (2.84), связанных с подмногообразием, оно будет иметь компоненты

$$X = (X^\alpha \partial_\alpha g^A) \partial_A + (X^\alpha \partial_\alpha f^M) \partial_M.$$

При этом первое слагаемое лежит в касательном пространстве $\mathbb{T}(\mathbb{M})$.

В дальнейшем нам понадобится следующее

Определение. Пусть $f \in C^\infty(\mathbb{N})$ и $\mathbb{M} \subset \mathbb{N}$ – подмногообразие, тогда отображение

$$f|_{\mathbb{M}} : \mathbb{M} \rightarrow \mathbb{R}.$$

называется *сужением* функции на подмногообразие. Функция f является *продолжением* функции g с некоторого подмногообразия $\mathbb{M}$ на все $\mathbb{N}$, если ее сужение на

$\mathbb{M}$ совпадает с g , $f|_{\mathbb{M}} = g$. Аналогично, сужением произвольного тензорного поля $T \in \mathcal{T}(\mathbb{N})$ на подмногообразие называется тензорное поле $T|_{\mathbb{M}}$, рассматриваемое только в точках $x \in \mathbb{M}$. Обратно, тензорное поле $T \in \mathcal{T}(\mathbb{N})$ является продолжением тензорного поля $K \in \mathcal{T}(\mathbb{N})$, заданного на подмногообразии $\mathbb{M}$, если его сужение совпадает с K , $T|_{\mathbb{M}} = K$. $\square$

Определение. Два подмногообразия $\mathbb{M}_1$ и $\mathbb{M}_2$ многообразия $\mathbb{N}$ называются *трансверсальными* в точке $x \in \mathbb{M}_1 \cap \mathbb{M}_2$, если касательные подпространства $T_x(\mathbb{M}_1)$ и $T_x(\mathbb{M}_2)$ порождают все касательное пространство $T_x(\mathbb{N})$, т.е. касательные подпространства $T_x(\mathbb{M}_1)$ и $T_x(\mathbb{M}_2)$ трансверсальны. $\square$

Трансверсальность подмногообразий означает, что в некоторой окрестности $\mathbb{U}$ точки $x \in \mathbb{M}_1 \cap \mathbb{M}_2$ существует такая система координат x^α , $\alpha = 1, \dots, n$, что подмногообразия задаются условиями:

$$\begin{aligned}\mathbb{M}_1 \cap \mathbb{U} &= \{x \in \mathbb{U} : x^{\dim \mathbb{M}_1 + 1} = 0, \dots, x^n = 0\}, \\ \mathbb{M}_2 \cap \mathbb{U} &= \{x \in \mathbb{U} : x^1 = 0, \dots, x^{n - \dim \mathbb{M}_2} = 0\}.\end{aligned}$$

Пример 2.10.5. Пусть $\mathbb{N} := \mathbb{M}_1 \times \mathbb{M}_2$. Тогда подмногообразия $(\mathbb{M}_1, x_2) \subset \mathbb{N}$, где $x_2 \in \mathbb{M}_2$, и $(x_1, \mathbb{M}_2) \subset \mathbb{N}$, где $x_1 \in \mathbb{M}_1$, трансверсальны в точке $x = (x_1, x_2)$. Если в окрестностях $\mathbb{U}_1 \subset \mathbb{M}_1$ и $\mathbb{U}_2 \subset \mathbb{M}_2$, где $x_1 \in \mathbb{U}_1$ и $x_2 \in \mathbb{U}_2$, заданы системы координат, $x_1 = \{x_1^\mu\}$ и $x_2 = \{x_2^\lambda\}$, то подмногообразия задаются уравнениями:

$$\begin{aligned}(\mathbb{U}_1, x_2) &= \{x \in \mathbb{U}_1 \times \mathbb{U}_2 : x_2^\lambda = \text{const}, \forall \lambda\}, \\ (x_1, \mathbb{U}_2) &= \{x \in \mathbb{U}_1 \times \mathbb{U}_2 : x_1^\mu = \text{const}, \forall \mu\}.\end{aligned}$$

Конечно, системы координат можно выбрать таким образом, что точки x_1 и x_2 будут находиться в началах систем отсчета. $\square$

Определение. Отображение $f : \mathbb{M} \rightarrow \mathbb{N}$ трансверсально к подмногообразию $\mathbb{L} \subset \mathbb{N}$ в точке $x \in f^{-1}(\mathbb{L} \cap f(\mathbb{M}))$, если образ $f_*(T_x(\mathbb{M}))$ трансверсален к $T_{f(x)}(\mathbb{L})$. Два отображения $f_1 : \mathbb{M}_1 \rightarrow \mathbb{N}$ и $f_2 : \mathbb{M}_2 \rightarrow \mathbb{N}$ трансверсальны друг к другу в точке $(x_1, x_2) \in \mathbb{M}_1 \times \mathbb{M}_2$, где $f_1(x_1) = f_2(x_2)$, если образы $f_{1*}(T_{x_1}(\mathbb{M}_1))$ и $f_{2*}(T_{x_2}(\mathbb{M}_2))$ порождают все касательное пространство $T_{f_1(x_1)}(\mathbb{N})$. $\square$

Понятия трансверсальных подмногообразий и трансверсальных отображений естественным образом сводятся друг к другу. Если $f_{1,2} : \mathbb{M}_{1,2} \hookrightarrow \mathbb{N}$ – вложения, то трансверсальность вложений эквивалентна трансверсальности подмногообразий $f_1(\mathbb{M}_1) \subset \mathbb{N}$ и $f_2(\mathbb{M}_2) \subset \mathbb{N}$.

2.11 Теорема Фробениуса

В разделе 2.6.5 было показано, что у любого дифференцируемого векторного поля X существуют интегральные кривые. Если векторное поле нигде не обращается в нуль на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, то интегральная кривая, проходящая через некоторую точку $x \in \mathbb{M}$, представляет собой одномерное подмногообразие, и в окрестности точки x существует такая система координат, что $X = \partial_1$. Поставим более общую задачу. Пусть в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ задано k векторных полей $(X_1, \dots, X_k) = \{X_\alpha\}$, $\alpha = 1, \dots, k < n$, которые линейно независимы в каждой точке

$x \in \mathbb{U}$. Возникает вопрос о том, существуют ли такие подмногообразия в $\mathbb{M}$, касательные пространства к которым в каждой точке $x \in \mathbb{U}$ совпадают с подпространством, натянутым на векторы $\{X_A\}$? Ответ на этот вопрос дает теорема Фробениуса. Чтобы ее сформулировать введем несколько новых понятий.

Определение. *Распределением $\mathcal{L}_k(\mathbb{M})$ размерности k на многообразии $\mathbb{M}$ называется сопоставление каждой точке $x \in \mathbb{M}$ k -мерного подпространства в касательном пространстве $\mathbb{L}_x(\mathbb{M}) \subset \mathbb{T}_x(\mathbb{M})$. Распределение называется *дифференцируемым*, если каждая точка $x \in \mathbb{M}$ имеет окрестность $\mathbb{U}$ и k дифференцируемых векторных полей $(X_1, \dots, X_k)$ на $\mathbb{U}$, которые образуют базис в $\mathbb{L}_x(\mathbb{M})$ для всех $x \in \mathbb{U}$. Множество векторных полей $(X_1, \dots, X_k)$ называется *локальным базисом* распределения $\mathcal{L}_k(\mathbb{M})$ на $\mathbb{U}$. Векторное поле X принадлежит распределению, если $X(x) \in \mathbb{L}_x(\mathbb{M})$ для всех $x \in \mathbb{M}$. Распределение называется *инволютивным* или *вполне интегрируемым*, если для любых векторных полей из распределения $X, Y \in \mathcal{L}_k(\mathbb{M})$ их коммутатор также принадлежит распределению: $[X, Y] \in \mathcal{L}_k(\mathbb{M})$. $\square$*

Ни одно из векторных полей X_A не может обращаться в нуль, т.к. в соответствующей точке векторные поля были бы линейно зависимы.

Инволютивность векторных полей, принадлежащих распределению, означает, что эти векторные поля образуют подалгебру Ли алгебры Ли векторных полей $\mathcal{X}(\mathbb{M})$.

Определение. Связное подмногообразие $f : \mathbb{N} \hookrightarrow \mathbb{M}$ называется *интегральным многообразием* распределения $\mathcal{L}_k(\mathbb{M})$, если $f_*(\mathbb{T}_x(\mathbb{N})) = \mathbb{L}_{f(x)}(\mathbb{M})$, для всех $x \in \mathbb{N}$. Если не существует других интегральных многообразий, содержащих $\mathbb{N}$, то $\mathbb{N}$ называется *максимальным интегральным многообразием* для распределения $\mathcal{L}_k(\mathbb{M})$. $\square$

Замечание. Поскольку интегральное многообразие – это подмногообразие и, следовательно, задается парой $(f, \mathbb{N})$, где f – вложение $\mathbb{N}$ в $\mathbb{M}$, то определение максимального интегрального многообразия требует уточнения. А именно, мы говорим, что интегральное многообразие $(f_1, \mathbb{N}_1)$ содержит интегральное многообразие $(f_2, \mathbb{N}_2)$, если $f_2(\mathbb{N}_2)$ есть подмножество в $f_1(\mathbb{N}_1)$: $f_2(\mathbb{N}_2) \subset f_1(\mathbb{N}_1)$. Из определения максимального интегрального многообразия сразу следует его единственность. $\square$

Предложение 2.11.1. Пусть $\mathcal{L}_k(\mathbb{M})$ – гладкое распределение на многообразии $\mathbb{M}$ такое, что через каждую точку $x \in \mathbb{M}$ проходит интегральное многообразие, тогда распределение $\mathcal{L}_k(\mathbb{M})$ инволютивно.

Доказательство. Пусть X, Y – два произвольных векторных поля, принадлежащих распределению $\mathcal{L}_k(\mathbb{M})$. Пусть $(f, \mathbb{N})$ – интегральное многообразие, проходящее через точку $f(x) \in \mathbb{M}$, $x \in \mathbb{N}$. Поскольку отображение f в каждой точке интегрального многообразия $x \in \mathbb{N}$ задает изоморфизм $\mathbb{T}_x(\mathbb{N}) \simeq \mathbb{L}_{f(x)}(\mathbb{M})$, то на подмногообразии $\mathbb{N}$ существуют векторные поля $\tilde{X}(x) = f_*^{-1}X(f(x))$ и $\tilde{Y} = f_*^{-1}Y(f(x))$. Поскольку коммутатор касательных векторов к $\mathbb{N}$ также касателен к $\mathbb{N}$, то

$$[X, Y](f(x)) = f_*[\tilde{X}, \tilde{Y}](x) \in \mathbb{L}_x(\mathbb{M}). \quad \square$$

Теорема 2.11.1 (Фробениус). Пусть $\mathcal{L}_k(\mathbb{M})$ – k -мерное гладкое инволютивное распределение на многообразии $\mathbb{M}$. Тогда для каждой точки $x \in \mathbb{M}$ существует интегральное многообразие распределения $\mathcal{L}_k(\mathbb{M})$, проходящее через точку x . Кроме того, существует такая система координат в некоторой окрестности точки x , что базис распределения имеет вид $(\partial_1, \dots, \partial_k)$.

Доказательство. См., например, [42], теорема 1.60. $\square$

Предложение 2.11.1 дает необходимое условие для существования интегральных многообразий, а теорема 2.11.1 – достаточное. Обе теоремы локальны, т.к. в них говорится об интегральных многообразиях, проходящих через точку многообразия.

Если распределение $\mathcal{L}_k(\mathbb{U})$ задано набором гладких линейно независимых векторных полей $(X_1, \dots, X_k) = \{X_A\}$, $A = 1, \dots, k < n$, в каждой точке $x \in \mathbb{U} \subset \mathbb{M}$, то будем писать $\mathcal{L}_k(\mathbb{U}) = \{X_A\}$. В этом случае будем говорить, что распределение задано *распределением векторных полей*. Тогда теорему Фробениуса вместе с предложением 2.11.1 можно переформулировать.

Теорема 2.11.2 (Фробениус). Пусть $\mathcal{L}_k(\mathbb{U}) = \{X_A\}$, $A = 1, \dots, k$, является k -мерным гладким распределением векторных полей в области $\mathbb{U} \subset \mathbb{M}$. Тогда для существования такой системы координат в некоторой подобласти $\mathbb{V} \subset \mathbb{U}$, что распределение имеет вид

$$\mathcal{L}_k(\mathbb{V}) = (\partial_1, \dots, \partial_k) \quad (2.85)$$

необходимо и достаточно, чтобы коммутатор векторных полей также принадлежал распределению,

$$[X_A, X_B] = f_{AB}^C X_C \in \mathbb{L}_x(\mathbb{M}), \quad \forall x \in \mathbb{U}, \quad (2.86)$$

где $f_{AB}^C(x)$ – некоторые гладкие функции на $\mathbb{U}$.

Доказательство. См., например, [49], теорема 4.4. $\square$

Замечание. Конечно, функции f_{AB}^C в теореме Фробениуса не могут быть произвольными. Из определения (2.86) следует, что они антисимметричны по нижним индексам и удовлетворяют тождествам, вытекающим из тождеств Якоби для коммутатора векторных полей. $\square$

Допустим, что на многообразии $\mathbb{M}$ существует гладкое распределение $\mathcal{L}_k(\mathbb{M})$. Тогда существование системы координат (2.85) в окрестности каждой точки $x \in \mathbb{M}$ означает следующее. Уравнения

$$x^\alpha = \text{const}, \quad \alpha = k + 1, \dots, n$$

определяют k -мерные подмногообразия, которые являются интегральными многообразиями распределения $\mathcal{L}_k(\mathbb{M})$, причем каждое векторное поле X_A касается одного из подмногообразий. Интегральные многообразия не пересекаются, т.к. в точке пересечения у нас не было бы подмногообразия. Кроме того, через каждую точку $\mathbb{M}$ проходит одно подмногообразие. Тем самым k -мерное распределение $\mathcal{L}_k(\mathbb{M})$ расслаивает область $\mathbb{V} \subset \mathbb{M}$, представляя ее в виде объединения несчетного числа k -мерных подмногообразий. Очевидно, что интегральная кривая для любого векторного поля из заданного инволютивного распределения, проходящая через точку $x \in \mathbb{M}$, целиком лежит в интегральном подмногообразии, содержащем эту точку.

Пример 2.11.1. Произвольное дифференцируемое векторное поле $X \in \mathcal{X}(\mathbb{M})$, не равное нулю ни в одной точке, задает одномерное распределение на $\mathbb{M}$. Это распределение инволютивно. Интегральными многообразиями одномерного распределения являются интегральные кривые векторного поля X . $\square$

Пример 2.11.2. Рассмотрим трехмерное евклидово пространство с выколотым началом координат, $\mathbb{R}^3 \setminus \{0\}$, в сферической системе координат r, θ, φ . Отличные от нуля векторные поля ∂_θ и ∂_φ коммутируют и, значит, находятся в инволюции. Интегральными подмногообразиями этого распределения являются сферы $r = \text{const}$ с центром в начале координат. Угловые координаты θ, φ и есть те координаты, которые фигурируют в теореме Фробениуса.

Векторные поля

$$J_1 = \frac{1}{2}(y\partial_z - z\partial_y), \quad J_2 = \frac{1}{2}(z\partial_x - x\partial_z), \quad J_3 = \frac{1}{2}(x\partial_y - y\partial_x), \quad (2.87)$$

заданные в декартовых координатах, являются генераторами алгебры $\mathfrak{so}(3)$ и находятся в инволюции. Интегральными подмногообразиями для них также являются сферы. Однако поля (2.87) не задают базис распределения, т.к. линейно зависимы:

$$xJ_1 + yJ_2 + zJ_3 = 0. \quad \square$$

Сформулируем глобальный вариант теоремы Фробениуса.

Теорема 2.11.3 (Фробениус). Пусть $\mathcal{L}_k(\mathbb{M})$ – гладкое инволютивное k -мерное распределение. Тогда через каждую точку $x \in \mathbb{M}$ проходит единственное максимальное k -мерное интегральное многообразие $f : \mathbb{N} \hookrightarrow \mathbb{M}$ распределения $\mathcal{L}_k(\mathbb{M})$. Любое другое связное интегральное многообразие этого распределения, проходящее через точку x , содержится в $(f, \mathbb{N})$.

Доказательство. См., например, [50]), § VIII, теорема 2. $\square$

Замечание. В общем случае интегральное многообразие распределения $\mathcal{L}_k(\mathbb{M})$ может не быть регулярным подмногообразием размерности k в $\mathbb{M}$. Примером является иррациональная обмотка тора в примере 2.10.2. $\square$

Эта теорема означает, что, если на всем многообразии $\mathbb{M}$ существует гладкое инволютивное k -мерное распределение, то оно представляет собой слоение, т.е. объединение несчетного числа k -мерных максимальных интегральных многообразий (листов) данного распределения (см. следующий раздел).

Переформулируем теорему Фробениуса на языке дифференциальных форм. Дополним поля $\{X_A\} = (X_1, \dots, X_k)$ в окрестности $\mathbb{U} \subset \mathbb{M}$ $n - k$ векторными полями $\{X_M\} = (X_{k+1}, \dots, X_n)$ так, чтобы вся совокупность векторных полей $\{X_A, X_M\} = (X_1, \dots, X_n)$ была линейно независима в каждой точке $x \in \mathbb{U}$. Пусть $\{A^A, A^M\} = (A^1, \dots, A^n)$ – соответствующие дуальные 1-формы. Тогда в каждой точке $x \in \mathbb{U}$ 1-формы A^M определяют ортогональное дополнение $(\mathcal{L}_k)^\perp$. Это значит, что локально задание распределения векторных полей эквивалентно нахождению решения системы $n - k$ уравнений на дифференциалы dx^α ,

$$A^M = dx^\alpha A_\alpha^M = 0, \quad M = k + 1, \dots, n. \quad (2.88)$$

Эта система называется *пфаффовской системой уравнений*.

Из тождества (3.37) следуют равенства

$$2dA^M(X_A, X_B) = X_A(A^M(X_B)) - X_B(A^M(X_A)) - A^M([X_A, X_B]) = 0, \quad (2.89)$$

т.к. распределение векторных полей удовлетворяет условию теоремы Фробениуса (2.86). Это значит, что распределение векторных полей $\mathcal{L}_k$ удовлетворяет условию теоремы Фробениуса тогда и только тогда, когда

$$dA^M(X_A, X_B) = 0, \quad \forall A, B = 1, \dots, k, \quad \forall M = k + 1, \dots, n.$$

Для любой 2-формы, в том числе и для dA^M , справедливо представление

$$dA^M = B^M_N \wedge A^N + g^M_{BA} A^A \wedge A^B,$$

где $\{B^M_N\}$ и $\{g^M_{BA}\}$ – некоторые наборы, соответственно, 1-форм и функций. Поскольку $A^A(X_B) = \delta^A_B$, то из равенства (2.89) следует, что $g^M_{BA} = 0$. Таким образом, справедливо разложение

$$dA^M = B^M_N \wedge A^N,$$

которое эквивалентно условию коммутативности векторных полей (2.86).

Определение. Система уравнений Пфаффа называется *вполне интегрируемой*, если существует такая система координат $y^\alpha(x)$, что для подмногообразий, определяемых системой уравнений

$$y^M = \text{const} \quad \Leftrightarrow \quad dy^M = dx^\alpha \partial_\alpha y^M = 0, \quad M = k+1, \dots, n, \quad (2.90)$$

выполнялась система уравнений Пфаффа (2.88). $\square$

Для вполне интегрируемой системы уравнений Пфаффа распределение векторных полей $\mathcal{L}_k$ задается в точности первыми k координатными полями $\partial_1, \dots, \partial_k$. Обратное утверждение также справедливо. Таким образом теорему Фробениуса можно переформулировать полностью в терминах 1-форм.

Теорема 2.11.4 (Фробениус). Для того, чтобы система уравнений Пфаффа

$$A^M = 0, \quad M = k+1, \dots, n, \quad (2.91)$$

была вполне интегрируемой, необходимо и достаточно, чтобы существовал такой набор 1-форм B^M_N , что

$$dA^M = B^M_N \wedge A^N. \quad (2.92)$$

Следствие. Если все 1-формы A^M замкнуты, $dA^M = 0$, то система уравнений Пфаффа (2.91) вполне интегрируема. $\square$

Пример 2.11.3. Рассмотрим уравнение Пфаффа на 1-форму заданную на евклидовой плоскости $\mathbb{R}^2$ в декартовой системе координат

$$dxP + dyQ = 0, \quad (2.93)$$

где $P(x, y), Q(x, y) \in \mathcal{C}^1(\mathbb{R}^2)$. Для определенности будем считать, что $P \neq 0$. Тогда 1-формы

$$A^1 := dxP + dyQ \quad \text{и} \quad A^2 := dy$$

можно выбрать в качестве базиса 1-форм на плоскости. Соответствующий ему дуальный базис векторных полей имеет вид

$$X_1 = \left(\frac{1}{P}, 0 \right), \quad X_2 = \left(-\frac{Q}{P}, 1 \right).$$

Действительно, нетрудно проверить равенство $A^i(X_j) = \delta^i_j$, где $i, j = 1, 2$. Векторное поле X_2 определяет ортогональное дополнение к A^1 и задает уравнения, определяющие интегральные кривые в параметрическом виде

$$\dot{x} = -\frac{Q}{P}, \quad \dot{y} = 1.$$

Отсюда следует уравнение на форму интегральной кривой

$$\frac{dx}{dy} = -\frac{Q}{P},$$

которое эквивалентно уравнению Пфаффа (2.93). Это уравнение всегда интегрируемо. Нетрудно проверить, что в двумерном случае 1-формы B в условии теоремы Фробениуса (2.92) всегда существуют.

Отметим, что интегрируемость уравнения Пфаффа (2.93) совсем не означает, что 1-форма A^1 является точной, т.е. $A^1 = df$ для некоторой функции $f \in C^\infty(\mathbb{R}^2)$. Критерием точности 1-формы на плоскости являются нетривиальные условия интегрируемости $\partial_x Q = \partial_y P$, которые не имеют отношения к интегрируемости уравнений Пфаффа. Полная интегрируемость уравнения Пфаффа $A^1 = 0$ означает существование такой функции $f \in C^\infty(\mathbb{R}^2)$, что $df = 0$ для любой траектории, определяемой уравнением Пфаффа (2.93), а не то, что $A = df$. $\square$

Пример 2.11.4. Рассмотрим уравнение Пфаффа на 1-форму в трехмерном евклидовом пространстве $\mathbb{R}^3$ в декартовой системе координат

$$A = dxP + dyQ + dzR = 0, \quad (2.94)$$

где $P, Q, R \in C^1(\mathbb{R}^3)$. Полная интегрируемость этого уравнения означает существование функции $f \in C^2(\mathbb{R}^3)$ такой, что равенство $f = \text{const}$ является первым интегралом уравнения Пфаффа (2.94). Из теоремы Фробениуса следует критерий интегрируемости, который можно (только для трехмерного многообразия) записать в виде

$$dA = B \wedge A \quad \Leftrightarrow \quad dA \wedge A = 0. \quad (2.95)$$

Поскольку

$$dA = dx \wedge dy(\partial_x Q - \partial_y P) + dy \wedge dz(\partial_y R - \partial_z Q) + dz \wedge dx(\partial_z P - \partial_x R),$$

то необходимым и достаточным условием интегрируемости уравнения (2.94) является равенство

$$dA \wedge A = dx \wedge dy \wedge dz [P(\partial_y R - \partial_z Q) + Q(\partial_z P - \partial_x R) + R(\partial_x Q - \partial_y P)] = 0. \quad \square$$

Пример 2.11.5. Покажем, что в общем случае, когда распределение не является инволютивным, интегральные многообразия отсутствуют. Пусть в трехмерном евклидовом пространстве $\mathbb{R}^3$ задана 1-форма

$$A = dx + ydz.$$

Ее внешняя производная равна

$$dA = dy \wedge dz.$$

Поскольку

$$dA \wedge A = dx \wedge dy \wedge dz \neq 0,$$

то условие теоремы Фробениуса не выполняется, которое в трехмерном случае имеет вид (2.95). Это значит, что уравнение Пфаффа $A = 0$ в данном случае неинтегрируемо.

Посмотрим на эту задачу с точки зрения векторных полей. Ортогональное дополнение к форме A натянуто, например, на векторные поля

$$X = -y\partial_x + \partial_z \quad \text{и} \quad Y = \partial_y,$$

поскольку они линейно независимы и

$$A(X) = 0, \quad A(Y) = 0.$$

Эти векторные поля не находятся в инволюции, т.к. их коммутатор нельзя представить в виде линейной комбинации

$$[X, Y] = \partial_x \neq fX + gY, \quad f, g \in \mathcal{C}^\infty(\mathbb{R}^3).$$

Это значит, что условие теоремы Фробениуса нарушено.

Докажем от противного, что у векторных полей X, Y интегральных поверхностей не существует. Допустим, что через начало координат проходит интегральная поверхность. Это значит, что, двигая начало координат вдоль интегральных кривых, задаваемых векторными полями, мы никогда не покинем эту поверхность. Векторные поля X, Y задают потоки $s_X(t), s_Y(t)$, проходящие через точку (x_0, y_0, z_0) (см. раздел 2.6.5),

$$\begin{aligned} s_X : \quad t \times (x_0, y_0, z_0) &\mapsto (-y_0 t + x_0, y_0, t + z_0), \\ s_Y : \quad t \times (x_0, y_0, z_0) &\mapsto (x_0, t + y_0, z_0). \end{aligned}$$

Коммутатору векторных полей $[Y, X]$ соответствует отображение

$$s_Y^{-1}(t_2)s_X^{-1}(t_1)s_Y(t_2)s_X(t_1)$$

При этом начало координат перейдет в точку

$$(0, 0, 0) \xrightarrow{s_X(t_1)} (0, 0, t_1) \xrightarrow{s_Y(t_2)} (0, t_2, t_1) \xrightarrow{s_X^{-1}(t_1)} (t_1 t_2, t_2, 0) \xrightarrow{s_Y^{-1}(t_2)} (t_1 t_2, 0, 0).$$

Ясно, что не существует такой линейной комбинации векторных полей $fX + gY$, чей поток переводил бы начало координат в точку $(t_1 t_2, 0, 0)$. Тем самым мы пришли к противоречию, что и доказывает отсутствие интегральной поверхности, проходящей через начало координат. Аналогичное построение можно провести для произвольной точки $\mathbb{R}^3$. По сути дела доказательство теоремы Фробениуса сводится к такому же построению в общем случае. $\square$

2.12 Слоения

Слоения обобщают понятие расслоения, введенное в разделе 2.4. Они часто возникают и играют большую роль в теории динамических систем. Мы начнем с простейшего примера слоения, который поможет понять данное ниже общее определение.

Пример 2.12.1 (Тривиальное слоение). Представим евклидово пространство $\mathbb{R}^n$ с естественной топологией и декартовой системой координат в виде прямого произведения:

$$\mathbb{R}^n \ni x = (y, z) \in \mathbb{R}^k \times \mathbb{R}^{n-k}, \quad 1 \leq k < n.$$

Построим многообразие $\mathbb{M}^k$ меньшей размерности k , которое как множество совпадает со всем $\mathbb{R}^n$, т.е. $\mathbb{M}^k = \mathbb{R}^n$, и задает на нем структуру слоения. Базу топологии $\mathbb{M}^k$ зададим открытыми множествами вида

$$\mathbb{B}_z := \{(y, z) \in \mathbb{R}^n : y \in \mathbb{O}, \},$$

где $\mathbb{O}$ – открытое подмножество в $\mathbb{R}^k$, для каждой фиксированной точки $z \in \mathbb{R}^{n-k}$. Эта топология тоньше исходной топологии в $\mathbb{R}^n$. Назовем листом тривиального слоения сечение евклидова пространства $\mathbb{R}^n$ гиперплоскостью $z = \text{const}$:

$$\mathbb{V}_z := \{(y, z) \in \mathbb{R}^n : y \in \mathbb{R}^k, z = \text{const}\}.$$

Тогда многообразие $\mathbb{M}^k$ представляет собой несвязное объединение всех листов

$$\mathbb{M}^k = \bigcup_{z \in \mathbb{R}^{n-k}} \mathbb{V}_z$$

и покрывается картами

$$\psi_z : \mathbb{V}_z \ni (y, z) \mapsto y \in \mathbb{R}^k.$$

Рассмотренную конструкцию можно представить по-другому. Рассмотрим два евклидова пространства: $\mathbb{R}^k$ с обычной топологией и $\mathbb{R}_D^{n-k}$ с дискретной топологией, что отмечено индексом D . Тогда многообразие $\mathbb{M}^k$ является топологическим произведением $\mathbb{M}^k = \mathbb{R}^k \times \mathbb{R}_D^{n-k}$ и как множество совпадает с $\mathbb{R}^n$. $\square$

В рассмотренном примере многообразие $\mathbb{M}^k$ несвязно и состоит из несчетного множества связных компонент (листов) $\mathbb{V}_z$, которые параметризуются точками $z \in \mathbb{R}^{n-k}$. Эти свойства лежат в основе общего определения.

Определение. Пусть на множестве $\mathbb{M}$ заданы две структуры многообразия разных размерностей. То есть у нас есть два многообразия: $\mathbb{M}^n$, $\dim \mathbb{M}^n = n$, и $\mathbb{M}^k$, $\dim \mathbb{M}^k = k$, где $1 \leq k < n$, которые совпадают как множества: $\mathbb{M}^n = \mathbb{M}^k = \mathbb{M}$. Многообразие $\mathbb{M}^n$ называется *слоением*, если на нем дополнительно определена другая более тонкая топология и вторая соответствующая ей структура многообразия меньшей размерности $\mathbb{M}^k$. Поскольку размерность многообразия $\mathbb{M}^k$ меньше размерности многообразия $\mathbb{M}^n$, то многообразие $\mathbb{M}^k$ не может быть связным, и $\mathbb{M}^n$ представляет собой несчетное объединение связных компонент $\mathbb{M}^k$. Связная компонента многообразия $\mathbb{M}^k$, содержащая точку $p \in \mathbb{M}^n$, называется *листом слоения*, проходящим через эту точку. Мы говорим, что многообразие $\mathbb{M}^k$ *слоит* многообразие $\mathbb{M}^n$. Числа k и $n - k$ называются, соответственно, *размерностью* и *контрразмерностью* слоения. Слоение мы будем обозначать парой $(\mathbb{M}^n, \mathbb{M}^k)$. $\square$

По определению, через каждую точку $p \in \mathbb{M}^n$ проходит один и только один лист. В общем случае листы устроены довольно сложно. Как покажет дальнейший пример слоения тора, листы могут не быть даже регулярными подмногообразиями в $\mathbb{M}^n$. Многообразие $\mathbb{M}^n$ является объединением своих листов, число которых бесконечно и не может быть счетным из-за разных размерностей. Многообразие $\mathbb{M}^k$ также является объединением листов и не может быть связным, т.к. каждый лист является открытым подмножеством и представляет собой компоненту связности. В этом смысле многообразие $\mathbb{M}^n$ слоится на листы.

Можно доказать, что если $(\mathbb{M}^n, \mathbb{M}^k)$ – слоение, то на многообразии $\mathbb{M}^n$ существует атлас $\{\mathbb{U}_i, \varphi_i\}$, $i \in I$, со следующими свойствами. Пусть

$$\varphi: \mathbb{M}^n \supset \mathbb{U} \rightarrow \varphi(\mathbb{U}) \subset \mathbb{R}^n$$

– карта исходной дифференцируемой структуры многообразия $\mathbb{M}^n$. Представим евклидово пространство $\mathbb{R}^n$ в виде прямого произведения $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$, тогда для всех $p \in \mathbb{U}$

$$\varphi(p) = (y, z) \in \mathbb{R}^k \times \mathbb{R}^{n-k}.$$

Для каждой фиксированной точки $z \in \mathbb{R}^{n-k}$ обозначим

$$\varphi(\mathbb{U})|_z = \{(y, z) \in \varphi(\mathbb{U}) : z = \text{const}\}$$

сечение образа $\varphi(\mathbb{U})$ гиперплоскостью $z = \text{const}$. Соответствующий прообраз

$$\mathbb{V}_z := \varphi^{-1}(\varphi(\mathbb{U})|_z)$$

называется *локальным листом*. Локальный лист является частью листа, проходящего через точку $\varphi^{-1}(y, z)$. Каждый локальный лист – это открытое подмногообразие (возможно пустое) в $\mathbb{M}^k$, а пара $(\mathbb{V}_z, \psi_z)$ для непустого $\mathbb{V}_z$ является картой на многообразии $\mathbb{M}^k$:

$$\psi_z: \mathbb{M}^k \supset \mathbb{V}_z \rightarrow \psi_z(\mathbb{V}_z) \subset \mathbb{R}^k,$$

где отображение $\psi_z := \pi_k \circ \varphi|_{\mathbb{V}_z}$ определено естественной проекцией

$$\pi_k: \mathbb{R}^k \times \mathbb{R}^{n-k} \ni (y, z) \mapsto y \in \mathbb{R}^k$$

и сужением $\varphi|_{\mathbb{V}_z}$ отображения φ на $\mathbb{V}_z$. Описанная конструкция изображена на рис.2.12.

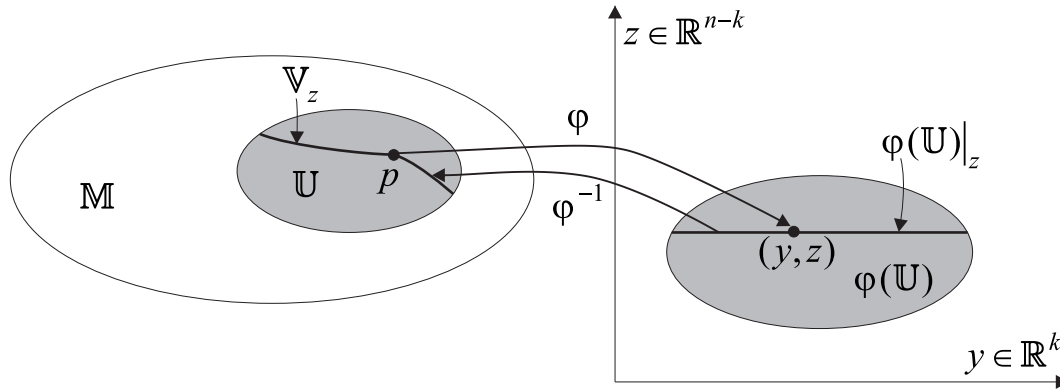


Рис. 2.12: Карта и локальный лист слоения

Функции склейки данного атласа для любых пересекающихся карт имеют вид

$$\varphi_j \circ \varphi_i^{-1}: \mathbb{R}^k \times \mathbb{R}^{n-k} \supset \varphi_i(\mathbb{U}_i \cap \mathbb{U}_j) \ni y, z \mapsto y', z' \in \varphi_j(\mathbb{U}_i \cap \mathbb{U}_j) \subset \mathbb{R}^k \times \mathbb{R}^{n-k},$$

где $y' = y'(y, z)$, но $z' = z'(z)$. В этом случае локальные листы для этих карт принадлежат одному листу слоения тогда и только тогда, когда они содержат хотя бы одну точку из пересечения $p \in \mathbb{U}_i \cap \mathbb{U}_j$. Условие того, что функции перехода $z' = z'(z)$ зависят только от z необходимо и достаточно, для того, чтобы склеивание локальных листов не зависело от выбора точки p из пересечения. Действительно, если $q, p \in \mathbb{U}_i \cap \mathbb{U}_j$

– две различные точки из пересечения координатных окрестностей, то в этих картах они имеют координаты (y_p, z_p) , (y'_p, z'_p) и (y_q, z_q) , (y'_q, z'_q) . Если $z_p = z_q$, то $z'_p = z'_q$ тогда и только тогда, когда $z' = z'(z)$.

Теперь рассмотрим несколько типичных примеров слоений.

Пример 2.12.2. Пусть $dy/dx = f(x, y)$ – обыкновенное дифференциальное уравнение, где правая часть $f(x, y)$ – функция класса C^r в некоторой области $U \subset \mathbb{R}^2$. Из теории обыкновенных дифференциальных уравнений известно, что через каждую точку области U проходит одна и только одна интегральная кривая (решение) данного уравнения класса C^{r+1} . Поскольку решение является функцией класса C^r от начальных данных, то объединение всех интегральных кривых образует одномерное многообразие V , которое слоит U .

Аналогичный результат справедлив и для систем обыкновенных дифференциальных уравнений. Для уравнений в частных производных ситуация существенно отличается. $\square$

Пример 2.12.3 (Распределение). Пусть на многообразии M^n задано k -мерное распределение векторных полей $\mathcal{L}_k(M^n)$, где $1 \leq k < n$, которое находится в инволюции (см. раздел 2.11). Напомним, что k -мерное распределение векторных полей ставит каждой точке многообразия некоторое подпространство размерности k в касательном пространстве к данной точке:

$$M^n \ni p \mapsto L_p(M^n) \subset T_p(M^n)$$

Тогда согласно теореме Фробениуса через каждую точку $p \in M^n$ проходит интегральное подмногообразие данного распределения размерности k . Эти интегральные подмногообразия можно рассматривать как листы слоения (M^n, M^k) . $\square$

Рассмотрим слоение M^n , которое слоится многообразием M^k . Тогда определено касательное пространство к листу $T_p(M^k)$ в произвольной точке $p \in M^n$. Размерность этого пространства равна k , и оно является k -мерным подпространством касательного пространства $T_p(M^n)$. Следовательно, определено k -мерное распределение векторных полей $\mathcal{L}_k(M^n)$ на многообразии M^n . Вместе с предыдущим примером это доказывает следующее утверждение.

Теорема 2.12.1. Любое гладкое слоение (M^n, M^k) , $1 \leq k < n$, находится во взаимно однозначном соответствии с k -мерным гладким инволютивным распределением векторных полей $\mathcal{L}_k(M^n)$, которое касается M^k , т.е. $L_p(M^n) = T_p(M^k)$. При этом лист слоения, проходящий через точку $p \in M^n$, совпадает с интегральным подмногообразием данного распределения, проходящим через ту же точку.

Замечание. В силу данной теоремы слоение часто определяют с помощью распределений и их интегральных подмногообразий. $\square$

Отметим, что не каждое топологически нетривиальное многообразие допускает структуру слоения.

Пример 2.12.4. Теорема 10.2.1 утверждает, что на четномерной сфере S^{2n} , $n = 1, 2, \dots$, не существует непрерывного векторного поля, которое всюду отлично от нуля. Это означает, что на S^{2n} невозможно задать одномерное распределение векторных полей $\mathcal{L}_1(S^{2n})$, и, следовательно одномерных слоений на четномерной сфере не существует. В частности, на двумерной сфере S^2 нельзя задать структуру слоения. $\square$

Далее мы рассмотрим более подробно связь слоений и расслоений.

Пример 2.12.5 (Расслоение). Рассмотрим дифференцируемое расслоение $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$ (см., раздел 2.4) с n -мерным пространством расслоения $\mathbb{E}$, k -мерным типичным слоем $\mathbb{F}$ и $(n - k)$ -мерной базой $\mathbb{M}$. Через каждую точку расслоения $p \in \mathbb{E}$ проходит слой $\mathbb{V}_{\pi(p)} := \pi^{-1}(\pi(p))$, диффеоморфный типичному слою $\mathbb{F}$, и, значит, имеющий структуру дифференцируемого k -мерного многообразия. Известно, что слои расслоения либо совпадают: $\mathbb{V}_{\pi(p)} = \mathbb{V}_{\pi(q)}$, если $\pi(p) = \pi(q)$, либо не пересекаются: $\mathbb{V}_{\pi(p)} \cap \mathbb{V}_{\pi(q)} = \emptyset$, если $\pi(p) \neq \pi(q)$. Рассмотрим множество, которое состоит из объединения всех слоев

$$\mathbb{E}^k = \bigcup_{x \in \mathbb{M}} \mathbb{V}_x.$$

Как множество оно совпадает с пространством расслоения $\mathbb{E}$, но мы введем на нем другую топологию и дифференцируемую структуру. А именно, будем считать, что в этой топологии и дифференцируемой структуре каждый слой является открытым подмножеством и открытым подмногообразием. Для того, чтобы проверить, что многообразие $\mathbb{E}^k$ слоит расслоение $\mathbb{E}$ следует воспользоваться картами на $\mathbb{E}$, которые тривиализируют расслоение. Как и в предыдущем примере многообразие $\mathbb{E}^k$ несвязно и состоит из несчетного числа листов, которые совпадают со слоями расслоения $\mathbb{V}_x$ и параметризуются точками базы $x \in \mathbb{M}$. В этом слоении все листы $\mathbb{V}_x$ диффеоморфны между собой и диффеоморфны типичному слою, $\mathbb{V}_x \approx \mathbb{F}$. $\square$

Рассмотренный пример показывает, что на каждом расслоении определена естественная структура слоения. Обратное утверждение неверно: не каждое слоение является расслоением. Чтобы построить соответствующий пример, нам понадобится

Определение. Дифференцируемое отображение двух многообразий $f : \mathbb{M} \rightarrow \mathbb{N}$ называется *субмерсией*, если дифференциал отображения $f_* : \mathbb{T}_p(\mathbb{M}) \rightarrow \mathbb{T}_{f(p)}(\mathbb{N})$ является сюръективным отображением для всех точек $p \in \mathbb{M}$. $\square$

Пример 2.12.6 (Субмерсия). Рассмотрим субмерсию двух многообразий $f : \mathbb{M} \rightarrow \mathbb{N}$. Из определения следует, что $\dim \mathbb{M} \geq \dim \mathbb{N}$. Мы положим $\dim \mathbb{M} = n$ и $\dim \mathbb{N} = n - k$, где $1 \leq k < n$. Каждая точка $p \in \mathbb{M}$ порождает лист или несвязное объединение листов $\mathbb{V}_p := f^{-1}(f(p))$. При этом прообразы либо совпадают: $\mathbb{V}_p = \mathbb{V}_q$, если $f(p) = f(q)$, либо не пересекаются: $\mathbb{V}_p \cap \mathbb{V}_q = \emptyset$, если $f(p) \neq f(q)$. Предположим, что отображение f сюръективно. Тогда слои можно параметризовать точками $x \in \mathbb{N}$. Из условия сюръективности $f_*(\mathbb{T}_p(\mathbb{M})) = \mathbb{T}_{f(p)}(\mathbb{N})$ для всех $p \in \mathbb{M}$ следует, что для каждой точки $p \in \mathbb{M}$ найдутся карты:

$$\begin{aligned} \varphi : \mathbb{M} \supset \mathbb{U} \ni \quad & p \mapsto \varphi(p) \in \mathbb{R}^n, \\ \chi : \mathbb{N} \supset \mathbb{W} \ni \quad & x = f(p) \mapsto \chi(x) \in \mathbb{R}^{n-k} \end{aligned}$$

со следующими свойствами:

1. $f(\mathbb{U}) \subset \mathbb{W}$,
2. $\varphi(q) = (y, z) \in \mathbb{R}^k \times \mathbb{R}^{n-k}$, $\forall q \in \mathbb{U}$,
3. $\chi(x) \in \mathbb{R}^{n-k}$, $\forall x \in \mathbb{W}$,
4. $(\chi \circ f \circ \varphi^{-1})(y, z) = z \in \chi(\mathbb{W})$, $\forall (y, z) \in \varphi(\mathbb{U})$,
5. $\varphi(\mathbb{V}_p) = \{(y, z) \in \varphi(\mathbb{U}) : z = \chi(f(p))\}$, $\forall f(p) \in \mathbb{W}$.

Отсюда легко выводится, что на множестве $\mathbb{M}$ определена структура k -мерного слоения $\mathbb{M}^k$, в котором листы, содержащиеся в $f^{-1}(x)$, являются открытыми подмногообразиями. $\square$

Ранее было показано, что любое расслоение является слоением. Приведем пример слоения, которое не является расслоением.

Пример 2.12.7. Рассмотрим субмерсию, изображенную на рис.2.13, где лента в $\mathbb{R}^2$ с вырезанной дыркой проектируется на ось x . Ленту с вырезанной дыркой мы рассматриваем, как открытое подмногообразие $\mathbb{M}$ в $\mathbb{R}^2$ с декартовыми координатами x, y . Субмерсия задается проекцией $f : (x, y) \mapsto x$ и определяет одномерное слоение $\mathbb{M}$. Соответствующее распределение векторных полей имеет вид ∂_y . Все интегральные кривые этого распределения, которые являются листами слоения, параллельны оси y . Если $x < a$ или $x > b$, то листами слоения $f^{-1}(x)$ являются интервалы. В то же время каждый прообраз $f^{-1}(x)$ для $a \leq x \leq b$ представляет несвязное объединение двух листов. Ясно, что не все прообразы $f^{-1}(x)$ данного слоения диффеоморфны между собой, и поэтому рассматриваемое слоение не может быть расслоением. $\square$

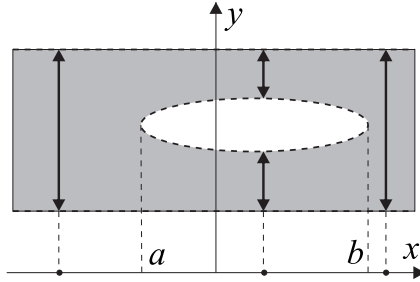


Рис. 2.13: Субмерсия ленты с вырезанной дыркой на ось x .

Следующий пример показывает, что листы слоения даже в простых случаях могут быть устроены достаточно сложно.

Пример 2.12.8 (Слоения тора). Рассмотрим двумерный тор, как квадрат на плоскости: $\mathbb{T}^2 = [0, 1] \times [0, 1] \subset \mathbb{R}^2$ с отождествленными противоположными сторонами. Пусть на нем задано векторное поле $X = X^x \partial_x + X^y \partial_y$ с постоянными компонентами $X^x = a \neq 0$ и $X^y = b$. Тем самым определено одномерное распределение векторного поля. Оно, конечно, находится в инволюции. Обозначим интегральные кривые этого векторного поля через γ_{ab} . Они же являются листами одномерного слоения тора. Если отношение b/a рационально, то каждый лист диффеоморфен окружности $\mathbb{S}^1$ и является регулярным замкнутым подмногообразием в $\mathbb{T}^2$. Если отношение b/a иррационально (всюду плотная обмотка тора в примере 2.10.2), то каждый лист диффеоморфен прямой $\mathbb{R}$. При этом отображение $\mathbb{R} \rightarrow \gamma_{ab} \in \mathbb{T}^2$ не является регулярным. В этом случае лист не является одномерным регулярным подмногообразием и замкнутым подмножеством в $\mathbb{T}^2$. $\square$

2.13 Бесконечно малые преобразования координат

В настоящем разделе рассматриваются свойства различных полей на многообразии при бесконечно малых преобразованиях координат. Полученные ниже формулы важны в приложениях, в частности, они будут использованы в следующем разделе для вычисления производных Ли.

Поскольку рассмотрение носит локальный характер, то для простоты и наглядности мы рассмотрим евклидово пространство $\mathbb{R}^n$ с фиксированной декартовой системой координат. Будем рассматривать преобразования координат (1.66) с активной точки зрения. То есть будем считать, что в евклидовом пространстве точка с координатами x^α перемещается в новую точку с координатами $x'^\alpha(x)$ в той же системе координат. Бесконечно малые (инфинитезимальные) преобразования координат

$$x'^\alpha := x^\alpha + u^\alpha(x) \quad (2.96)$$

определяются векторным полем

$$u = u^\alpha \partial_\alpha.$$

Обозначение $u^\alpha(x)$ заимствовано из теории упругости, где этот вектор называется *вектором смещения*. Будем считать, что малы сами смещения, $|u^\alpha| \ll 1^2$, и относительные смещения $|\partial_\alpha u^\beta| \ll 1$ для всех значений индексов α, β .

Рассмотрим, как меняются тензорные поля при инфинитезимальных преобразованиях координат (2.96). Начнем с простейшего случая скалярного поля (функции) $f \in C^\infty(\mathbb{R}^n)$. Из закона преобразования скалярного поля (2.5) следует равенство

$$f'(x + u) = f(x). \quad (2.97)$$

Разлагая левую часть этого равенства в ряд Тейлора³, в первом порядке по u получим следующее выражение

$$f' = f - u^\alpha \partial_\alpha f, \quad (2.98)$$

где все поля рассматриваются в точке x . Полученное соотношение интерпретируется следующим образом. При инфинитезимальном преобразовании координат (2.96) значение функции $f(x)$ в точке x меняется на бесконечно малую величину

$$\delta f(x) := f'(x) - f(x) = -u^\alpha \partial_\alpha f = -uf. \quad (2.99)$$

которая называется *приращением* или *вариацией формы* функции f в точке x . Правая часть этого равенства с обратным знаком представляет собой производную от функции f вдоль векторного поля u , определяющего преобразование координат.

Из закона преобразования векторных полей (2.11), который для инфинитезимальных преобразований в первом порядке по u можно переписать в виде

$$X'^\alpha(x + u) = X^\beta(\delta_\beta^\alpha + \partial_\beta u^\alpha),$$

следует, что вариация формы компонент векторного поля имеет вид

$$\delta X^\alpha = X^\beta \partial_\beta u^\alpha - u^\beta \partial_\beta X^\alpha = [X, u]^\alpha. \quad (2.100)$$

То есть приращение векторного поля в точке x равно коммутатору этого векторного поля с инфинитезимальным векторным полем смещения u .

Закон преобразования компонент 1-формы (2.12) приводит к следующему приращению компонент

$$\delta \omega_\alpha = -\partial_\alpha u^\beta \omega_\beta - u^\beta \partial_\beta \omega_\alpha \quad (2.101)$$

²В физике координаты являются размерными величинами, и условие малости смещений записывается в виде $|u^\alpha| \ll l$, где l – выбранная единица длины.

³Здесь и далее мы предполагаем, что разложение всех функций и компонент тензорных полей в ряд Тейлора имеет смысл в некоторой окрестности точки x .

или в ковариантном виде

$$\begin{aligned}\delta\omega_\alpha &= -\nabla_\alpha u^\beta \omega_\beta - u^\beta \nabla_\beta \omega_\alpha + u^\beta T_{\alpha\beta}{}^\gamma \omega_\gamma = \\ &= -\tilde{\nabla}_\alpha u^\beta \omega_\beta - u^\beta \tilde{\nabla}_\beta \omega_\alpha.\end{aligned}\quad (2.102)$$

При этом использовано явное выражение для аффинной связности (6.17). Напомним, что ∇ и $\tilde{\nabla}$ обозначают ковариантные производные с аффинной связностью и символами Кристоффеля, соответственно.

Замечание. Приращения компонент тензорных полей, а также связности при инфинитезимальных преобразованиях координат являются тензорными полями. Поскольку бесконечно малые преобразования координат параметризуются векторным полем u , то вариации полей всегда можно записать в явно ковариантном виде. $\square$

Аналогично находится явный вид бесконечно малых преобразований координат для тензоров произвольного ранга. Рассмотрим бесконечно малые преобразования основных геометрических объектов: метрики и аффинной связности, а также репера и линейной связности. В частности, для метрики справедливо равенство

$$\delta g_{\alpha\beta} = -\partial_\alpha u^\gamma g_{\gamma\beta} - \partial_\beta u^\gamma g_{\alpha\gamma} - u^\gamma \partial_\gamma g_{\alpha\beta} \quad (2.103)$$

или в ковариантном виде

$$\begin{aligned}\delta g_{\alpha\beta} &= -\nabla_\alpha u_\beta - \nabla_\beta u_\alpha + u^\gamma (T_{\alpha\gamma\beta} + T_{\beta\gamma\alpha} - Q_{\alpha\gamma\beta} - Q_{\beta\gamma\alpha} + Q_{\gamma\alpha\beta}) = \\ &= -\tilde{\nabla}_\alpha u_\beta - \tilde{\nabla}_\beta u_\alpha.\end{aligned}\quad (2.104)$$

Отсюда следует, что, если метрика инвариантна относительно бесконечно малых преобразований координат, $\delta g_{\alpha\beta} = 0$, то соответствующее векторное поле должно удовлетворять уравнению Киллинга (15.4).

Из уравнения (2.103) следует правило преобразования определителя метрики

$$\delta g = -u^\alpha \partial_\alpha g - 2\partial_\alpha u^\alpha g. \quad (2.105)$$

Рассмотрим, как меняются компоненты аффинной связности при инфинитезимальных преобразованиях координат. Из закона преобразования (6.5) следует выражение для вариации формы компонент аффинной связности:

$$\delta\Gamma_{\alpha\beta}{}^\gamma = -\partial_\alpha u^\delta \Gamma_{\delta\beta}{}^\gamma - \partial_\beta u^\delta \Gamma_{\alpha\delta}{}^\gamma + \Gamma_{\alpha\beta}{}^\delta \partial_\delta u^\gamma - \partial_{\alpha\beta}^2 u^\gamma - u^\delta \partial_\delta \Gamma_{\alpha\beta}{}^\gamma. \quad (2.106)$$

В ковариантном виде это изменение имеет вид

$$\delta\Gamma_{\alpha\beta}{}^\gamma = -\nabla_\alpha \nabla_\beta u^\gamma + \nabla_\alpha u^\delta T_{\beta\delta}{}^\gamma + u^\delta (R_{\alpha\delta\beta}{}^\gamma + \nabla_\alpha T_{\beta\delta}{}^\gamma). \quad (2.107)$$

В частном случае для символов Кристоффеля справедливо равенство

$$\delta\tilde{\Gamma}_{\alpha\beta}{}^\gamma = -\tilde{\nabla}_\alpha \tilde{\nabla}_\beta u^\gamma + u^\delta \tilde{R}_{\alpha\delta\beta}{}^\gamma. \quad (2.108)$$

Эту вариацию можно переписать в виде

$$\delta\tilde{\Gamma}_{\alpha\beta\gamma} = \frac{1}{2}(\tilde{\nabla}_\alpha \delta g_{\beta\gamma} + \tilde{\nabla}_\beta \delta g_{\alpha\gamma} - \tilde{\nabla}_\gamma \delta g_{\alpha\beta}). \quad (2.109)$$

В римановой геометрии, где связность однозначно определяется метрикой, инвариантность метрики является достаточным условием инвариантности символов Кристоффеля. Необходимое и достаточное условие инвариантности символов Кристоффеля относительно бесконечно малых преобразований координат имеет вид $\delta\tilde{\Gamma}_{\alpha\beta}^\gamma = 0$, где вариация определяется формулой (2.108).

Антисимметризация уравнения (2.107) по индексам α, β дает приращение компонент тензора кручения:

$$\begin{aligned}\delta T_{\alpha\beta}^\gamma &= \nabla_\alpha u^\delta T_{\beta\delta}^\gamma - \nabla_\beta u^\delta T_{\alpha\delta}^\gamma + T_{\alpha\beta}^\delta \nabla_\delta u^\gamma + \\ &+ u^\delta (-R_{\alpha\beta\delta}^\gamma - R_{\delta\alpha\beta}^\gamma - R_{\beta\delta\alpha}^\gamma + \nabla_\alpha T_{\beta\delta}^\gamma - \nabla_\beta T_{\alpha\delta}^\gamma)\end{aligned}\quad (2.110)$$

или, с учетом (6.79),

$$\begin{aligned}\delta T_{\alpha\beta}^\gamma &= \nabla_\alpha u^\delta T_{\beta\delta}^\gamma - \nabla_\beta u^\delta T_{\alpha\delta}^\gamma + T_{\alpha\beta}^\delta \nabla_\delta u^\gamma - \\ &- u^\delta (\nabla_\delta T_{\alpha\beta}^\gamma + T_{\alpha\beta}^\epsilon T_{\epsilon\delta}^\gamma + T_{\beta\delta}^\epsilon T_{\epsilon\alpha}^\gamma + T_{\delta\alpha}^\epsilon T_{\epsilon\beta}^\gamma)\end{aligned}\quad (2.111)$$

Нетрудно получить также вариацию формы компонент тензора неметричности при инфинитезимальных преобразованиях координат:

$$\begin{aligned}\delta Q_{\alpha\beta\gamma} &= -\nabla_\alpha u^\delta Q_{\delta\beta\gamma} - \nabla_\beta u^\delta Q_{\alpha\delta\gamma} - \nabla_\gamma u^\delta Q_{\alpha\beta\delta} + \\ &+ u^\delta (T_{\alpha\delta}^\epsilon Q_{\epsilon\beta\gamma} + T_{\beta\delta}^\epsilon Q_{\alpha\epsilon\gamma} + T_{\gamma\delta}^\epsilon Q_{\alpha\beta\epsilon}).\end{aligned}\quad (2.112)$$

В аффинной геометрии при заданной связности можно поставить задачу нахождения таких преобразований координат, которые не меняют связность, кручение или неметричность. Для этого необходимо решить, соответственно, уравнения

$$\delta\Gamma_{\alpha\beta}^\gamma = 0, \quad \delta T_{\alpha\beta}^\gamma = 0, \quad \delta Q_{\alpha\beta}^\gamma = 0.$$

Репер и лоренцева связность являются коекторами относительно преобразований координат. Поэтому их приращения такие же, как и приращения 1-форм (2.101). Приведем их явный вид, т.к. они часто встречаются в приложениях

$$\delta e_\alpha^a = -\partial_\alpha u^\beta e_\beta^a - u^\beta \partial_\beta e_\alpha^a = -\tilde{\nabla}_\alpha u^a + u^\beta \tilde{\omega}_{\beta b}^a e_\alpha^b, \quad (2.113)$$

$$\delta\omega_\alpha^{ab} = -\partial_\alpha u^\beta \omega_\beta^{ab} - u^\beta \partial_\beta \omega_\alpha^{ab} = -\tilde{\nabla}_\alpha u^\beta \omega_\beta^{ab} - u^\beta (\partial_\beta \omega_\alpha^{ab} - \tilde{\Gamma}_{\beta\alpha}^\gamma \omega_\gamma^{ab}). \quad (2.114)$$

Из формулы (2.113) следует выражение для вариации формы элемента объема

$$\delta\sqrt{|g|} = -u^\alpha \partial_\alpha \sqrt{|g|} - \sqrt{|g|} \partial_\alpha u^\alpha = -\sqrt{|g|} \tilde{\nabla}_\alpha u^\alpha, \quad (2.115)$$

Бесконечно малые приращения различных полей, рассмотренные в настоящем разделе, называются вариацией формы, т.к. показывают как меняется значение полей в фиксированной точке многообразия. Следующее свойство проверяется прямой проверкой и используется при доказательстве теоремы Нетер.

Предложение 2.13.1. *Вариация формы производной скалярного поля равно производной от вариации*

$$\delta(\partial_\alpha f) = \partial_\alpha(\delta f).$$

□

Рассмотрим коммутатор двух бесконечно малых преобразований координат. Пусть эти преобразования задаются векторными полями $u = u^\alpha \partial_\alpha$ и $v = v^\alpha \partial_\alpha$. Для простоты, ограничимся их последовательным действием на функцию $f(x)$. Поскольку

преобразования координат евклидова пространства $\mathbb{R}^n$ образуют псевдогруппу, то коммутатор двух преобразований также является преобразованием координат с параметром, который квадратичен по компонентам u^α и v^α . Поэтому, в отличие от предыдущего рассмотрения, для вычисления коммутатора необходимо удерживать квадратичные слагаемые.

Сначала совершим преобразование координат, определяемое векторным полем u . Из закона преобразования функции (2.97) с учетом квадратичных слагаемых получаем равенство

$$f' + u^\alpha \partial_\alpha f' + \frac{1}{2} u^\alpha u^\beta \partial_{\alpha\beta}^2 f' = f.$$

Во втором слагаемом в левой части равенства f' можно выразить через f , воспользовавшись линейным приближением (2.98), а в третьем слагаемом заменим f' на f , потому что коэффициент перед ним уже квадратичен по u . В результате с точностью до квадратичных слагаемых получаем разложение

$$f' := (1 + T_u)f = f - u^\alpha \partial_\alpha f + u^\alpha \partial_\alpha u^\beta \partial_\beta f + \frac{1}{2} u^\alpha u^\beta \partial_{\alpha\beta}^2 f + \dots, \quad (2.116)$$

где мы ввели генератор общих преобразований координат T_u и выписали его действие на функцию в квадратичном приближении.

Совершим теперь второе преобразование координат с параметром v

$$\begin{aligned} (1 + T_v)f' &= (1 + T_v)(1 + T_u)f = \\ &= f' - v^\alpha \partial_\alpha f' + v^\alpha \partial_\alpha u^\beta \partial_\beta f' + \frac{1}{2} v^\alpha v^\beta \partial_{\alpha\beta}^2 f' + \dots = \\ &= f - (u^\alpha + v^\alpha - u^\beta \partial_\beta u^\alpha - v^\beta \partial_\beta v^\alpha - v^\beta \partial_\beta u^\alpha) \partial_\alpha f + \\ &\quad + \frac{1}{2} (u^\alpha u^\beta + 2u^\alpha v^\beta + v^\alpha v^\beta) \partial_{\alpha\beta}^2 f + \dots \end{aligned}$$

Вычитая из этого выражения результат тех же преобразований в обратном порядке, получим явное выражение для коммутатора двух преобразований координат

$$[T_v, T_u]f = T_{[v, u]}f = (v^\beta \partial_\beta u^\alpha - u^\beta \partial_\beta v^\alpha) \partial_\alpha f. \quad (2.117)$$

Таким образом, коммутатор двух преобразований координат в евклидовом пространстве $\mathbb{R}^n$, определяемых инфинитезимальными векторными полями u и v , является преобразованием координат, которое задается коммутатором векторных полей $[v, u]$.

Из групповых соображений следует, что коммутатор двух преобразований координат не зависит от представления. Другими словами, это же выражение для коммутатора имеет место не только для функций, но и для тензорных полей или плотностей более высокого ранга. В последнем случае вычисления являются более громоздкими.

Рассмотрим, как меняются при бесконечно малых преобразованиях тензорные плотности, заданные на многообразии $\mathbb{M}$. Мы уже получили правила преобразования определителей метрики (2.105) и репера (2.115), которые являются тензорными плотностями, соответственно, степеней -2 и -1 . В общем случае, если на $\mathbb{M}$ задана скалярная плотность f степени $\deg f = p$, то из правила преобразования тензорных плотностей (2.7) следует правило

$$\delta f = -u^\alpha \partial_\alpha f + p \partial_\alpha u^\alpha f,$$

т.к. якобиан бесконечно малых преобразований координат (2.96) в первом порядке по вектору смещения имеет вид

$$J = \det(\delta_\alpha^\beta + \partial_\alpha u^\beta) \approx 1 + \partial_\alpha u^\alpha. \quad (2.118)$$

Если на многообразии $\mathbb{M}$ задана тензорная плотность T произвольного ранга и степени p , то ее вариация формы равна

$$\delta T = \dots + p \partial_\alpha u^\alpha T,$$

где точки обозначают набор обычных тензорных слагаемых для всех ковариантных и контравариантных индексов.

2.14 Производная Ли

Понятие потока $\mathbb{R} \times \mathbb{M} \xrightarrow{s} \mathbb{M}$ или экспоненциального отображения $s^\alpha(t, x) = \exp(tX)x^\alpha$, генерируемого отличным от нуля дифференцируемым векторным полем $X \in \mathcal{X}(\mathbb{M})$ (раздел 2.6.5), позволяет определить производную Ли L_X от произвольного тензорного поля T вдоль векторного поля X . Для этого определения достаточно, чтобы поток существовал локально.

Рассмотрим произвольную точку $x \in \mathbb{M}$. Тогда у нее существует окрестность $\mathbb{U}_x$ такая, что отображение $s : \mathbb{U}_x \rightarrow s(\mathbb{U}_x)$ является диффеоморфизмом. Поэтому для него определены дифференциал отображения s_* и обратное отображение s^{*-1} к возврату отображения s^* . Пусть в некоторой окрестности точки $x \in \mathbb{M}$ задано тензорное поле $T \in \mathcal{T}_s^r(\mathbb{M})$ типа (r, s) . В результате экспоненциального отображения $s(\epsilon, x)$ с малым параметром ϵ тензор $T(x)$ в точке x отобразится в тензор $\tilde{s}T(x)$ в точке $s(\epsilon, x)$, где $\tilde{s}T$ обозначает продолжение отображения $(s_*)^r(s^{*-1})^s T$, заданного в касательном и кокасательном пространствах, на всю тензорную алгебру (в компонентах: на каждый контравариантный индекс действует дифференциал отображения s_* , а на каждый ковариантный индекс – обратное отображение s^{*-1}). Это значит, что в точку x отобразится тензор из точки $s(-\epsilon, x)$:

$$T(s(-\epsilon, x)) \mapsto \tilde{s}(\epsilon, s(-\epsilon, x))T(s(-\epsilon, x)).$$

Определение. *Производной Ли* от тензорного поля T вдоль векторного поля X в точке x называется предел

$$L_X T := \lim_{\epsilon \rightarrow 0} \frac{T(x) - \tilde{s}(\epsilon, s(-\epsilon, x))T(s(-\epsilon, x))}{\epsilon}. \quad (2.119)$$

В упрощенной записи мы пишем

$$L_X T = \lim_{\epsilon \rightarrow 0} \frac{T(x) - \tilde{s}(\epsilon)T}{\epsilon}, \quad (2.120)$$

где мы опустили аргумент $s(-\epsilon, x) \in \mathbb{M}$. $\square$

Для дифференцируемых векторных и тензорных полей этот предел существует.

Замечание. При малых ϵ экспоненциальное отображение имеет вид $x^\alpha \mapsto s^\alpha(\epsilon, x) = x^\alpha + \epsilon X^\alpha + \dots$, т.е. соответствует бесконечно малым преобразованиям координат, рассмотренным в предыдущем разделе, с вектором смещения $u^\alpha = \epsilon X^\alpha$. При этом выражение, стоящее в числителе производной Ли (2.119), совпадает с вариацией формы тензорного поля T , взятой с обратным знаком. Это следует непосредственно из определения вариации формы тензорного поля. Поэтому определение (2.119) для компонент тензорного поля типа (r, s) принимает вид

$$L_X T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}(x) = - \lim_{\epsilon \rightarrow 0} \frac{\delta T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}(x)}{\epsilon}, \quad (2.121)$$

где $\delta T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}(x)$ – вариация формы компонент тензорного поля типа (r, s) , рассмотренная в разделе 2.13. Несложные вычисления приводят к следующему выражению для производной Ли компонент тензорного поля

$$\begin{aligned} \mathcal{L}_X T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r} = & X^\gamma \partial_\gamma T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r} + \\ & + \partial_{\beta_1} X^\gamma T_{\gamma \beta_2 \dots \beta_s}^{\alpha_1 \dots \alpha_r} + \dots + \partial_{\beta_s} X^\gamma T_{\beta_1 \dots \beta_{s-1} \gamma}^{\alpha_1 \dots \alpha_r} - \\ & - T_{\beta_1 \dots \beta_s}^{\gamma \alpha_2 \dots \alpha_r} \partial_\gamma X^{\alpha_1} - \dots - T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_{r-1} \gamma} \partial_\gamma X^{\alpha_r}. \end{aligned} \quad (2.122)$$

Первое слагаемое в правой части (2.122) соответствует смещению самой точки x , слагаемые во второй и третьей строках (2.122) возникают при действии отображений s^{*-1} и s_* на каждый ковариантный и контравариантный индекс, соответственно. $\square$

Формула (2.122) конструктивна и позволяет получить явные выражения для производных Ли различных тензорных полей в координатах.

Пример 2.14.1. В простейшем случае скалярного поля $f(x)$ производная Ли совпадает с производной функции вдоль векторного поля. Покажем это. Поскольку на скалярное поле ни дифференциал отображения, ни его возврат не действуют, то из определения (2.119) следует

$$\mathcal{L}_X f = \lim_{\epsilon \rightarrow 0} \frac{f(x) - f(x - \epsilon X)}{\epsilon} = X^\alpha \partial_\alpha f.$$

Таким образом, производная Ли от функции – это просто производная вдоль векторного поля. $\square$

Пример 2.14.2. Из уравнения (2.122) следует выражение для производной Ли от векторного поля Y в координатах:

$$\mathcal{L}_X Y^\alpha = X^\beta \partial_\beta Y^\alpha - Y^\beta \partial_\beta X^\alpha,$$

что совпадает с коммутатором векторных полей (2.49)

$$\mathcal{L}_X Y = [X, Y]. \quad (2.123)$$

Определение производной Ли (2.119) для векторного поля принимает вид

$$\mathcal{L}_X Y = \lim_{\epsilon \rightarrow 0} \frac{Y - s_*(\epsilon)Y}{\epsilon}. \quad \square \quad (2.124)$$

Производная Ли от векторного поля обладает следующими свойствами:

$$\begin{aligned} \mathcal{L}_{fX} Y &= f[X, Y] - (Yf)X, \\ \mathcal{L}_X (fY) &= f[X, Y] + (Xf)Y, \end{aligned}$$

где $f(x) \in \mathcal{C}^1(\mathbb{M})$ – произвольное скалярное поле.

Отметим, что тождества Якоби (2.48) в алгебре Ли векторных полей можно переписать в эквивалентном виде, используя производную Ли,

$$\mathcal{L}_X [Y, Z] = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z].$$

В таком виде тождества Якоби аналогичны правилу Лейбница.

Можно также проверить следующее свойство производной Ли:

$$\mathcal{L}_{[X, Y]} = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X. \quad (2.125)$$

Действие этого равенства на векторное поле сводится к тождеству Якоби.

Производная Ли отличается от ковариантной производной вдоль векторного поля (6.12).

Предложение 2.14.1. Для компонент векторного поля справедливо равенство

$$\mathbb{L}_X Y^\alpha - \nabla_X Y^\alpha = -Y^\beta (\nabla_\beta X^\alpha + X^\gamma T_{\beta\gamma}^\alpha), \quad (2.126)$$

где ∇_α – ковариантная производная и $T_{\beta\gamma}^\alpha$ – тензор кручения.

Доказательство. Простая проверка. $\square$

Пример 2.14.3. Из формулы (2.101) следует, что производная Ли от 1-формы $A = dx^\alpha A_\alpha$ имеет вид

$$\mathbb{L}_X A_\alpha = X^\beta \partial_\beta A_\alpha + \partial_\alpha X^\beta A_\beta.$$

При этом разность с ковариантной производной вдоль векторного поля имеет другой знак по сравнению с (2.126):

$$\mathbb{L}_X A_\alpha - \nabla_X A_\alpha = (\nabla_\alpha X^\beta + X^\gamma T_{\alpha\gamma}^\beta) A_\beta. \quad \square$$

Определение. Дифференцированием D тензорной алгебры $\mathcal{T}(\mathbb{M})$ называется линейный эндоморфизм $\mathcal{T}(\mathbb{M})$, удовлетворяющий следующим свойствам:

- 1) D сохраняет тип тензорных полей: $D\mathcal{T}_s^r(\mathbb{M}) \subset \mathcal{T}_s^r(\mathbb{M})$.
- 2) Дифференцирование удовлетворяет правилу Лейбница. Если Y и Z – два произвольных тензорных поля на $\mathbb{M}$, то

$$D(Y \otimes Z) = DY \otimes Z + Y \otimes DZ.$$

- 3) D перестановочен с каждым свертыванием: $DC = CD$. $\square$

Предложение 2.14.2. Производная Ли $\mathbb{L}_X$ является дифференцированием в тензорной алгебре $\mathcal{T}(\mathbb{M})$.

Доказательство. Прямая проверка. $\square$

Рассмотрим тензорное поле $S = S_\beta^\alpha(x) e_\alpha \otimes e^\beta$ типа $(1, 1)$. В каждой точке $x \in \mathbb{M}$ оно задает линейный эндоморфизм касательного $\mathbb{T}_x(\mathbb{M})$ и кокасательного $\mathbb{T}_x^*(\mathbb{M})$ пространств, который естественным образом продолжается до линейного эндоморфизма D_S тензорной алгебры в этой точке. Если $T \in \mathcal{T}_s^r(\mathbb{M})$ – тензорное поле типа (r, s) , то в компонентах

$$(D_S T)_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r} = S_{\beta_1}^{\gamma} T_{\gamma \beta_2 \dots \beta_s}^{\alpha_1 \dots \alpha_r} + \dots + S_{\beta_s}^{\gamma} T_{\beta_1 \dots \beta_{s-1} \gamma}^{\alpha_1 \dots \alpha_r} - T_{\beta_1 \dots \beta_s}^{\gamma \alpha_2 \dots \alpha_r} S_{\gamma}^{\alpha_1} - \dots - T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_{r-1} \gamma} S_{\gamma}^{\alpha_r}.$$

Тогда D_S представляет собой дифференцирование тензорной алгебры $\mathcal{T}(\mathbb{M})$, индуцированное тензорным полем S типа $(1, 1)$.

Теорема 2.14.1. Каждое дифференцирование D тензорной алгебры $\mathcal{T}(\mathbb{M})$ допускает единственное разложение

$$D = \mathbb{L}_X + D_S,$$

где X есть векторное поле, а S – тензорное поле типа $(1, 1)$.

Доказательство. См., например, [48], предложение 3.3. $\square$

Замечание. Поскольку любое дифференцирование D линейно отображает алгебру функций $\mathcal{C}^\infty(\mathbb{M})$ в себя, и выполняется правило Лейбница, то векторное поле X существует и определяется единственным образом из условия $Df = Xf$ для всех $f \in \mathcal{C}^\infty(\mathbb{M})$. $\square$

Следствие. Разность двух произвольных дифференцирований есть дифференцирование, индуцированное некоторым тензорным полем типа $(1, 1)$. $\square$

Таким образом, у нас есть два дифференцирования в тензорной алгебре $\mathcal{T}(\mathbb{M})$: производная Ли вдоль векторного поля L_X и дифференцирование D_S , индуцированное тензорным полем S типа $(1, 1)$. В разделе 5.3 будет введено еще одно дифференцирование в тензорной алгебре – ковариантное дифференцирование ∇ . Согласно теореме 2.14.1, эти три вида дифференцирований связаны между собой. Нетрудно убедиться, что в общем случае производная Ли L_X и ковариантная производная вдоль векторного поля ∇_X связаны следующим соотношением

$$\begin{aligned} (L_X - \nabla_X)Y_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r} = \\ = (\nabla_{\beta_1} X^\gamma + T_{\beta_1 \delta}^\gamma X^\delta)Y_{\gamma \beta_2 \dots \beta_s}^{\alpha_1 \dots \alpha_r} + \dots + (\nabla_{\beta_s} X^\gamma + T_{\beta_s \delta}^\gamma X^\delta)Y_{\beta_1 \dots \beta_{s-1} \gamma}^{\alpha_1 \dots \alpha_r} - \\ - Y_{\beta_1 \dots \beta_s}^{\gamma \alpha_2 \dots \alpha_r} (\nabla_\gamma X^{\alpha_1} + T_{\gamma \delta}^{\alpha_1} X^\delta) - \dots - Y_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_{r-1} \gamma} (\nabla_\gamma X^{\alpha_r} + T_{\gamma \delta}^{\alpha_r} X^\delta). \end{aligned}$$

То есть разность ковариантного дифференцирования и производной Ли является дифференцированием, индуцированным тензорным полем с компонентами $S_\beta^\alpha := \nabla_\beta X^\alpha + T_{\beta \gamma}^\alpha X^\gamma$.

Как следствие этой формулы или непосредственной проверкой можно убедиться в том, что в правой части (2.122) все частные производные можно выразить через ковариантные. Это естественно, т.к. производная Ли является инвариантным оператором. В этом случае появляются дополнительные слагаемые, содержащие тензор кручения для каждого тензорного индекса. Если кручение аффинной связности равно нулю, то все частные производные в производной Ли (2.122) можно просто заменить на ковариантные.

Пример 2.14.4. Для иллюстрации приведем формулу производной Ли от компонент тензора Y_α^β типа $(1, 1)$

$$\begin{aligned} L_X Y_\alpha^\beta &= X^\gamma \partial_\gamma Y_\alpha^\beta + \partial_\alpha X^\gamma Y_\gamma^\beta - Y_\alpha^\gamma \partial_\gamma X^\beta = \\ &= X^\gamma \nabla_\gamma Y_\alpha^\beta + \nabla_\alpha X^\gamma Y_\gamma^\beta - Y_\alpha^\gamma \nabla_\gamma X^\beta + X^\gamma Y_\alpha^\delta T_{\delta \gamma}^\beta - X^\gamma Y_\delta^\beta T_{\alpha \gamma}^\delta = \\ &= X^\gamma \tilde{\nabla}_\gamma Y_\alpha^\beta + \tilde{\nabla}_\alpha X^\gamma Y_\gamma^\beta - Y_\alpha^\gamma \tilde{\nabla}_\gamma X^\beta. \end{aligned}$$

Аналогичные формулы справедливы для произвольных тензорных полей. $\square$

Замечание. Производная Ли не зависит ни от метрики, ни от аффинной связности, которые могут быть заданы на многообразии совершенно независимо. $\square$

Пример 2.14.5. Приведем также явное выражение для производной Ли от метрики

$$L_X g_{\alpha\beta} = X^\gamma \partial_\gamma g_{\alpha\beta} + \partial_\alpha X^\gamma g_{\gamma\beta} + \partial_\beta X^\gamma g_{\alpha\gamma}.$$

Это выражение можно переписать в виде

$$L_X g_{\alpha\beta} = \tilde{\nabla}_\alpha X_\beta + \tilde{\nabla}_\beta X_\alpha, \quad (2.127)$$

где $\tilde{\nabla}_\alpha$ – ковариантная производная со связностью Леви–Чивиты (раздел 6.2). $\square$

Определение. Пусть $T \in \mathcal{T}_s^r(\mathbb{M})$ и $X \in \mathcal{X}(\mathbb{M})$ – произвольное тензорное поле типа (r, s) и полное векторное поле на многообразии $\mathbb{M}$. Полное векторное поле порождает однопараметрическую группу преобразований $s(t, x)$. Если значение тензорного поля $T(s)$ в точке $s(t, x)$ равно $\tilde{s}(t, x)T(x)$, то мы говорим, что тензорное поле $T(x)$ *инвариантно* при действии однопараметрической группы преобразований. $\square$

Из определения производной Ли следует

Предложение 2.14.3. *Тензорное поле $T \in \mathcal{T}_s^r(\mathbb{M})$ инвариантно относительно однопараметрической группы преобразований $s(t, x)$, порожденной векторным полем X , для всех t тогда и только тогда, когда производная Ли равна нулю, $\mathcal{L}_X T = 0$.*

Глава 3

Дифференциальные формы и интегрирование

В дифференциальной геометрии и приложениях важную роль играют дифференциальные формы. Они используются для определения связностей на расслоениях, интегрирования по многообразию, характеристических классов, когомологий де Рама, в гамильтоновом формализме и других областях. В настоящей главе будут даны определения и рассмотрены основные свойства дифференциальных форм. Будут также сформулированы две фундаментальные теоремы: теорема Дарбу и формула Стокса, играющие исключительно важную роль в приложениях.

3.1 Внешняя алгебра

Рассмотрим векторное и сопряженное к нему пространства $\mathbb{V}$ и $\mathbb{V}^*$. В разделе 1.3.3 над произвольным векторным пространством, например, над $\mathbb{V}^*$, была построена тензорная алгебра $\otimes \mathbb{V}^*$. В этой алгебре существуют двусторонние идеалы, и соответствующие фактор пространства также представляют собой алгебры. Наиболее важными из этих фактор алгебр являются внешняя алгебра и алгебры Клиффорда, которые будут рассмотрены в главе 26. В настоящем разделе мы рассмотрим внешнюю алгебру $\Lambda(\mathbb{V})$, лежащую в основе теории дифференциальных форм.

Рассмотрим n -мерное векторное пространство $\mathbb{V}$ над полем вещественных чисел с базисом e_a , $a = 1, \dots, n$. Дуальный базис сопряженного пространства $\mathbb{V}^*$ обозначим через e^a , $e^a(e_b) = \delta_b^a$. В следующем разделе мы отождествим пространства $\mathbb{V}$ и $\mathbb{V}^*$ с касательным $\mathbb{T}_x(\mathbb{M})$ и кокасательным $\mathbb{T}_x^*(\mathbb{M})$ пространствами в фиксированной точке $x \in \mathbb{M}$ многообразия $\mathbb{M}$. Пока же под $\mathbb{V}$ будем понимать абстрактное векторное пространство.

Начнем с длинного, но конструктивного определения внешней алгебры над векторным пространством $\mathbb{V}$. Сначала введем вспомогательные понятия.

Определение. Множество полностью антисимметричных ковариантных тензоров типа $(0, r)$, $0 \leq r \leq n$, называется множеством *форм степени r* над векторным пространством $\mathbb{V}$ и обозначается $\Lambda_r(\mathbb{V})$.

В данном определении антисимметрия формы означает следующее. Пусть $A \in \Lambda_r(\mathbb{V})$. Тогда ее значение на произвольном наборе векторов $X_1, \dots, X_r \in \mathbb{V}$ антисимметрично относительно перестановки любой пары векторов:

$$A(X_1, \dots, X_i, \dots, X_j, \dots, X_r) = -A(X_1, \dots, X_j, \dots, X_i, \dots, X_r), \quad 1 \leq i < j \leq r.$$

В компонентах это условие означает антисимметрию относительно перестановки любой пары индексов.

Множество форм $\Lambda_r(\mathbb{V}) \subset \otimes \mathbb{V}^*$ степени r является векторным пространством над полем вещественных чисел размерности C_n^r .

По определению, вещественные числа (скаляры) образуют множество 0-форм, $\Lambda_0(\mathbb{V}) = \mathbb{R}$. Отметим, что 1-формы $A \in \Lambda_1(\mathbb{V}) = \mathbb{V}^*$ (ковекторы или линейные формы) являются элементами сопряженного пространства и представляются в виде $A = e^a A_a$. Максимальная степень формы совпадает с размерностью векторного пространства n , поскольку для более высоких степеней по крайней мере два индекса будут совпадать, а это невозможно для антисимметричных тензоров.

Из определения тензора следует, что r -форма $A \in \Lambda_r(\mathbb{V})$ представляет собой полилинейное отображение r экземпляров векторного пространства $\mathbb{V}$ в вещественную прямую

$$A : \underbrace{\mathbb{V} \times \dots \times \mathbb{V}}_r \rightarrow \mathbb{R}.$$

Полилинейность означает, что значение r -формы на r векторах $X_1, \dots, X_r \in \mathbb{V}$ линейно по каждому аргументу при фиксированных остальных.

Ввиду антисимметрии r -форм относительно перестановки индексов, они имеют меньше независимых компонент, чем ковариантные тензоры типа $(0, r)$. Введем обозначение

$$e^{a_1} \wedge \dots \wedge e^{a_r} := \sum_{\sigma} e^{\sigma(a_1)} \otimes \dots \otimes e^{a_r} \operatorname{sgn} \sigma, \quad (3.1)$$

где символ $\otimes$ обозначает тензорное произведение, сумма берется по всем перестановкам $\sigma(a_1 \dots a_r)$ индексов, а $\operatorname{sgn} \sigma = \pm 1$ обозначает знак перестановки. Здесь мы ввели знак внешнего умножения $\wedge$, которое будет определено ниже. Пока он рассматривается просто, как некоторый символ.

По построению, выражение (3.1) является ковектором типа $(0, r)$ и антисимметрично относительно перестановки любой пары индексов:

$$e^{\sigma(a_1)} \wedge \dots \wedge e^{a_r} = e^{a_1} \wedge \dots \wedge e^{a_r} \operatorname{sgn} \sigma.$$

Поэтому множество ковекторов с упорядоченным набором индексов

$$e^{a_1} \wedge \dots \wedge e^{a_r}, \quad 1 \leq a_1 < \dots < a_r \leq n, \quad (3.2)$$

образует базис в пространстве r -форм. Тогда произвольная форма степени r имеет вид

$$A = \sum_{a_1 < \dots < a_r} e^{a_1} \wedge \dots \wedge e^{a_r} A_{a_1 \dots a_r} = \quad (3.3)$$

$$= \frac{1}{r!} e^{a_1} \wedge \dots \wedge e^{a_r} A_{a_1 \dots a_r}. \quad (3.4)$$

В разложении (3.3) при фиксированной выборке индексов суммирование по индексам отсутствует, а сумма подразумевается только по различным выборкам. В последнем выражении (3.4) суммирование проводится по всем значениям индексов, и поэтому введен компенсирующий множитель $1/r!$. Это часто бывает удобнее, поскольку запись r -формы A в виде (3.4) имеет инвариантный вид в отличие от (3.3).

Пример 3.1.1. Рассмотрим 2-форму $A = e^a \otimes e^b A_{ab}$, где $A_{ab} = -A_{ba}$. Тогда

$$A = \frac{1}{2} e^a \otimes e^b (A_{ab} - A_{ba}) = \frac{1}{2} e^a \wedge e^b A_{ab},$$

где

$$e^a \wedge e^b = -e^b \wedge e^a = e^a \otimes e^b - e^b \otimes e^a. \quad \square \quad (3.5)$$

В компонентах значение r -формы A на векторных полях $\{X_1, \dots, X_r\}$ определяется сверткой компонент:

$$A(X_1, \dots, X_r) := X_1^{a_1} \dots X_r^{a_r} A_{a_1 \dots a_r}.$$

Построим 2^n -мерное векторное пространство, равное прямой сумме форм всех степеней

$$\Lambda(\mathbb{V}) := \bigoplus_{r=0}^n \Lambda_r(\mathbb{V}). \quad (3.6)$$

Базис этого пространства имеет вид

$$1, \quad e^a, \quad e^{a_1} \wedge e^{a_2}, \quad \dots, \quad e^{a_1} \wedge \dots \wedge e^{a_r}, \dots, \quad e^1 \wedge \dots \wedge e^n, \quad (3.7)$$

где последовательность $a_1 \dots a_r$ является упорядоченным набором различных индексов, расположенных в порядке возрастания $1 \leq a_1 < \dots < a_r \leq n$. Первый базисный вектор является единицей, поскольку поле вещественных чисел $\mathbb{R}$ является подпространством рассматриваемого векторного пространства. Последний базисный вектор единственен, т.к. содержит ровно n индексов. Элементы из $\Lambda(\mathbb{V})$ представимы в виде

$$A = \sum_{r=0}^n \frac{1}{r!} e^{a_1} \wedge \dots \wedge e^{a_r} A_{a_1 \dots a_r}, \quad (3.8)$$

где подразумевается суммирование по всем значениям индексов.

Введем на множестве форм *внешнее умножение*, отображающее $\Lambda(\mathbb{V}) \rightarrow \Lambda(\mathbb{V})$. Пусть A и B – формы фиксированных степеней r и s , соответственно. Их тензорное произведение $A \otimes B$ является ковариантным тензором типа $(0, r+s)$, однако не будет формой, т.к. не будет антисимметрично относительно всех перестановок индексов. Чтобы исправить ситуацию после тензорного умножения форм необходимо произвести полную антисимметризацию. Тогда внешнее умножение станет отображением $\Lambda_r \times \Lambda_s \rightarrow \Lambda_{r+s}$.

Определение. Внешним произведением двух форм $A \in \Lambda_r(\mathbb{V})$ и $B \in \Lambda_s(\mathbb{V})$ называется форма $C \in \Lambda_{r+s}(\mathbb{V})$, построенная по правилу:

$$A \wedge B = \frac{1}{(r+s)!} e^{a_1} \wedge \dots \wedge e^{a_{r+s}} C_{a_1 \dots a_{r+s}}, \quad (3.9)$$

где компоненты внешнего произведения равны

$$\begin{aligned} C_{a_1 \dots a_{r+s}} &:= \frac{1}{r!s!} \sum_{\sigma} A_{\sigma(a_1 \dots a_r)} B_{a_{r+1} \dots a_{r+s}} \operatorname{sgn} \sigma = \\ &= C_{r+s}^r A_{[a_1 \dots a_r} B_{a_{r+1} \dots a_{r+s}]} \end{aligned} \quad (3.10)$$

Здесь $\sigma(a_1 \dots a_r a_{r+1} \dots a_{r+s})$ обозначает перестановку σ всех индексов $(a_1 \dots a_{r+s})$, $\operatorname{sgn} \sigma$ – знак перестановки, суммирование ведется по всем перестановкам, а квадратные скобки обозначают антисимметризацию всех индексов. $\square$

Формулу внешнего произведения (3.9) можно переписать в эквивалентном виде

$$A \wedge B = \frac{1}{r!s!} e^{a_1} \wedge \dots \wedge e^{a_{r+s}} A_{a_1 \dots a_r} B_{a_{r+1} \dots a_{r+s}}. \quad (3.11)$$

Замечание. Множитель $1/(r+s)!$ в (3.9) связан с тем, что суммирование проводится по всем значениям индексов. Необходимость введения множителя $1/r!s!$ в (3.10) будет ясна из дальнейшего. $\square$

Пример 3.1.2. Внешнее умножение произвольной r -формы $A \in \Lambda_r(\mathbb{V})$ на 0-форму (число) $f \in \mathbb{R} = \Lambda_0(\mathbb{V})$ сводится к обычному умножению компонент формы A на число,

$$f \wedge A = \frac{1}{r!} e^{a_1} \wedge \dots \wedge e^{a_r} (f A_{a_1 \dots a_r}). \quad \square$$

Пример 3.1.3. Внешнее произведение двух 1-форм $A = e^a A_a$ и $B = e^b B_b$ равно

$$A \wedge B = \frac{1}{2} e^a \wedge e^b (A_a B_b - A_b B_a) = e^a \wedge e^b A_a B_b. \quad \square \quad (3.12)$$

Пример 3.1.4. Внешнее произведение 1-формы $A = e^a A_a$ на 2-форму $B = \frac{1}{2} e^a \wedge e^b B_{ab}$ равно

$$\begin{aligned} A \wedge B &= \frac{1}{6} e^a \wedge e^b \wedge e^c (A_a B_{bc} + A_b B_{ca} + A_c B_{ab}) = \\ &= \frac{1}{2} e^a \wedge e^b \wedge e^c A_a B_{bc}. \quad \square \end{aligned} \quad (3.13)$$

Пример 3.1.5. Внешнее произведение 2-формы $A = \frac{1}{2} e^2 \wedge e^b A_{ab}$ на 2-форму $B = \frac{1}{2} e^a \wedge e^b B_{ab}$ равно

$$A \wedge B = \frac{1}{4} e^a \wedge e^b \wedge e^c \wedge e^d A_{ab} B_{cd}. \quad \square \quad (3.14)$$

Если суммарная степень форм превосходит размерность векторного пространства $r + s > n$, то внешнее произведение этих форм дает нуль.

Можно проверить, что внешнее умножение обладает следующими свойствами:

$$A \wedge B \quad \text{линейно по } A \text{ и } B, \quad (3.15)$$

$$A \wedge B = (-1)^{rs} B \wedge A, \quad (3.16)$$

$$(A \wedge B) \wedge C = A \wedge (B \wedge C). \quad (3.17)$$

Первые два свойства очевидны. Из второго свойства (3.16) следует, что, если A – форма нечетной степени, то $A \wedge A = 0$.

Третье свойство доказывается прямым вычислением. Для трех форм, степеней r, s и t , соответственно, справедливо равенство

$$\begin{aligned} (A \wedge B) \wedge C &= A \wedge (B \wedge C) = A \wedge B \wedge C = \\ &= \frac{1}{(r+s+t)!} e^{a_1} \wedge \dots \wedge e^{a_{r+s+t}} \frac{1}{r!s!t!} \sum_{\sigma} \text{sgn}(\sigma) A_{\sigma(a_1 \dots a_r)} B_{\sigma(a_{r+1} \dots a_{r+s})} C_{\sigma(a_{r+s+1} \dots a_{r+s+t})} = \\ &= \frac{1}{r!s!t!} e^{a_1} \wedge \dots \wedge e^{a_{r+s+t}} A_{a_1 \dots a_r} B_{a_{r+1} \dots a_{r+s}} C_{a_{r+s+1} \dots a_{r+s+t}}. \end{aligned}$$

Последнее выражение симметрично относительно перестановок r, s и t , что соответствует ассоциативности внешнего умножения и определяют выбор коэффициента в равенстве (3.10): они подобраны таким образом, чтобы внешнее умножение было ассоциативным.

Определение. Форма фиксированной степени называется *однородной*. Прямая сумма отличных от нуля форм разных степеней называется *неоднородной*. $\square$

Внешнее умножение (3.9), определенное для однородных форм, продолжается на неоднородные формы общего вида (3.8) по линейности. Тем самым множество форм $\Lambda(\mathbb{V})$ над произвольным векторным пространством $\mathbb{V}$, $\dim \mathbb{V} = n$, с операцией внешнего умножения представляет собой ассоциативную алгебру над полем вещественных чисел.

Пример 3.1.6. Пусть A, B, C, D – четыре однородные формы различных степеней, тогда

$$(A \oplus B) \wedge (C \oplus D) := (A \wedge C) \oplus (B \wedge C) \oplus (A \wedge D) \oplus (B \wedge D). \quad \square$$

Определение. Множество форм вида (3.8) с внешним умножением (3.9), продолженным на формы общего вида по линейности, называется *внешней алгеброй* $\Lambda(\mathbb{V})$ над векторным пространством $\mathbb{V}$, $\dim \mathbb{V} = n$. $\square$

Эта алгебра ассоциативна, градуирована, антикоммутативна и содержит единичный элемент $1 \in \Lambda_0(\mathbb{V})$. Внешняя алгебра является алгеброй Грассмана [51] с образующими e^a (частным случаем алгебр Клиффорда, которые рассмотрены в главе 26).

В обозначении базисных векторов внешней алгебры (3.2) был использован знак внешнего умножения, которое было определено позже. Покажем, что это обозначение обосновано. Рассмотрим две базисные 1-формы $A^a = e^a$ и $B^b = e^b$. Тогда их внешнее произведение равно

$$A^a \wedge B^b = e^a \wedge e^b. \quad (3.18)$$

Более общо, для двух базисных форм $A^{a_1 \dots a_r} = e^{a_1} \wedge \dots \wedge e^{a_r}$ и $B^{b_1 \dots b_s} = e^{b_1} \wedge \dots \wedge e^{b_s}$ справедливо равенство

$$A^{a_1 \dots a_r} \wedge B^{b_1 \dots b_s} = e^{a_1} \wedge \dots \wedge e^{a_r} \wedge e^{b_1} \wedge \dots \wedge e^{b_s}.$$

Это оправдывает обозначение для базиса (3.1) и выбор коэффициентов для внешнего умножения (3.10).

Замечание. Для определения ассоциативного умножения в пространстве форм $\Lambda(\mathbb{V})$ выбор коэффициента в определении внешнего умножения не является единственным. При определении внешнего умножения (3.9), (3.10) можно было бы выбрать другой коэффициент в (3.10):

$$C_{a_1 \dots a_{r+s}} := \frac{1}{(r+s)!} \sum_{\sigma} \text{sgn } \sigma A_{\sigma(a_1 \dots a_r)} B_{a_{r+1} \dots a_{r+s}}. \quad (3.19)$$

В этом случае произведение трех форм степеней r, s и t

$$\begin{aligned} (A \wedge B) \wedge C &= A \wedge (B \wedge C) = A \wedge B \wedge C = \\ &= \frac{1}{(r+s+t)!} e^{a_1} \wedge \dots \wedge e^{a_{r+s+t}} A_{a_1 \dots a_r} B_{a_{r+1} \dots a_{r+s}} C_{a_{r+s+1} \dots a_{r+s+t}} \end{aligned}$$

также симметрично относительно перестановок r, s, t и, значит, ассоциативно. Однако в произведении базисных векторов (3.18) появился бы дополнительный множитель. $\square$

К определению внешней алгебры можно подойти с другой, более абстрактной, точки зрения. Рассмотрим две произвольные 1-формы $A, B \in \Lambda_1(\mathbb{V}) = \mathbb{V}^*$. Тогда для их тензорного произведения справедливо тождество

$$A \otimes B = A \wedge B + \frac{1}{2}[(A + B) \otimes (A + B) - A \otimes A - B \otimes B]. \quad (3.20)$$

Выражение в квадратных скобках является элементом тензорной алгебры, и каждое слагаемое представляет собой тензорное произведение одинаковых 1-форм. Рассмотрим элементы более общего вида $X \otimes A \otimes A \otimes Y$, где $A \in \mathbb{V}^*$ – линейная форма, а $X, Y \in \otimes \mathbb{V}^*$ – произвольные элементы тензорной алгебры. Нетрудно проверить, что элементы такого вида образуют двусторонний идеал $\mathbb{I}$ в тензорной алгебре $\otimes \mathbb{V}^*$, которая рассматривается, как кольцо по отношению к сложению и тензорному умножению. Тогда из представления (3.20) следует, что внешняя алгебра является факторпространством

$$\Lambda(\mathbb{V}) = \otimes \mathbb{V}^* / \mathbb{I}.$$

То есть два элемента $A, B \in \otimes \mathbb{V}^*$ являются эквивалентными, $A \sim B$, если $A = B + C$, где $C \in \mathbb{I}$. Выражение в квадратных скобках в (3.20) принадлежит идеалу, и, значит,

$$A \otimes B \sim A \wedge B.$$

Таким образом каждый элемент внешней алгебры взаимно однозначно определяет класс эквивалентности в тензорной алгебре.

Сформулируем четыре утверждения, доказательства которых сводятся к алгебраическим выкладкам.

Предложение 3.1.1. *Для того, чтобы набор 1-форм $\{A_i\}$, $i = 1, \dots, N$, был линейно зависим, необходимо и достаточно, чтобы выполнялось равенство*

$$A_1 \wedge \dots \wedge A_N = 0. \quad (3.21)$$

Предложение 3.1.2. *Пусть $\{A_i\}$ и $\{B_i\}$, $i = 1, \dots, N$, – два набора 1-форм таких, что выполнено равенство*

$$\sum_{i=1}^N A_i \wedge B_i = 0.$$

Если 1-формы $\{A_i\}$ линейно независимы, то 1-формы $\{B_i\}$ можно выразить в виде линейной комбинации 1-форм $\{A_i\}$:

$$B_i = \sum_{j=1}^N C_{ij} A_j,$$

причем $C_{ij} = C_{ji}$.

Предложение 3.1.3. *Пусть $\{A_i\}$, $i = 1, \dots, N$, – набор N линейно независимых 1-форм и $B \in \Lambda_s$. Для того, чтобы s -форму B можно было представить в виде линейной комбинации*

$$B = A_1 \wedge C_1 + \dots + A_N \wedge C_N,$$

где $\{C_i\}$ – некоторый набор $(s-1)$ -форм, необходимо и достаточно, чтобы

$$A_1 \wedge \dots \wedge A_N \wedge B = 0.$$

Замечание. При $N + s > n$ эта теорема тривиально выполняется. $\square$

Предложение 3.1.4. Пусть $\{A_i, B_i\}$ и $\{A'_i, B'_i\}$, $i = 1, \dots, N$, – два набора по $2N$ 1-форм. Если 1-формы $\{A_i, B_i\}$ линейно независимы и выполнено равенство

$$\sum_{i=1}^N A_i \wedge B_i = \sum_{i=1}^N A'_i \wedge B'_i,$$

то набор 1-форм $\{A'_i, B'_i\}$ также линейно независим и представляется в виде линейных комбинаций 1-форм $\{A_i, B_i\}$.

Пусть некоторое семейство 1-форм $\{A_i\}$, $i = 1, \dots, N$, выражается в виде линейной комбинации 1-форм $\{B_i\}$:

$$A_i = M_i^j B_j,$$

где M_i^j – некоторая квадратная $N \times N$ матрица. Тогда нетрудно проверить следующую формулу

$$A_1 \wedge \dots \wedge A_N = \det(M_i^j) B_1 \wedge \dots \wedge B_N.$$

Определение. Пусть $A \in \Lambda_r(\mathbb{V})$ – r -форма, тогда внешнее произведение

$$A^s := \underbrace{A \wedge \dots \wedge A}_s$$

называется *внешней степенью* r -формы A .

На множестве 2-форм $\Lambda_2(\mathbb{V})$ можно определить *внешнюю экспоненту* $\widehat{\exp}$, которая каждой 2-форме $A \in \Lambda_2(\mathbb{V})$ ставит в соответствие прямую сумму форм четной степени по правилу

$$\widehat{\exp} A := 1 \oplus A \oplus \frac{1}{2} A \wedge A \oplus \frac{1}{3!} A \wedge A \wedge A \oplus \dots \quad \square \quad (3.22)$$

Ряд (3.22) содержит конечное число слагаемых, т.к. максимальное число сомножителей в последнем отличном от нуля слагаемом не может превосходить $n/2$. Внешняя экспонента обладает следующими свойствами:

$$\widehat{\exp} A \wedge \widehat{\exp} B = \widehat{\exp} (A + B), \quad \forall A, B \in \Lambda_2(\mathbb{V}).$$

$$\frac{d}{dt} \widehat{\exp} (tA) = A \wedge \widehat{\exp} (tA) = \widehat{\exp} (tA) \wedge A.$$

При доказательстве этих формул использовано равенство $A \wedge B = B \wedge A$, справедливое для форм четной степени.

Определение. В пространстве векторов и r -форм введем *внутреннее умножение* как билинейное отображение

$$i: \quad \mathbb{V} \times \Lambda_r(\mathbb{V}) \ni (X, A) \mapsto i_X A \in \Lambda_{r-1}(\mathbb{V}).$$

Оно определяется следующим образом. Пусть $X \in \mathbb{V}$ и $A \in \Lambda_r(\mathbb{V})$. Положим

$$i_X A(X_1, \dots, X_{r-1}) := A(X, X_1, \dots, X_{r-1}). \quad (3.23)$$

Другими словами, компоненты вектора X нужно просто свернуть с первым индексом r -формы. Правая часть выражения (3.23) $(r-1)$ -линейна и антисимметрична, как функция $X_1, \dots, X_{r-1}$, а также билинейна по X и A . Удобно считать, что $i_X A = 0$, если $A \in \Lambda_0$ – форма нулевой степени. Внутреннее умножение обозначают также $i_X A = X \rfloor A$. Определение (3.23) можно переписать в эквивалентном виде

$$i_X A = X \rfloor A = \frac{1}{(r-1)!} e^{a_1} \wedge \dots \wedge e^{a_{r-1}} X^b A_{ba_1 \dots a_{r-1}}. \quad \square \quad (3.24)$$

Внутреннее умножение линейно

$$\begin{aligned} i_{X+Y} &= i_X + i_Y, \\ i_{aX} &= a i_X, \quad a \in \mathbb{R}, \end{aligned}$$

и антисимметрично

$$i_X i_Y = -i_Y i_X \quad \Leftrightarrow \quad i_X i_X = 0,$$

для всех $X, Y \in \mathbb{V}$.

Нетрудно проверить, что внутреннее умножение удовлетворяет правилу Лейбница

$$i_X(A \wedge B) = i_X A \wedge B + (-1)^r A \wedge i_X B, \quad A \in \Lambda_r(\mathbb{V}).$$

Поэтому внутреннее умножение называют также *внутренним дифференцированием*.

Внутреннее умножение удобно использовать для характеристики внешней алгебры, построенной на фактор пространстве $\mathbb{V}/\mathbb{U}$, где $\mathbb{U} \subset \mathbb{V}$ – некоторое линейное подпространство в $\mathbb{V}$. А именно, внешняя алгебра $\Lambda(\mathbb{V}/\mathbb{U})$ состоит из тех форм $A \in \Lambda(\mathbb{V})$, для которых выполнено равенство $i_X A = 0$ для всех $X \in \mathbb{U}$. Верно и обратное утверждение. Пусть задано подпространство $\mathbb{U} \subset \mathbb{V}$. Тогда множество векторов X , удовлетворяющих условию $i_X A = 0$ для всех $A \in \Lambda(\mathbb{V}/\mathbb{U})$ образует подпространство, которое совпадает с $\mathbb{U}$.

В заключение обсудим свойства внешнего умножения при отображении векторных пространств. Пусть f – линейное отображение векторных пространств

$$f: \mathbb{V} \rightarrow \mathbb{W},$$

и пусть на векторном пространстве-мишени $\mathbb{W}$ заданы две формы $A, B \in \Lambda(\mathbb{W})$, тогда для возврата отображения справедливо равенство

$$f^*(A \wedge B) = (f^*A) \wedge (f^*B) \in \Lambda(\mathbb{V}),$$

которое просто проверяется.

3.2 Дифференциальные формы

Рассмотрим многообразие $\mathbb{M}$ размерности $\dim \mathbb{M} = n$. В каждой точке многообразия мы имеем касательное и кокасательное векторные пространства. Построим над касательным векторным пространством $\mathbb{V} = \mathbb{T}_x(\mathbb{M})$ с координатным базисом ∂_α в каждой точке $x \in \mathbb{M}$ векторное пространство форм степени r : $\Lambda_{r,x}(\mathbb{M})$, как это было сделано в предыдущем разделе. Объединение форм степени r по всем точкам многообразия $\bigcup_{x \in \mathbb{M}} \Lambda_{r,x}(\mathbb{M})$ приведет к векторному расслоению, сечениями которого являются полностью антисимметричные ковариантные тензорные поля ранга r (дифференциальные формы). Очевидно, что множество дифференциальных форм фиксированной степени с поточечным сложением и умножением на числа образует бесконечномерное линейное пространство. Прямая сумма по r множества всех дифференциальных форм даст внешнюю алгебру дифференциальных форм $\Lambda(\mathbb{M})$ на многообразии $\mathbb{M}$, в которой все операции (умножение на числа, сложение и внешнее умножение) определены поточечно. Ясно, что все свойства внешнего умножения без изменений переносятся на внешнюю алгебру дифференциальных форм, и мы не будем их переписывать. Отметим изменение обозначений и терминологии.

Определение. Множество полностью антисимметричных ковариантных тензорных полей типа $(0, r)$, $0 \leq r \leq n$ называется множеством *дифференциальных форм степени r* на многообразии $\mathbb{M}$ и обозначается $\Lambda_r(\mathbb{M}) \subset \mathcal{T}_r(\mathbb{M})$. $\square$

В дальнейшем дифференциальные формы на многообразии $\mathbb{M}$ степени r мы, для краткости, будем называть просто r -формами.

По определению, гладкие функции на многообразии образуют множество гладких 0-форм, $\Lambda_0(\mathbb{M}) = \mathcal{C}^\infty(\mathbb{M})$. 1-формы $A \in \Lambda_1(\mathbb{M}) = \mathcal{T}_1(\mathbb{M})$ (ковариантные векторные поля = сечения кокасательных расслоений) представляются в виде

$$A(x) = dx^\alpha A_\alpha(x),$$

где дифференциалы dx^α представляют собой координатный базис кокасательного пространства, который дуален к ∂_α . Дифференциальные 1-формы называются также *формами Пфаффа*. Максимальная степень нетривиальной формы совпадает с размерностью многообразия n , поскольку для более высоких степеней по крайней мере два индекса будут совпадать, а это невозможно для антисимметричных тензоров.

Замечание. Напомним, что ∂_α это сокращенное обозначение набора векторных полей $(\partial_\alpha)_p \in \mathcal{X}(\mathbb{U})$, $p \in \mathbb{U}$, на карте $(\mathbb{U}, \varphi)$ многообразия $\mathbb{M}$ (см. раздел 2.6.4), а $dx^\alpha = (dx^\alpha)_p$ – дуальный базис 1-форм, $dx^\alpha(\partial_\beta) = \delta_\beta^\alpha$. $\square$

Как и для векторных пространств, алгебра дифференциальных форм равна прямой сумме форм всех степеней

$$\Lambda(\mathbb{M}) := \bigoplus_{r=0}^n \Lambda_r(\mathbb{M}). \quad (3.25)$$

Координатный базис этого пространства имеет вид

$$1, \quad dx^\alpha, \quad dx^{\alpha_1} \wedge dx^{\alpha_2}, \quad \dots, \quad dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r}, \dots, \quad dx^1 \wedge \dots \wedge dx^n, \quad (3.26)$$

где индексы упорядочены по возрастанию $1 \leq \alpha_1 < \dots < \alpha_r \leq n$. Элементы внешней алгебры $\Lambda(\mathbb{M})$, в соответствии с разложением (3.25), представимы в виде

$$A = \sum_{r=0}^n \frac{1}{r!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r} A_{\alpha_1 \dots \alpha_r}, \quad (3.27)$$

где подразумевается суммирование по всем значениям индексов. Форма A называется *дифференцируемой*, если все антисимметричные тензорные поля $A_{\alpha_1 \dots \alpha_r}(x)$ (компоненты формы) дифференцируемы. Если не оговорено противное, мы будем предполагать, что все компоненты форм гладко зависят от точки многообразия.

Замечание. Запись дифференциальной формы в виде (3.27) имеет смысл только в определенной карте многообразия, где определен координатный базис кокасательного пространства dx^α . Однако она универсальна, т.к. инвариантна относительно преобразований координат, поскольку в точках пересечения двух карт с координатами x^α и $x^{\alpha'}$ справедливо равенство

$$dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r} A_{\alpha_1 \dots \alpha_r} = dx^{\alpha'_1} \wedge \dots \wedge dx^{\alpha'_r} A_{\alpha'_1 \dots \alpha'_r}.$$

Запись дифференциальных форм в виде (3.27) является общепринятой. $\square$

Введем несколько новых понятий, которые являются полезными с точки зрения дифференциальной геометрии. Конечно, часть из них можно было бы ввести и для форм над векторным пространством еще в предыдущем разделе.

Можно рассматривать дифференциальные формы со значениями в произвольном векторном пространстве $\mathbb{W}$. Пусть e_i – базис векторного пространства $\mathbb{W}$. Тогда r -форма A со значениями в $\mathbb{W}$ имеет вид $A = A^i e_i$. При каждом значении индекса i коэффициенты этого разложения представляют собой r -формы $A^i \in \Lambda_r(\mathbb{M})$. По индексу i эти формы преобразуются, как компоненты вектора из пространства $\mathbb{W}$.

Пример 3.2.1. Произвольную r -форму на многообразии $\mathbb{M}$

$$A = \frac{1}{r!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r} A_{\alpha_1 \dots \alpha_r} \in \Lambda_r(\mathbb{M}),$$

можно поточечно рассматривать, как внешнее произведение базисной r -формы $dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r}$ на 0-форму $A_{\alpha_1 \dots \alpha_r}$ со значениями в векторном пространстве $\mathbb{W} = \mathbb{T}_x^*(\mathbb{M}) \wedge \cdots \wedge \mathbb{T}_x^*(\mathbb{M})$. $\square$

Пусть на многообразии $\mathbb{M}$ помимо форм заданы векторные поля. Тогда значение произвольной r -формы A на r векторных полях $X_1, \dots, X_r \in \mathcal{X}(\mathbb{M})$ равно свертке компонент:

$$A(X_1, \dots, X_r) = X_1^{\alpha_1} \cdots X_r^{\alpha_r} A_{\alpha_1 \dots \alpha_r} \in \mathcal{C}^\infty(\mathbb{M}).$$

Данное определение не зависит от выбора карты и потому корректно. Отсюда следует, что произвольная r -форма в каждой точке $x \in \mathbb{M}$ определяет полилинейное отображение ($\mathcal{C}^\infty$ -модуль): $(\mathbb{T}_x(\mathbb{M}))^r \rightarrow \mathbb{R}$, которое антисимметрично относительно перестановки аргументов.

Пример 3.2.2. Значение 1-формы $A = dx^\alpha A_\alpha$ на векторном поле $X = X^\alpha \partial_\alpha$ определяет функцию, которая равна сумме компонент: $A(X) = X^\alpha A_\alpha$. $\square$

Пусть A_i и X_i , $i = 1, \dots, r \leq n$ – произвольный набор 1-форм и векторных полей на $\mathbb{M}$. Тогда из определения внешнего умножения следует, что значение внешнего произведения 1-форм на векторных полях равно определителю матрицы с элементами $A_i(X_j)$

$$(A_1 \wedge \cdots \wedge A_r)(X_1 \dots X_r) = \det(A_i(X_j)).$$

Пусть две 1-формы A и B линейно зависимы, т.е. $A = fB$, где $f \in \Lambda_0(\mathbb{M})$, тогда их внешнее произведение, очевидно, равно нулю. Верно также и обратное утверждение, из которого следует, что, если $A \wedge B \neq 0$ в некоторой области $\mathbb{U} \subset \mathbb{M}$, то 1-формы линейно независимы в каждой точке $\mathbb{U}$. Более общо, справедливо следующее утверждение.

Предложение 3.2.1. Для того, чтобы линейные формы $e^i = dx^\alpha e_\alpha^i$, $i = 1, \dots, r \leq n$ были линейно независимы в области $\mathbb{U} \subset \mathbb{M}$, необходимо и достаточно, чтобы в каждой точке этой области $e^1 \wedge \cdots \wedge e^r \neq 0$.

Дифференциальные формы часто бывает удобно рассматривать в неголономном базисе. Выберем в качестве базиса кокасательного пространства 1-формы $e^a := dx^\alpha e_\alpha^a$, $a = 1, \dots, n$, где $e_\alpha^a(x)$ – поле репера, тогда соответствующий базис r -форм имеет вид

$$e^{a_1} \wedge \cdots \wedge e^{a_r}, \quad 1 \leq a_1 < \cdots < a_r \leq n. \quad (3.28)$$

Мы используем для неголономного базиса те же обозначения, что и для базиса форм над векторным пространством (3.2). Разница заключается в том, что теперь базис $e^a(x)$ зависит от точки многообразия, и это всегда ясно из контекста.

Для того, чтобы 1-формы e^a действительно образовывали базис, необходимо и достаточно, чтобы

$$\det e_\alpha^a \neq 0 \quad \Leftrightarrow \quad e^1 \wedge \dots \wedge e^n \neq 0$$

во всех точках многообразия $\mathbb{M}$, что совпадает с определением репера, который будет введен позже в разделе 5.4. Компоненты разложения r -формы

$$A = \frac{1}{r!} e^{a_1} \wedge \dots \wedge e^{a_r} A_{a_1 \dots a_r}$$

по этому базису связаны с компонентами в голономном базисе простым соотношением

$$A_{a_1 \dots a_r} := e^{\alpha_1}_{a_1} \dots e^{\alpha_r}_{a_r} A_{\alpha_1 \dots \alpha_r},$$

где e^α_a – матрица обратного репера: $e^\alpha_a e^\beta_a = \delta^\beta_\alpha$.

3.3 Внешнее дифференцирование

Во внешней алгебре дифференциальных форм $\Lambda(\mathbb{M})$ на многообразии $\mathbb{M}$ определим *внешнее дифференцирование* как линейное отображение

$$d: \Lambda_r(\mathbb{M}) \rightarrow \Lambda_{r+1}(\mathbb{M}) \quad (3.29)$$

множества r -форм в множество $(r+1)$ -форм. Начнем с локального определения. Пусть в некоторой карте $(\mathbb{U}, \varphi)$ задана r -форма

$$A = \frac{1}{r!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r} A_{\alpha_1 \dots \alpha_r} \in \Lambda_r(\mathbb{U}).$$

Положим

$$\begin{aligned} dA &:= \frac{1}{r!} dA_{\alpha_1 \dots \alpha_r} \wedge dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r} = \\ &= \frac{1}{r!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_{r+1}} \partial_{\alpha_1} A_{\alpha_2 \dots \alpha_{r+1}}. \end{aligned} \quad (3.30)$$

Здесь под $dA_{\alpha_1 \dots \alpha_r} := dx^\alpha \partial_\alpha A_{\alpha_1 \dots \alpha_r}$ понимается обычный дифференциал. Внешнее дифференцирование обладает следующими свойствами:

$$1) \quad \text{если } f \text{ — 0-форма, то } df = dx^\alpha \partial_\alpha f, \quad (3.31)$$

$$2) \quad \text{если } A \text{ — } r\text{-форма, то } d(A \wedge B) = dA \wedge B + (-1)^r A \wedge dB, \quad (3.32)$$

$$3) \quad d(dA) = 0 \quad \text{для любой формы } A. \quad (3.33)$$

Первые два свойства легко проверяются, исходя из определения (3.30). Доказательство третьего свойства сводится к тому, что вторая частная производная $\partial_{\alpha\beta}^2$ симметрична по индексам и дает нуль при свертке с антисимметричным базисным вектором.

Формула для внешнего дифференцирования (3.30) является прямым следствием свойств 1)–3). Действительно, представив r -форму как внешнее произведение 0-формы $A_{\alpha_1 \dots \alpha_n}$ со значениями в векторном пространстве на базисную r -форму $dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_n}$ сразу приходим к (3.30). Поэтому она задает единственное дифференцирование во внешней алгебре, удовлетворяющее условиям 1)–3).

Глобальное определение внешнего дифференцирования сформулируем в виде теоремы.

Теорема 3.3.1 (Внешнее дифференцирование). *Существует единственное линейное отображение (3.29) такое, что выполнены условия (3.31)–(3.33). Оно называется внешним дифференцированием.*

Доказательство. Выше мы доказали существование и единственность внешнего дифференцирования в определенной карте. Осталось доказать, что существование и единственность не зависят от выбора карты. Действительно, компоненты в разложении внешней производной от r -формы (3.30)

$$\partial_{[\alpha_1} A_{\alpha_2 \dots \alpha_{r+1}]},$$

где квадратные скобки обозначают антисимметризацию всех индексов, являются компонентами полностью антисимметричного тензора типа $(0, r+1)$. Это утверждение нетривиально, потому что обычная частная производная от компонент тензора не дает тензора. Для 1-формы имеем следующий закон преобразования частной производной в области пересечения двух карт

$$\partial_{\alpha'} A_{\beta'} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} \frac{\partial x^\beta}{\partial x^{\beta'}} \partial_\alpha A_\beta + \frac{\partial^2 x^\alpha}{\partial x^{\alpha'} \partial x^{\beta'}} A_\alpha.$$

Первое слагаемое соответствует тензорному закону преобразования, а второе – нет. Однако второе слагаемое исчезает при антисимметризации и мы получаем тензорный закон преобразования для внешней производной. Аналогично сокращаются дополнительные слагаемые для произвольной r -формы. Подробности доказательства приведены в [42]. $\square$

Для практических вычислений внешней производной, которые проводятся в координатах, достаточно формулы (3.30).

Пример 3.3.1. Внешний дифференциал от 1-формы $A = dx^\alpha A_\alpha$ равен

$$dA = dA_\alpha \wedge dx^\alpha = \frac{1}{2} dx^\alpha \wedge dx^\beta (\partial_\alpha A_\beta - \partial_\beta A_\alpha) = dx^\alpha \wedge dx^\beta \partial_\alpha A_\beta. \quad \square \quad (3.34)$$

Пример 3.3.2. Внешний дифференциал от 2-формы $B = \frac{1}{2} dx^\alpha \wedge dx^\beta B_{\alpha\beta}$ равен

$$\begin{aligned} dB &= \frac{1}{2} dB_{\alpha\beta} \wedge dx^\alpha \wedge dx^\beta = \frac{1}{2} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \partial_\alpha B_{\beta\gamma} = \\ &= \frac{1}{12} dx^\alpha \wedge dx^\beta \wedge dx^\gamma (\partial_\alpha B_{\beta\gamma} + \partial_\beta B_{\gamma\alpha} + \partial_\gamma B_{\alpha\beta} - \partial_\alpha B_{\gamma\beta} - \partial_\gamma B_{\beta\alpha} - \partial_\beta B_{\alpha\gamma}). \quad \square \end{aligned} \quad (3.35)$$

Формулу внешнего дифференцирования легко запомнить с помощью следующего мнемонического правила. Сравнивая последнее выражение в (3.30) с формулой для внешнего умножения (3.11), можно записать

$$dA = \partial \wedge A.$$

Из определения 3.30 следует, что внешний дифференциал можно записать с помощью ковариантной производной, определяемой символами Кристоффеля или другой произвольной симметричной аффинной связностью

$$dA = \frac{1}{r!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_{r+1}} \tilde{\nabla}_{\alpha_1} A_{\alpha_2 \dots \alpha_{r+1}}. \quad (3.36)$$

Для несимметричной связности возникнут дополнительные слагаемые с тензором кручения.

Приведем некоторые свойства внешнего дифференцирования. С помощью прямых вычислений доказывается формула

$$2dA(X, Y) = X(A(Y)) - Y(A(X)) - A([X, Y]), \quad (3.37)$$

где A – произвольная 1-форма, а X, Y – любые векторные поля на $\mathbb{M}$. В общем случае справедлива

Теорема 3.3.2. Пусть A – гладкая дифференциальная форма степени r и $X_1 \dots X_{r+1}$ – гладкие векторные поля. Тогда значение $r+1$ -формы dA на этих векторных полях может быть найдено по формуле

$$\begin{aligned} dA(X_1, \dots, X_{r+1}) = & \frac{1}{r+1} \sum_i (-1)^{i-1} X_i^\alpha \partial_\alpha A(X_1, \dots, \check{X}_i, \dots, X_{r+1}) + \\ & + \frac{1}{r+1} \sum_{i < j} (-1)^{i+j} A([X_i, X_j], X_1, \dots, \check{X}_i, \dots, \check{X}_j, \dots, X_{r+1}), \end{aligned} \quad (3.38)$$

где символ $\check{X}_i$ означает, что соответствующее векторное поле отсутствует.

Доказательство. Прямая проверка. $\square$

Определение. Форма A степени $r \geq 1$ называется *точной*, если существует такая $(r-1)$ -форма B , что

$$A = dB. \quad (3.39)$$

Форма A степени $r \geq 0$ называется *замкнутой*, если выполнено условие

$$dA = 0. \quad (3.40)$$

По определению, точных 0-форм не существует. $\square$

Предложение 3.3.1. Множество всех замкнутых 0-форм (функций) на связном многообразии состоит из функций, которые постоянны на всем многообразии.

Доказательство. Пусть $f \in C^\infty(\mathbb{M})$. Тогда замкнутость 0-формы в произвольной системе координат означает, что

$$df = 0 \quad \Leftrightarrow \quad \partial_\alpha f = 0.$$

Откуда следует $f = \text{const}$ и в данной карте и на всем многообразии. $\square$

Любая точная форма является замкнутой в силу свойства внешнего дифференцирования (3.33). Обратное утверждение в общем случае неверно. Для замкнутой формы представление (3.39) справедливо только локально, т.е. в некоторой окрестности произвольной точки (лемма Пуанкаре), которая будет сформулирована ниже. При этом глобальное представление может не иметь места. В общем случае множество замкнутых форм больше множества точных форм. В дальнейшем будет показано, что отличие множества замкнутых и точных форм связано с топологическими свойствами многообразий.

Замечание. Если форма A точна, $A = dB$, и принадлежит классу $\mathcal{C}^k$, то это не значит, что форма B принадлежит классу $\mathcal{C}^{k+1}$, потому что в определении внешней производной (3.30) содержится только часть частных производных. Например, если $B \in \Lambda_n(\mathbb{M})$ – форма максимальной степени и класса $\mathcal{C}^1$, то $A = dB = 0$, и, следовательно, принадлежит классу $\mathcal{C}^\infty$, а это ничего не говорит о степени гладкости формы B . В лемме Пуанкаре нам встретится ситуация, когда обе формы A и B одного класса гладкости. $\square$

Пример 3.3.3. Пусть на компактном многообразии $\mathbb{M}$ задана гладкая функция $f \in \mathcal{C}^\infty(\mathbb{M})$. Тогда 1-форма $A = dx^\alpha \partial_\alpha f$ точна, замкнута и имеет по крайней мере два нуля. Наличие нулей следует из теоремы 1.3.4 о том, что на компактном многообразии всякая непрерывная функция принимает свое минимальное и максимальное значения, где все частные производные обращаются в нуль. Этот пример показывает, что нули замкнутых форм связаны с топологией многообразий. $\square$

Предложение 3.3.2. Внешнее произведение замкнутых форм замкнуто, а если одна из форм, кроме того, точна, то это произведение является точной формой.

Доказательство. Прямая проверка. $\square$

Приведем важную теорему, имеющую многочисленные применения в дифференциальной геометрии.

Теорема 3.3.3 (Лемма Пуанкаре). Пусть $A \in \Lambda_r(\mathbb{R}^n)$, $r \geq 1$, – замкнутая r -форма, $dA = 0$, класса $\mathcal{C}^1$, задана в евклидовом пространстве $\mathbb{R}^n$. Тогда существуют $(r-1)$ -формы $B \in \Lambda_{r-1}(\mathbb{R}^n)$, класса $\mathcal{C}^1$, такие, что $A = dB$. При этом все формы B отличаются друг от друга на точные формы $B_1 = B_2 + dC$, где $C \in \Lambda_{r-2}(\mathbb{R}^n)$, класса $\mathcal{C}^1$ при $r \geq 2$. При $r = 1$, 0-формы B определены с точностью до константы.

Доказательство. Доказательство проводится по индукции по размерности евклидова пространства путем явного построения формы B с помощью интеграла вдоль кривой [22], глава 6, теорема 19. $\square$

Эта теорема локальна, т.к. все евклидово пространство диффеоморфно, например, шару $\mathbb{B}^n \hookrightarrow \mathbb{R}^n$. В случае, когда многообразие $\mathbb{M}$ нетривиально и не покрывается одной картой, лемма Пуанкаре справедлива для каждой карты, но не для всего $\mathbb{M}$. Отличие классов замкнутых и точных форм служит для глобальной характеристики многообразий и лежит в основе теории кохомологий де Рама – одного из подходов к теории гомологий.

Замечание. Согласно лемме Пуанкаре, если форма A определена на области $U \subset \mathbb{M}$, гомеоморфной евклидову пространству $\mathbb{R}^n$, то она всегда точна. Поэтому различие между точными и замкнутыми формами проявляется только для форм, заданных на всем многообразии $\mathbb{M}$. $\square$

Пример 3.3.4. Поясним лемму Пуанкаре на примере 1-формы A и 0-формы f . В этом случае условие точности формы (3.39) сводится к уравнению

$$A_\alpha = \partial_\alpha f. \quad (3.41)$$

Из теории дифференциальных уравнений известно, что это уравнение локально разрешимо тогда и только тогда, когда

$$\partial_\alpha A_\beta - \partial_\beta A_\alpha = 0 \quad \text{или} \quad dA = 0. \quad (3.42)$$

Таким образом, замкнутость некоторой формы обеспечивает ее точность только локально, но не глобально.

Дифференциальная 1-форма $dx^\alpha A_\alpha$ на n -мерном многообразии задается n независимыми компонентами, а скалярное поле f – только одной. С другой стороны, если форма A точна, то она параметризуется только одной функцией, т.к. выполнено равенство (3.41). На первый взгляд, число уравнений на компоненты A_α в (3.42) равно C_n^2 , что превышает разность в числе независимых компонент, которая равна $n - 1$, при $n \geq 3$. Кажущееся противоречие заключается в наивном подсчете числа уравнений. Подсчитаем число независимых уравнений в (3.42). Общее число уравнений равно C_n^2 , однако независимых уравнений меньше, поскольку имеется дифференциальное тождество $d^2 A = 0$. Число этих тождеств равно C_n^3 , однако среди них также есть зависящие в силу тождества $d^3 A = 0$. Продолжая этот процесс до порядка $n - 1$ (дальнейшая антисимметризация по индексам не имеет смысла) получим, что число независимых уравнений среди (3.42) равно

$$C_n^2 - C_n^3 + C_n^4 - \dots (-1)^n C_n^n = C_n^1 - 1 = n - 1.$$

Это равенство следует из формулы бинома Ньютона (37.3). То есть среди n компонент замкнутой 1-формы A (ковекторного поля) только одна является независимой, в качестве которой можно выбрать f . $\square$

Пример 3.3.5. Продемонстрируем локальность леммы Пуанкаре и важность нетривиальных замкнутых форм для описания глобальных свойств многообразий. Рассмотрим евклидову плоскость $\mathbb{R}^2$ в декартовых координатах x, y . Тогда дифференциал полярного угла φ имеет вид (1.80)

$$A = \frac{xdy - ydx}{x^2 + y^2}.$$

1-форма A определена и замкнута, $dA = 0$, всюду, за исключением начала координат, и, значит, не удовлетворяет условиям леммы Пуанкаре. Отсюда следует, что не существует дифференцируемой функции $\varphi(x, y)$, определенной на всей плоскости $\mathbb{R}^2$, такой, что $A = d\varphi$. Действительно, при обходе начала координат по замкнутому контуру один раз, функция φ получит приращение 2π . В то же время на любой односвязной области, не содержащей начала координат, такая функция существует. $\square$

Предложение 3.3.3. Если на многообразии задано векторное поле X , то производная Ли (см. раздел 2.14) от произвольной r -формы $A \in \Lambda_r(\mathbb{M})$ представляет собой симметричную комбинацию внешней и внутренней производной

$$\mathcal{L}_X = d \mathbf{i}_X + \mathbf{i}_X d. \quad (3.43)$$

Это соотношение называется основной формулой гомотопии.

Доказательство. Достаточно показать, что правая часть этого выражения является дифференцированием, коммутирующим с внешней производной d , и что оно совпадает с производной Ли на функциях. Действительно, пусть $A \in \Lambda_r(\mathbb{M})$ и $B \in \Lambda_s(\mathbb{M})$, тогда нетрудно проверить равенство

$$(d \mathbf{i}_X + \mathbf{i}_X d)(A \wedge B) = (d \mathbf{i}_X + \mathbf{i}_X d)A \wedge B + A \wedge (d \mathbf{i}_X + \mathbf{i}_X d)B. \quad (3.44)$$

На функциях $f \in \Lambda_0(\mathbb{M})$

$$(d \mathbf{i}_X + \mathbf{i}_X d)f = \mathbf{i}_X df = Xf = \mathcal{L}_X f.$$

Коммутативность $(d \mathbf{i}_X + \mathbf{i}_X d)$ с внешней производной легко проверяется. $\square$

Формулу (3.44) можно переписать в виде

$$\mathbf{L}_X(A \wedge B) = \mathbf{L}_X A \wedge B + A \wedge \mathbf{L}_X B.$$

Предложение 3.3.4. Пусть задан диффеоморфизм двух многообразий $h : \mathbb{M} \rightarrow \mathbb{N}$. Тогда возврат отображения

$$h^* : \mathbf{T}^*(\mathbb{N}) \rightarrow \mathbf{T}^*(\mathbb{M}),$$

действует на формы в обратную сторону и имеет следующие свойства.

1) Для произвольных форм $A, B \in \Lambda(\mathbb{N})$, заданных на многообразии $\mathbb{N}$, справедливо равенство

$$h^*(A \wedge B) = h^*A \wedge h^*B.$$

2) Возврат отображения коммутирует с внешней производной

$$h^*dA = d(h^*A), \quad \forall A \in \Lambda_r(\mathbb{N}). \quad (3.45)$$

3) Пусть задан набор векторных полей $X_1, \dots, X_r$ на $\mathbb{M}$. Тогда

$$h^*A(X_1, \dots, X_r) = A(h_*X_1, \dots, h_*X_r), \quad \forall A \in \Lambda_r(\mathbb{N}).$$

Доказательство. Прямая проверка. □

Поскольку производная Ли от форм определяется возвратом отображения, то из свойства 1) вытекает, что производная Ли от произвольной r -формы коммутирует с внешним дифференцированием:

$$\mathbf{L}_X d = d \mathbf{L}_X.$$

Это свойство было использовано в доказательстве равенства (3.44).

3.4 Теорема Дарбу

Пусть на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, задана произвольная r -форма

$$A = \frac{1}{n!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r} A_{\alpha_1 \dots \alpha_r} \in \Lambda_r(\mathbb{M}).$$

Зададим вопрос о том, каково минимальное число функций $f^1(x), \dots, f^q(x)$ таких, что r -форма A представима в виде

$$A = \sum_{1 \leq \alpha_1 < \dots < \alpha_r \leq q} df^{\alpha_1} \wedge \dots \wedge df^{\alpha_r} A_{\alpha_1 \dots \alpha_r} (f^1 \dots f^q). \quad (3.46)$$

Ясно, что при $q = n$ любая форма уже имеет такой вид. С другой стороны, число функций не может быть меньше степени формы. Поэтому представляет интерес случай $r \leq q < n$. Важность этого вопроса заключается в том, что функции $f^1, \dots, f^q$ можно выбрать в качестве части координат некоторой новой координатной системы и тем самым привести форму к каноническому виду, в котором она будет зависеть не от всех координат, а только от их части. Ответ на поставленный вопрос дает теорема 3.47, сформулированная в настоящем разделе. Для 1- и 2-форм ответ дают теоремы Дарбу.

Для формулировки теорем мы введем новые понятия: ранг и класс дифференциальных форм. Начнем с форм над произвольным векторным пространством $\mathbb{V}$ с базисом e_α , $\alpha = 1, \dots, n$. Дадим несколько определений.

Определение. Назовем r -форму $A \in \Lambda_r(\mathbb{V})$ *разложимой*, если существует r линейно независимых 1-форм $a^1, \dots, a^r \in \Lambda_1(\mathbb{V})$ таких, что

$$A = a^1 \wedge \dots \wedge a^r. \quad \square$$

Ясно, что всякую r -форму можно представить в виде суммы разложимых r -форм, например, разложив по базису. Однако, не всякая r -форма разложима.

Пример 3.4.1. Рассмотрим четырехмерное векторное пространство $\mathbb{V}$ с базисом e_α , $\alpha = 1, 2, 3, 4$. Базис сопряженного пространства $\mathbb{V}^*$ обозначим через e^α . Рассмотрим 2-форму

$$A = e^1 \wedge e^2 + e^3 \wedge e^4 \in \Lambda_2(\mathbb{V}).$$

Если форма A разложима, то существуют две 1-формы $a = e^\alpha a_\alpha$ и $b = e^\beta b_\beta$ такие, что выполнено условие

$$A = a \wedge b = e^\alpha \wedge e^\beta a_\alpha b_\beta = e^1 \wedge e^2 + e^3 \wedge e^4.$$

Отсюда вытекает следующая система уравнений на компоненты

$$\begin{aligned} a_1 b_2 &= 1, & a_2 b_3 &= 0, \\ a_1 b_3 &= 0, & a_2 b_4 &= 0, \\ a_1 b_4 &= 0, & a_3 b_4 &= 1. \end{aligned}$$

Не ограничивая общности, можно положить $a_1 = 1$. Тогда из трех уравнений в первой колонке следует $b_2 = 1$, $b_3 = b_4 = 0$. Подставив эти решения в остальные уравнения, приходим к противоречию. Следовательно, 2-форма A является неразложимой.

Рассмотрим другую 2-форму над тем же векторным пространством $\mathbb{V}$

$$A = e^1 \wedge e^2 + e^1 \wedge e^3 \in \Lambda_2(\mathbb{V}).$$

Ее можно представить в виде $A = e^1 \wedge (e^2 + e^3)$. Отсюда следует, что эта форма разложима: $A = e^1 \wedge b$, где $b := e^2 + e^3 \in \Lambda_1(\mathbb{V})$. $\square$

Определение. Назовем подпространство $\mathbb{A}(A) \subset \mathbb{V}$ *ассоциированным* с формой A , если это наибольшее из тех подпространств $\mathbb{H} \subset \mathbb{V}$, для которых $A \in \Lambda_r(\mathbb{V}/\mathbb{H})$. Другими словами, пространство $\mathbb{A}(A)$ образовано теми векторами $X \in \mathbb{V}$, чье внутреннее произведение с формой равно нулю, $i_X A = 0$. Это значит, что подпространство, ассоциированное с формой A , есть подпространство в $\mathbb{V}$, которое определяется системой уравнений $i_X A = 0$.

Каждому ассоциированному пространству $\mathbb{A}(A)$ можно поставить в соответствие *ассоциированную с формой систему* $\mathbb{A}^*(A)$, как ортогональное дополнение сопряженного пространства

$$\mathbb{A}^*(A) := \mathbb{A}^\perp(A) \in \mathbb{V}^*.$$

Размерность ассоциированной системы называется *рангом* формы A

$$\text{rank } A := \dim \mathbb{A}^*(A). \quad \square$$

Ясно, что ранг r -формы A равен минимальному числу линейно независимых 1-форм таких, что форма A равна линейной комбинации их внешних произведений.

Пример 3.4.2. Для 2-форм $A = \frac{1}{2}e^\alpha \wedge e^\beta A_{\alpha\beta}$ понятие ранга особенно наглядно: это просто ранг матрицы компонент,

$$\dim \mathbb{A}^*(A) = \text{rank } A_{\alpha\beta}.$$

□

Пример 3.4.3. Рассмотрим 0-формы $A \in \Lambda_0(\mathbb{V})$, т.е. $A \in \mathbb{R}$. Тогда ассоциированное подпространство $\mathbb{A}(A) = \mathbb{V}$, т.к. факторпространство $\mathbb{V}/\mathbb{V}$ состоит из одной точки – начала координат, и $\Lambda_0(\mathbb{V}/\mathbb{V}) = \mathbb{R}$. Соответствующая система $\mathbb{A}^*(A)$, ассоциированная с формой A , также состоит из одной точки – начала координат сопряженного пространства $\mathbb{A}^*(A) = \{0\}$. Таким образом, ранг произвольной 0-формы равен нулю. □

Пример 3.4.4. Пусть $\mathbb{V} = \mathbb{R}^n$ с базисом $\{e_1, \dots, e_n\}$ и $A = e^1 + e^2 \in \Lambda_1(\mathbb{R}^n)$, тогда ассоциированное подпространство $\mathbb{A}(A) = \mathbb{R}^{n-1}$ порождено базисными векторами $\{e_1 - e_2, e_3, \dots, e_n\}$. Соответствующая ассоциированная система $\mathbb{A}^*(A)$ представляет собой одномерное подпространство сопряженного пространства $\mathbb{V}^* = \mathbb{R}^n$, порожденное формой A , т.е. натянутое на базисный вектор $e^1 + e^2$. Ранг 1-формы $e^1 + e^2$ равен единице. Ясно, что ранг произвольной 1-формы равен единице. □

Пример 3.4.5. Пусть $\mathbb{V} = \mathbb{R}^n$ и $A = e^1 \wedge e^2 + e^3 \wedge e^4 \in \Lambda_2(\mathbb{R}^n)$. Эта 2-форма неразложима, поэтому ассоциированное с ней подпространство имеет размерность $n - 4$. Следовательно, ассоциированное подпространство имеет вид $\mathbb{A}(A) = \mathbb{R}^{n-4}$ и порождено векторами $e_5, \dots, e_n$. Соответствующая ассоциированная система четырехмерна и порождена векторами e^1, e^2, e^3, e^4 . Таким образом, ранг 2-формы $e^1 \wedge e^2 + e^3 \wedge e^4$ равен четырем. □

В общем случае справедлива следующая теорема, дающая ограничения на возможный ранг формы степени r .

Теорема 3.4.1. Если $A \in \Lambda_r(\mathbb{V})$ – ненулевая r -форма на $\mathbb{V}$, то ее ранг не меньше r и не больше n

$$r \leq \text{rank } A \leq n.$$

Ранг r -формы A равен r тогда и только тогда, когда форма A разложима.

Доказательство. См., например, [44], предложение 7.13. □

Как следствие получаем, что, поскольку любая n -форма разложима, то ее ранг равен n . Можно доказать, что любая $(n - 1)$ -форма разложима и ее ранг, следовательно, равен $n - 1$. А также, что ранг $(n - 2)$ -формы равен либо $n - 2$, либо n .

Доказательства следующих трех предложений для внешних форм второй степени $A \in \Lambda_2(\mathbb{V})$, $\dim(\mathbb{V}) = n$, приведены в [44].

Предложение 3.4.1. Пусть $A \in \Lambda_2(\mathbb{V})$ – внешняя 2-форма. Тогда существует целое четное число $2s$ такое, что $2 \leq 2s \leq n$, и $2s$ независимых линейных форм $a^1, \dots, a^{2s}$ таких, что

$$A = a^1 \wedge a^2 + \dots + a^{2s-1} \wedge a^{2s}.$$

При этом 1-форму a^1 можно выбрать произвольно в ассоциированной системе $\mathbb{A}^*(A)$.

Предложение 3.4.2. Внешняя 2-форма имеет четный ранг.

Предложение 3.4.3. Пусть $A \in \Lambda_2(\mathbb{V})$ – внешняя 2-форма. Для того, чтобы 2-форма A имела ранг $2s$, необходимо и достаточно, чтобы

$$A^s \neq 0 \quad \text{и} \quad A^{s+1} = 0,$$

где в левой части соотношений стоят внешние степени формы A .

Приведенные выше определения и утверждения даны для внешней алгебры $\Lambda(\mathbb{V})$ над произвольным векторным пространством $\mathbb{V}$ и применимы в каждой точке многообразия $\mathbb{M}$ для внешней алгебры $\Lambda(\mathbb{M})$. При рассмотрении форм на многообразии эти понятия, в том числе ранг формы, могут меняться от точки к точке.

Введем несколько новых понятий, для которых уже важна зависимость компонент формы от точки многообразия.

Определение. Характеристическим подпространством r -формы $A \in \Lambda_r(\mathbb{M})$ в точке $x \in \mathbb{M}$ называется подпространство касательного пространства $\mathbb{C}_x(A) \subset \mathbb{T}_x(\mathbb{M})$, которое является пересечением ассоциированных подпространств для формы A и ее внешней производной dA ,

$$\mathbb{C}_x(A) := \mathbb{A}_x(A) \cap \mathbb{A}_x(dA),$$

в точке x . Характеристической системой формы $A \in \Lambda_r(\mathbb{M})$ в точке $x \in \mathbb{M}$ называется подпространство $\mathbb{C}_x^*(A) \subset \mathbb{T}_x^*(\mathbb{M})$, которое ортогонально характеристическому подпространству этой формы $\mathbb{C}_x(A)$. $\square$

Из определения и формулы (1.47) вытекает, что характеристическая система $\mathbb{C}_x^*(A)$ строится по ассоциированным системам $\mathbb{A}_x^*(A)$ и $\mathbb{A}_x^*(dA)$ следующим образом

$$\begin{aligned} \mathbb{C}_x^*(A) &= \mathbb{A}_x^*(A) \oplus [\mathbb{A}_x^*(dA) \setminus (\mathbb{A}_x^*(A) \cap \mathbb{A}_x^*(dA))] = \\ &= [\mathbb{A}_x^*(A) \setminus (\mathbb{A}_x^*(A) \cap \mathbb{A}_x^*(dA))] \oplus \mathbb{A}_x^*(dA). \end{aligned}$$

Определение. Классом r -формы $A \in \Lambda_r(\mathbb{M})$ в точке $x \in \mathbb{M}$ называется размерность характеристической системы: $\text{class}(A) := \dim \mathbb{C}_x^*(A)$. $\square$

Из определения следует, что класс ненулевой r -формы A не меньше ее ранга и, следовательно, не меньше степени r

$$r \leq \text{rank}(A) \leq \text{class}(A).$$

В общем случае класс формы может зависеть от точки многообразия.

Пример 3.4.6. Рассмотрим 1-форму $A = dy(x^2 + y^2)$ на евклидовой плоскости $\mathbb{R}^2$ с декартовыми координатами x, y . Ее внешний дифференциал равен $dA = dx \wedge dy$. Существуют три случая.

1) $x \neq 0$. В этом случае ассоциированное подпространство $\mathbb{A}(A)$ порождено вектором ∂_x , а ассоциированное подпространство внешнего дифференциала нульмерно, $\mathbb{A}(dA) = \{0\}$. Поэтому характеристическое подпространство состоит из одной точки $\mathbb{C}_x(A) = \{0\}$, а характеристическая система совпадает с кокасательным пространством, $\mathbb{C}_x^*(A) = \mathbb{T}_x^*(\mathbb{R}^2)$. Следовательно, класс формы A равен двум, $\text{class}(A) = 2$.

2) $x = 0, y \neq 0$. Ассоциированное подпространство $\mathbb{A}(A)$ по-прежнему порождено вектором ∂_x . Внешний дифференциал формы A равен нулю, $dA = 0$, и ассоциированное подпространство внешнего дифференциала совпадает с касательным пространством, $\mathbb{A}(dA) = \mathbb{T}_x(\mathbb{R}^2)$. Следовательно, характеристическое подпространство

$\mathbb{C}_x(A)$ порождено вектором ∂_x , а характеристическая система – дифференциалом dy . Поэтому класс формы A равен единице, $\text{class}(A) = 1$.

3) $x = y = 0$. В этом случае и форма, и ее внешний дифференциал обращаются в нуль: $A = 0$, $dA = 0$. Поэтому ассоциированное подпространство совпадает с касательным пространством, $\mathbb{C}_x(A) = \mathbb{T}_x(\mathbb{R}^2)$, а ассоциированная система состоит из одной точки $\mathbb{C}_x^* = \{0\}$. Следовательно, класс формы A равен нулю, $\text{class}(A) = 0$.

Таким образом, класс рассматриваемой формы зависит от точки плоскости и равен:

$\text{class}(A)=2$, если $x \neq 0$;

$\text{class}(A)=1$, если $x = 0$ и $y \neq 0$;

$\text{class}(A)=0$, если $x = y = 0$. □

Продолжим общее рассмотрение. Нетрудно проверить, что если класс r -формы A равен ее степени r в некоторой точке $x \in \mathbb{M}$, то эта форма замкнута, $dA(x) = 0$. В обратную сторону справедливо другое утверждение. Если форма замкнута, $dA(x) = 0$, в точке x , то в этой точке ее класс равен рангу, а не степени. Отсюда, в частности, следует, что замкнутая форма второй степени имеет четный класс.

Теперь сформулируем основное утверждение настоящего раздела.

Теорема 3.4.2. Пусть $A \in \Lambda_r(\mathbb{M})$ – дифференциальная r -форма на многообразии $\mathbb{M}$ постоянного класса, $\text{class}(A) = q$, $r \leq q \leq n$. Тогда в окрестности любой точки $x \in \mathbb{M}$ существует такая система координат $y^1, \dots, y^n$, что форма A представима в виде

$$A = \sum_{1 \leq \alpha_1 < \dots < \alpha_r \leq q} dy^{\alpha_1} \wedge \dots \wedge dy^{\alpha_r} A_{\alpha_1 \dots \alpha_r}(y^1 \dots y^q). \quad (3.47)$$

Доказательство. См., например, [44], предложение 3.2. □

Нетривиальность этой теоремы заключается в том, что форма записана в виде, содержащим только первые q координат. Эта теорема дает ответ на вопрос, поставленный в начале раздела. Другими словами, класс формы равен минимальному числу независимых функций, необходимых для явного выражения r -формы постоянного класса.

Следствие. Пусть $A \in \Lambda_r(\mathbb{M})$ – дифференциальная r -форма на многообразии $\mathbb{M}$ постоянного класса, который равен степени, $\text{class}(A) = r$. Тогда в окрестности любой точки $x \in \mathbb{M}$ существует такая система координат $y^1, \dots, y^n$, что форма A имеет вид

$$A = dy^{\alpha_1} \wedge \dots \wedge dy^{\alpha_r} \quad (3.48)$$

Доказательство. См., например, [44], следствие 3.4. □

Для замкнутых 1- и 2-форм постоянного класса теорему 3.4.2 можно уточнить.

Теорема 3.4.3 (Дарбу). Пусть A – произвольная 1-форма без нулей на многообразии $\mathbb{M}$ постоянного класса $2s + 1$ (или $2s$). Тогда для любой точки $x \in \mathbb{M}$ существуют $2s + 1$ (или $2s$) функций $y^1, \dots, y^{2s+1}$ (или $y^1, \dots, y^{2s}$), заданных в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$, содержащей точку x , такие, что

$$\begin{aligned} y^1(x) = \dots = y^{2s+1}(x) &= 0, \\ \text{или} \quad y^1(x) = \dots = y^{2s}(x) &= 0, \end{aligned}$$

и

$$A|_{\mathbb{U}} = dy^1 + y^2 dy^3 + \dots + y^{2s} dy^{2s+1},$$

или

$$A|_{\mathbb{U}} = (1 + y^1) dy^2 + y^3 dy^4 + \dots + y^{2s-1} dy^{2s}.$$

Доказательство. См., например, [44], теорема 4.1. $\square$

Замечание. Если 1-форма имеет постоянный нечетный класс, то она не имеет нулей на многообразии. Напротив, если 1-форма имеет постоянный четный класс, то она может иметь нули. В этом случае невозможно привести общую локальную модель 1-формы. $\square$

Теорема 3.4.4 (Дарбу). Пусть ω – замкнутая дифференциальная 2-форма на многообразии $\mathbb{M}$ постоянного класса $\text{class } \omega = 2s$. Тогда для любой точки $x \in \mathbb{M}$ существуют $2s$ дифференцируемых функций $y^1, \dots, y^{2s}$, заданных в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$, содержащей точку x , такие, что

$$y^1(x) = \dots = y^{2s}(x) = 0,$$

$$\omega|_{\mathbb{U}} = dy^1 \wedge dy^2 + \dots + dy^{2s-1} \wedge dy^{2s}.$$

Доказательство. Из замкнутости 2-формы ω по лемме Пуанкаре следует, что она локально представима в виде $\omega = dA$, где 1-форма A в точке x имеет класс $2s$ или $2s + 1$. Тогда утверждение теоремы следует из теоремы 3.4.3. $\square$

Пример 3.4.7. Рассмотрим симплектическое многообразие $(\mathbb{M}, \omega)$, $\dim \mathbb{M} = 2n$. По определению, симплектическая форма ω на многообразии $\mathbb{M}$ невырождена, и, следовательно, ее ранг и класс постоянны и равны размерности многообразия

$$\text{rank } \omega = \text{class } \omega = \dim \mathbb{M} = 2n.$$

Симплектическая форма ω также замкнута. Поэтому из теоремы Дарбу следует, что для любой точки $x \in \mathbb{M}$ существуют $2n$ дифференцируемых функций $y^1, \dots, y^{2n}$, заданных в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$, содержащей точку x , таких, что

$$y^1(x) = \dots = y^{2n}(x) = 0,$$

$$\omega|_{\mathbb{U}} = dy^1 \wedge dy^2 + \dots + dy^{2n-1} \wedge dy^{2n}. \quad (3.49)$$

Другими словами, для любой симплектической формы в окрестности произвольной точки существует такая система координат, в которой она имеет канонический вид. Именно это следствие теоремы Дарбу, которое часто тоже называют теоремой Дарбу, обычно используют в гамильтоновой динамике. $\square$

Определение. Система координат, в которой симплектическая форма имеет канонический вид (3.49), называется *координатами Дарбу*. $\square$

Замечание. В римановой геометрии $(\mathbb{M}, g)$ с метрикой g общего вида всегда можно выбрать такие координаты, чтобы метрика совпала с евклидовой метрикой (единичной матрицей) в любой заданной точке многообразия. В общем случае этого можно добиться вдоль экстремали, но не в какой-либо окрестности. На симплектическом многообразии симплектическую форму можно привести к каноническому виду в окрестности произвольной точки, что является значительно более сильным утверждением. $\square$

3.5 Оператор Лапласа–Бельтрами

Определение. Пространства форм $\Lambda_r(\mathbb{M})$ и $\Lambda_{n-r}(\mathbb{M})$ имеют одинаковую размерность $C_n^r = C_n^{n-r}$, и при наличии метрики $g_{\alpha\beta}$ между их элементами можно установить взаимно однозначное соответствие (изоморфизм $\mathcal{C}^\infty(\mathbb{M})$ -модулей), которое задается оператором

$$* : \Lambda_r(\mathbb{M}) \leftrightarrow \Lambda_{n-r}(\mathbb{M}).$$

Рассмотрим дифференциальную r -форму $A = \frac{1}{r!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r} A_{\alpha_1 \dots \alpha_r}$ в (псевдо)римановом пространстве, т.е. при наличии метрики. Поставим ей в соответствие $(n-r)$ -форму по следующему правилу

$$*A := \frac{1}{(n-r)!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_{n-r}} (*A)_{\alpha_1 \dots \alpha_{n-r}}, \quad (3.50)$$

где компоненты $(n-r)$ -формы определены выражением

$$(*A)_{\alpha_{r+1} \dots \alpha_n} := \frac{1}{r!} \varepsilon_{\alpha_1 \dots \alpha_n} A^{\alpha_1 \dots \alpha_r}, \quad A^{\alpha_1 \dots \alpha_r} := g^{\alpha_1 \beta_1} \dots g^{\alpha_r \beta_r} A_{\beta_1 \dots \beta_r}. \quad (3.51)$$

Действие оператора $*$ можно записать также в виде

$$*A = \frac{1}{(n-r)!} dx^{\alpha_{r+1}} \wedge \dots \wedge dx^{\alpha_n} \frac{1}{r!} \varepsilon_{\alpha_1 \dots \alpha_r \alpha_{r+1} \dots \alpha_n} A^{\alpha_1 \dots \alpha_r}. \quad (3.52)$$

Эти формулы определяют *оператор $*$* , который также называется *оператором Ходжа*. Форма $*A$ называется *дуальной* к форме A . $\square$

Замечание. Метрика риманова пространства входит в определение оператора $*$ дважды. Во-первых, с ее помощью определяется полностью антисимметричный тензор $\varepsilon_{\alpha_1 \dots \alpha_n} = \sqrt{|g|} \hat{\varepsilon}_{\alpha_1 \dots \alpha_n}$. Во-вторых, метрика используется для подъема индексов в уравнении (3.51). Отметим, что для определения оператора $*$ необходимо не только наличие метрики на многообразии, но и его ориентируемость. В противном случае плотность $\sqrt{|g|}$ глобально не определена. $\square$

Оператор Ходжа $*$ линеен:

$$*(fA + gB) = f(*A) + g(*B), \quad \forall f, g \in \mathcal{C}^\infty(\mathbb{M}), \quad A, B \in \Lambda_r(\mathbb{M}).$$

Оператор $*$, действуя на r -форму, обладает следующим свойством

$$**A = (-1)^{r(n-r)} A \operatorname{sgn}, \quad (3.53)$$

где множитель sgn равен знаку определителя метрики (2.67). Вид этой формулы сохраняется при замене $r \mapsto n-r$. Отсюда следует, что обратный оператор, действующий на $(n-r)$ -форму в порядке $*^{-1}*$ или на r -форму в обратном порядке $**^{-1}$, имеет одинаковый вид

$$*^{-1} = (-1)^{r(n-r)} * \operatorname{sgn}.$$

Замечание. Отметим, что при проведении расчетов необходимо учитывать, что множитель $(-1)^{r(n-r)}$ зависит от степени формы, на которую действует оператор $*^{-1}$. Для форм четной степени он всегда равен единице. $\square$

Замечание. Оператор $*$ можно определить в произвольной внешней алгебре, снабженной метрикой, поскольку он не содержит дифференцирования. $\square$

Поскольку на $\mathbb{M}$ задана метрика, то определена свертка двух r -форм A и B :

$$(A, B) := g^{\alpha_1 \beta_1} \dots g^{\alpha_r \beta_r} A_{\alpha_1 \dots \alpha_r} B_{\beta_1 \dots \beta_r},$$

которая дает скалярное поле (функцию).

Предложение 3.5.1. Для двух r -форм $A, B \in \Lambda_r(\mathbb{M})$ справедлива формула

$$A \wedge *B = dx^1 \wedge \dots \wedge dx^n \sqrt{|g|} \frac{(A, B)}{r!}. \quad (3.54)$$

Доказательство. Прямая проверка. $\square$

Определение. Введем скалярное произведение

$$\langle \cdot, \cdot \rangle : \Lambda_r(\mathbb{M}) \times \Lambda_r(\mathbb{M}) \rightarrow \mathbb{R}$$

в пространстве r -форм $A, B \in \Lambda_r(\mathbb{M})$ с помощью интеграла

$$\langle A, B \rangle := \int_{\mathbb{M}} A \wedge *B = \int_{\mathbb{M}} B \wedge *A = \frac{1}{r!} \int_{\mathbb{M}} dx \sqrt{|g|} (A, B) \operatorname{sgn}. \quad (3.55)$$

Мы предполагаем, что этот интеграл сходится. В противном случае скалярное произведение неопределено. $\square$

Скалярное произведение симметрично, $\langle A, B \rangle = \langle B, A \rangle$. Если метрика риманова, т.е. положительно определена, то скалярное произведение $\langle A, B \rangle$ также положительно определено.

Предложение 3.5.2. Оператор $*$ ортогонален:

$$\langle *A, *B \rangle = \langle A, B \rangle \operatorname{sgn}.$$

Доказательство. Прямая проверка. $\square$

Определение. Введем дифференциальный оператор *кограницы* $\delta: \Lambda_r(\mathbb{M}) \rightarrow \Lambda_{r-1}(\mathbb{M})$

$$\delta := (-1)^{n(r+1)+1} \operatorname{sgn} * d*, \quad (3.56)$$

понижающий степень формы на единицу. Его можно записать также в виде

$$\delta = (-1)^r *^{-1} d*,$$

учитывая, что обратный оператор $*^{-1}$ действует на $(n-r+1)$ -форму. По определению, действие оператора кограницы на функцию дает нуль $\delta f = 0$. $\square$

В компонентах действие оператора δ на r -форму A имеет вид

$$\begin{aligned} \delta A &= -\frac{1}{(r-1)!} dx^{\alpha_2} \wedge \dots \wedge dx^{\alpha_r} \frac{1}{\sqrt{|g|}} (g_{\alpha_2 \beta_2} \dots g_{\alpha_r \beta_r}) \partial_{\alpha_1} (\sqrt{|g|} A^{\alpha_1 \beta_2 \dots \beta_r}) = \\ &= -\frac{1}{(r-1)!} dx^{\alpha_2} \wedge \dots \wedge dx^{\alpha_r} \tilde{\nabla}_{\alpha_1} A^{\alpha_1}_{\alpha_2 \dots \alpha_r}. \end{aligned} \quad (3.57)$$

Отсюда следует, что оператор δ можно интерпретировать, как ковариантную дивергенцию антисимметричного тензора с верхними индексами.

Предложение 3.5.3. Для замкнутых (компактных и без края) многообразий оператор δ сопряжен с оператором внешнего дифференцирования относительно скалярного произведения (3.55)

$$\langle dA, B \rangle = \langle A, \delta B \rangle, \quad (3.58)$$

где $A \in \Lambda_{r-1}(\mathbb{M})$ и $B \in \Lambda_r(\mathbb{M})$.

Доказательство. Прямая проверка. Компактность многообразия достаточна для существования интеграла (3.55), а отсутствие края позволяет отбросить граничные слагаемые, возникающие при интегрировании по частям. $\square$

Предложение 3.5.4. Квадрат оператора кограницы δ равен нулю:

$$\delta\delta = 0, \quad (3.59)$$

Доказательство. Прямое следствием определения (3.56) и нильпотентности внешнего дифференцирования (3.33). $\square$

Определение. Оператором Лапласа–Бельтрами в римановом пространстве называется оператор

$$\Delta_F := -(d + \delta)^2 = -d\delta - \delta d, \quad (3.60)$$

который действует в пространстве форм, $\Delta_F : \Lambda_r(\mathbb{M}) \rightarrow \Lambda_r(\mathbb{M})$, $0 \leq r \leq n$, $\square$

Предложение 3.5.5. Для замкнутых многообразий оператор Лапласа–Бельтрами самосопряжен:

$$\langle \Delta_F A, B \rangle = \langle A, \Delta_F B \rangle,$$

Доказательство. Из симметричности скалярного произведения и формулы (3.58) следуют равенства

$$\langle d\delta A, B \rangle = \langle \delta A, \delta B \rangle = \langle A, d\delta B \rangle.$$

Аналогичные равенства справедливы для оператора δd . Отсюда вытекает самосопряженность оператора Лапласа–Бельтрами. $\square$

Замечание. Оператор Лапласа–Бельтрами является дифференциальным оператором второго порядка эллиптического типа в римановом пространстве с положительно определенной метрикой. Если на многообразии задана метрика лоренцевой сигнатуры, то инвариантный оператор (3.60) также определен. В этом случае он будет гиперболического типа. Если для метрики лоренцевой сигнатуры скалярное произведение определено, то оператор (3.60) будет также самосопряжен, поскольку в доказательстве сигнатура метрики не используется. $\square$

Предложение 3.5.6. Оператор Лапласа–Бельтрами удовлетворяет следующим перестановочным соотношениям:

$$\Delta_F d = d\Delta_F, \quad \Delta_F \delta = \delta\Delta_F, \quad \Delta_F * = * \Delta_F, \quad (3.61)$$

Доказательство. Прямое следствие определения (3.60). $\square$

Предложение 3.5.7. *Формулы действия оператора Лапласа–Бельтрами на 0-форму f , 1-форму A и произвольную r -форму B в компонентах можно выразить через ковариантные производные и тензор кривизны следующим образом:*

$$\begin{aligned}\Delta_F f &= \tilde{\Delta} f, \\ \Delta_F A &= dx^\alpha (\tilde{\Delta} A_\alpha + \tilde{R}_\alpha{}^\beta A_\beta), \\ \Delta_F B &= \frac{1}{r!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r} \left(\tilde{\Delta} B_{\alpha_1 \dots \alpha_r} + r \tilde{R}_{\alpha_1}{}^\beta B_{\beta \alpha_2 \dots \alpha_r} - \frac{r(r-1)}{2} \tilde{R}_{\alpha_1 \alpha_2}{}^{\beta \gamma} B_{\beta \gamma \alpha_3 \dots \alpha_r} \right),\end{aligned}\tag{3.62}$$

где $\tilde{\Delta} := g^{\alpha\beta} \tilde{\nabla}_\alpha \tilde{\nabla}_\beta$ – оператор Лапласа–Бельтрами, построенный с помощью символов Кристоффеля.

Доказательство. Прямые вычисления. □

Замечание. В аффинной геометрии мы имеем три инвариантных дифференциальных оператора второго порядка Δ_F , $\tilde{\Delta}$ и Δ , действующих из $\Lambda_r(\mathbb{M}) \rightarrow \Lambda_r(\mathbb{M})$. Первый оператор определен только на формах и поэтому помечен индексом F . Два других определены для произвольных тензорных полей. В (псевдо)римановой геометрии при нулевом кручении и неметричности $\tilde{\Delta} = \Delta$. Отметим также, что действие операторов Δ_F и $\tilde{\Delta}$ на функции совпадает. Как правило, под оператором Лапласа–Бельтрами мы понимаем оператор $\tilde{\Delta}$, действующий на произвольные тензорные поля и зависящий только от метрики. □

Замечание. В геодезических (римановых) координатах (см. раздел 16.9) символы Кристоффеля и все симметризованные частные производные от них обращаются в нуль в некоторой точке $p \in \mathbb{M}$, поэтому в этой точке оператор Лапласа–Бельтрами принимает вид $\tilde{\Delta}|_{x=p} = g^{\alpha\beta} \partial_\alpha \partial_\beta$ □

Пример 3.5.1. Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовыми координатами $x^\alpha = \{x, y, z\}$. Приведем явные формулы действия операторов, введенных ранее, на формы. В трехмерном пространстве могут быть заданы только 0-, 1-, 2- и 3-формы:

$$\begin{aligned}f & & - 0\text{-форма}, \\ A &= dx^\alpha A_\alpha = dx A_x + dy A_y + dz A_z & - 1\text{-форма}, \\ B &= \frac{1}{2} dx^{\alpha_1} \wedge dx^{\alpha_2} B_{\alpha_1 \alpha_2} = dx \wedge dy B_{xy} + dy \wedge dz B_{yz} + dz \wedge dx B_{zx} & - 2\text{-форма}, \\ C &= \frac{1}{3!} dx^{\alpha_1} \wedge dx^{\alpha_2} \wedge dx^{\alpha_3} C_{\alpha_1 \alpha_2 \alpha_3} = dx \wedge dy \wedge dz C_{xyz} & - 3\text{-форма}.\end{aligned}$$

Формы более высокого порядка тождественно равны нулю в силу трехмерности пространства. Пространства 0- и 3-форм являются одномерными, а пространства 1- и 2-форм – трехмерными.

Оператор внешнего дифференцирования (3.30) повышает на единицу степень формы и действует следующим образом:

$$\begin{aligned}df &= dx \partial_x f + dy \partial_y f + dz \partial_z f, \\ dA &= dx \wedge dy (\partial_x A_y - \partial_y A_x) + dy \wedge dz (\partial_y A_z - \partial_z A_y) + dz \wedge dx (\partial_z A_x - \partial_x A_z), \\ dB &= dx \wedge dy \wedge dz (\partial_x B_{yz} + \partial_y B_{zx} + \partial_z B_{xy}), \\ dC &= 0.\end{aligned}$$

Нетрудно проверить, что оператор $*$, определенный соотношением (3.50), действует на формы по правилам:

$$\begin{aligned} *f &= dx \wedge dy \wedge dz f, \\ *A &= dx \wedge dy A_z + dy \wedge dz A_x + dz \wedge dx A_y, \\ *B &= dx B_{yz} + dy B_{zx} + dz B_{xy}, \\ *C &= C_{xyz}. \end{aligned}$$

Этот оператор устанавливает взаимно однозначное соответствие между пространствами $\Lambda_0(\mathbb{R}^3)$ и $\Lambda_3(\mathbb{R}^3)$, а также между $\Lambda_1(\mathbb{R}^3)$ и $\Lambda_2(\mathbb{R}^3)$. Заметим, что для любой формы в трехмерном евклидовом пространстве квадрат оператора $*$ является тождественным оператором:

$$** = 1.$$

Оператор кограницы δ , определенный соотношением (3.56), понижает степень форм на единицу:

$$\begin{aligned} \delta f &= 0, \\ \delta A &= -(\partial_x A_x + \partial_y A_y + \partial_z A_z), \\ \delta B &= dx(\partial_y B_{xy} - \partial_z B_{zx}) + dy(\partial_z B_{yz} - \partial_x B_{xy}) + dz(\partial_x B_{zx} - \partial_y B_{yz}), \\ \delta C &= -(dx \wedge dy \partial_z C_{xyz} + dy \wedge dz \partial_x C_{xyz} + dz \wedge dx \partial_y C_{xyz}). \end{aligned}$$

Теперь нетрудно найти явное действие оператора Лапласа–Бельтрами в пространстве форм:

$$\begin{aligned} \Delta_F f &= \Delta f, \\ \Delta_F A &= dx \Delta A_x + dy \Delta A_y + dz \Delta A_z, \\ \Delta_F B &= dx \wedge dy \Delta B_{xy} + dy \wedge dz \Delta B_{yz} + dz \wedge dx \Delta B_{zx}, \\ \Delta_F C &= dx \wedge dy \wedge dz \Delta C_{xyz}, \end{aligned}$$

где

$$\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$$

– обычный оператор Лапласа в трехмерном евклидовом пространстве. $\square$

Определение. Действие оператора внешнего дифференцирования на 0-формы совпадает с определением градиента функции, поэтому назовем *градиентом* произвольной r -формы ее внешнюю производную

$$\text{grad} := d. \quad (3.63)$$

Действие оператора δ на 1-формы с точностью до знака совпадает с дивергенцией 1-формы. Поэтому примем в качестве инвариантного определения *дивергенции* от произвольной r -формы соотношение

$$\text{div} := -\delta. \quad \square \quad (3.64)$$

В заключение рассмотрим еще один инвариантный оператор, действующий в пространстве форм произвольной степени,

$$\text{rot} := *d : \quad \Lambda_r \rightarrow \Lambda_{n-r-1}. \quad (3.65)$$

В частном случае трехмерного евклидова пространства (пример 3.5.1) он отображает 1-формы в 1-формы:

$$*dA = dx(\partial_y A_z - \partial_z A_y) + dy(\partial_z A_x - \partial_x A_z) + dz(\partial_x A_y - \partial_y A_x),$$

что совпадает с определением ротора от 1-формы. Поэтому примем выражение (3.65) в качестве инвариантного определения *ротора*. Тогда определение оператора Лапласа–Бельтрами (3.60) сводится к хорошо известному тождеству из векторного анализа

$$\Delta_F = \text{grad div} - \text{rot rot}. \quad (3.66)$$

3.6 Разложение Ходжа

Если на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, задана положительно определенная риманова метрика, то можно доказать ряд важных результатов для дифференциальных форм, которые не имеют места в общем случае. Для формулировки этих результатов введем несколько понятий.

Определение. Форма $A \in \Lambda_r(\mathbb{M})$ является *коточной*, если ее можно представить в виде

$$A = \delta B, \quad \text{где } B \in \Lambda_{r+1}(\mathbb{M}).$$

Форма $A \in \Lambda_r(\mathbb{M})$ называется *козамкнутой*, если

$$\delta A = 0.$$

Форма $A \in \Lambda_r(\mathbb{M})$ называется *гармонической*, если

$$\Delta_F A = 0. \quad \square$$

Любая функция $f \in C^1(\mathbb{M})$, очевидно, является козамкнутой, $\delta f = 0$. Любая коточная форма является козамкнутой в силу свойства (3.59).

Теорема 3.6.1. Любая гармоническая форма $A \in \Lambda_r(\mathbb{M})$ на компактном ориентируемом римановом многообразии является замкнутой и козамкнутой.

Доказательство. Из определения оператора Лапласа–Бельтрами (3.60) для скалярного произведения (3.55) следует равенство

$$\langle \Delta_F A, A \rangle = -\langle \delta A, \delta A \rangle - \langle dA, dA \rangle. \quad (3.67)$$

Отсюда следует, что условие гармоничности $\Delta_F A = 0$ влечет за собой равенства $\delta A = 0$ и $dA = 0$. При этом ориентируемость и компактность многообразия достаточны для существования интеграла, входящего в скалярное произведение. Положительная определенность метрики необходима для того, чтобы оба слагаемых в правой части равенства (3.67) обращались в нуль по отдельности, и из равенства $\langle \delta A, \delta A \rangle = 0$ вытекало $\delta A = 0$ (аналогично для dA). $\square$

Следствие. Любая гармоническая функция на компактном ориентируемом римановом многообразии является замкнутой $df = 0$ и, следовательно, равна константе. $\square$

Замечание. Сформулированная теорема справедлива также для компактных неориентируемых многообразий. Действительно, она верна для двулистного ориентируемого накрытия $\tilde{\mathbb{M}} \rightarrow \mathbb{M}$, которое всегда существует и тоже компактно, и при проекции замкнутость и козамкнутость форм сохраняется. $\square$

Теорема 3.6.2 (Разложение Ходжа). Пусть $\mathbb{M}$ – компактное ориентируемое риманово многообразие. Тогда любую r -форму $A \in \Lambda_r(\mathbb{M})$ можно представить в виде суммы точной, коточной и гармонической формы

$$A = dB + \delta C + H, \quad B \in \Lambda_{r-1}(\mathbb{M}), \quad C \in \Lambda_{r+1}(\mathbb{M}), \quad H \in \Lambda_r(\mathbb{M}). \quad (3.68)$$

Это разложение единственно.

Доказательство. См., например, [52], глава V, §31, следствие 1. $\square$

Единственность разложения вытекает из ортогональности трех подпространств относительно скалярного произведения форм:

$$\langle A, A \rangle = \langle dB, dB \rangle + \langle \delta C, \delta C \rangle + \langle H, H \rangle.$$

Следствие. Уравнение

$$\Delta_{\mathbb{F}} A = D, \quad A, D \in \Lambda_r(\mathbb{M}),$$

на компактном ориентируемом многообразии имеет решение тогда и только тогда, когда r -форма D ортогональна пространству гармонических r -форм. $\square$

Теорема 3.6.3. На компактном ориентируемом римановом многообразии существует только конечное число линейно независимых гармонических форм.

Доказательство. См., например, [53], теорема 22. $\square$

Определение. Размерность пространства гармонических r -форм

$$\mathcal{H}_r(\mathbb{M}) = \{H \in \Lambda_r(\mathbb{M}) : \Delta_{\mathbb{F}} H = 0\}$$

называется r -тым числом Бетти

$$b_r(\mathbb{M}) := \dim \mathcal{H}_r(\mathbb{M}). \quad \square \quad (3.69)$$

Отметим несколько свойств чисел Бетти. Во-первых, числа Бетти конечны. Поскольку оператор Лапласа–Бельтрами коммутирует с оператором Ходжа (3.61), а оператор Ходжа устанавливает изоморфизм $\mathcal{H}_r(\mathbb{M})$ и $\mathcal{H}_{n-r}(\mathbb{M})$, то $b_r(\mathbb{M}) = b_{n-r}(\mathbb{M})$. Так как все гармонические 0-формы на компактном ориентируемом римановом многообразии являются константами, то $b_0(\mathbb{M}) = b_n(\mathbb{M}) = 1$. Отметим, что пространство $\mathcal{H}_n(\mathbb{M})$ состоит из n -форм, пропорциональных форме объема v (3.79). Если многообразие $\mathbb{M}$ связно, то $b_n(\mathbb{M}) = 1$ или $b_n(\mathbb{M}) = 0$, соответственно, для ориентируемых и неориентируемых многообразий. В общем случае нулевое число Бетти $b_0(\mathbb{M})$ равно числу компонент связности многообразия $\mathbb{M}$.

Определение. Число

$$\chi(\mathbb{M}) = \sum_{r=0}^n (-1)^r b_r(\mathbb{M}) \quad (3.70)$$

называется *эйлеровой характеристикой* компактного ориентируемого риманова многообразия $\mathbb{M}$. $\square$

Теорема 3.6.4. Пусть $\mathbb{M}$ – компактное ориентируемое риманово многообразие и A – дифференцируемая 1-форма. Тогда

$$\int v \delta A = 0.$$

В частности, для произвольной функции $f \in C^2(\mathbb{M})$

$$\int v \Delta_{\mathbb{F}} f = 0.$$

Доказательство. Справедливо равенство

$$\int_{\mathbb{M}} v \delta A = \langle \delta A, 1 \rangle = \langle A, d1 \rangle = 0. \quad \square$$

Теорема 3.6.5. Пусть $\mathbb{M}$ – компактное риманово ориентируемое многообразие с положительной гауссовой кривизной $K > 0$. Тогда первое число Бетти равно нулю $b_1(\mathbb{M}) = 0$.

Доказательство. Доказательство приведено, например, в [41], теорема 5.25. $\square$

Предложение 3.6.1. Любая ковариантно постоянная форма является замкнутой, козамкнутой и гармонической.

Доказательство. Это следует из выражений для внешней производной (3.36), оператора кограницы (3.57) и определения оператора Лапласа–Бельтрами (3.60). $\square$

Пример 3.6.1. Форма объема v (3.79) замкнута, козамкнута и гармонична, поскольку $\tilde{\nabla}_{\alpha} \sqrt{|g|} = 0$. $\square$

3.7 Интегрирование дифференциальных форм

Начнем с обсуждения интегрирования в евклидовом пространстве $\mathbb{R}^n$. Для наших целей достаточно знакомства с определением интеграла по Риману, т.к. мы будем интегрировать непрерывные функции по достаточно хорошим областям.

Напомним правило замены переменных интегрирования. Пусть h – диффеоморфизм ограниченного открытого подмножества $\mathbb{U} \subset \mathbb{R}^n$ на ограниченное открытое подмножество $h(\mathbb{U}) \subset \mathbb{R}^n$. В координатах диффеоморфизм h задается n гладкими функциями $x^{\alpha} \mapsto y^{\alpha}(x)$, $\alpha = 1, \dots, n$, где x^{α} – координаты точки $x \in \mathbb{U}$. Пусть f – ограниченная непрерывная функция на $\overline{\mathbb{U}}$, включая границу. Тогда справедлива формула замены переменных интегрирования, которая хорошо известна из курса математического анализа,

$$\int_{\mathbb{U}} dx^1 \dots dx^n f(x) = \int_{h(\mathbb{U})} dy^1 \dots dy^n |J|^{-1} f(x(y)). \quad (3.71)$$

где $J := \det(\partial y^{\beta} / \partial x^{\alpha})$ – якобиан преобразования координат. Поскольку мы предполагаем, что отображение h является диффеоморфизмом, то якобиан преобразования всюду отличен от нуля. В частном случае, если $f = 1$, то равенство (3.71) принимает вид

$$\int_{\mathbb{U}} dx^1 \dots dx^n = \int_{h(\mathbb{U})} dy^1 \dots dy^n |J|^{-1}.$$

Если предположить, что x^α – это декартовы координаты в $\mathbb{R}^n$, то последняя формула представляет собой объем области $\mathbb{U} \subset \mathbb{R}^n$. Его можно вычислить как в декартовой системе координат x^α , так и в произвольной криволинейной системе координат y^α . В последнем случае появляется нетривиальный множитель, который равен модулю якобиана преобразования координат.

В дальнейшем мы ограничимся преобразованиями координат с положительным якобианом, $J > 0$, т.е. теми преобразованиями координат, которые сохраняют ориентацию. Кроме этого будем использовать сокращенное выражение для элемента интегрирования

$$dx := dx^1 \dots dx^n.$$

Определение. Рассмотрим произвольную n -форму на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$. На карте $(\mathbb{U}, \varphi)$ она имеет вид

$$A = dx^1 \wedge \dots \wedge dx^n A_{1\dots n} = \frac{1}{n!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_n} A_{\alpha_1 \dots \alpha_n},$$

где в последнем выражении производится суммирование по всем значениям индексов. Положим

$$\int_{\mathbb{U}} A = \int_{\varphi(\mathbb{U})} dx^1 \wedge \dots \wedge dx^n A_{1\dots n} := \int_{\varphi(\mathbb{U})} dx^1 \dots dx^n A_{1\dots n}(x) = \int_{\varphi(\mathbb{U})} dx A_{1\dots n}(x). \quad (3.72)$$

Правая часть этого равенства определена, т.к. $\varphi(\mathbb{U}) \subset \mathbb{R}^n$, а запись $A_{1\dots n}(x)$ обозначает значение единственной независимой компоненты n -формы $A_{1\dots n} \circ \varphi^{-1}$ в точке $\{x^\alpha\} \in \varphi(\mathbb{U})$. Мы, конечно, предполагаем, что интеграл в правой части приведенной формулы сходится. Формула (3.72) корректна в том смысле, что не зависит от выбора системы координат, т.к. при преобразовании координат $x^\alpha \mapsto y^\alpha(x)$ (замене переменных интегрирования) единственная нетривиальная компонента n -формы $A_{1\dots n}$ преобразуется, как тензорная плотность степени -1 :

$$A_{1\dots n}(x) \mapsto J^{-1} A_{1\dots n}(y),$$

что совпадает с правилом замены переменных интегрирования (3.71) при положительном якобиане. Равенство (3.72) примем за определение *интеграла от n -формы* по области $\mathbb{U} \subset \mathbb{M}$. $\square$

Замечание. Определение интеграла от n -формы по области $\mathbb{U} \subset \mathbb{M}$ не зависит от того задана на многообразии метрика или нет. $\square$

Из равенства

$$dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_n} = dx^1 \wedge \dots \wedge dx^n \hat{\varepsilon}^{\alpha_1 \dots \alpha_n}, \quad (3.73)$$

где $\hat{\varepsilon}^{\alpha_1 \dots \alpha_n}$ – полностью антисимметричная тензорная плотность степени -1 , следует, что интеграл (3.72) можно записать в виде

$$\int_{\mathbb{U}} A = \frac{1}{n!} \int_{\varphi(\mathbb{U})} dx \hat{\varepsilon}^{\alpha_1 \dots \alpha_n} A_{\alpha_1 \dots \alpha_n}.$$

Поскольку компоненты тензорной плотности $\hat{\varepsilon}^{\alpha_1 \dots \alpha_n}$ во всех системах координат по модулю равны единице, то этот интеграл не зависит от метрики на $\mathbb{M}$.

Правило замены переменных интегрирования можно переписать иначе на языке возврата отображения. Пусть $h : \mathbb{U} \rightarrow h(\mathbb{U})$ диффеоморфизм двух областей многообразия $\mathbb{M}$, и пусть n -форма A задана на области $h(\mathbb{U})$. Тогда справедлива формула

$$\int_{h(\mathbb{U})} A = \int_{\mathbb{U}} h^* A, \quad (3.74)$$

где h^* – возврат отображения h .

Замечание. Запись интеграла $\int dx A_{1\dots n}$ в виде $\int dx^1 \wedge \dots \wedge dx^n A_{1\dots n}$ удобнее, т.к. якобиан преобразования координат возникает автоматически при замене дифференциалов:

$$dy^1 \wedge \dots \wedge dy^n = \left(\frac{\partial y^1}{\partial x^{\alpha_1}} dx^{\alpha_1} \right) \wedge \dots \wedge \left(\frac{\partial y^n}{\partial x^{\alpha_n}} dx^{\alpha_n} \right) = J dx^1 \wedge \dots \wedge dx^n. \quad \square \quad (3.75)$$

Теперь определим интеграл по произвольному ориентируемому многообразию с помощью разбиения единицы.

Определение. Пусть на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, заданы: ориентация, некоторый атлас $\{\mathbb{U}_i\}$, согласованный с ориентацией, и разбиение единицы $\{f_i\}$, подчиненное выбранному покрытию $\{\mathbb{U}_i\}$. Рассмотрим произвольную непрерывную n -форму $A \in \Lambda_n(\mathbb{M})$ с компактным носителем. Тогда

$$A = \left(\sum_i f_i \right) A = \sum_i (f_i A).$$

Поскольку $\text{supp}(f_i A) \subset \text{supp} f_i$, то справедливо равенство

$$\int_{\mathbb{M}} f_i A := \int_{\mathbb{U}_i} f_i A,$$

где интеграл в правой части был определен ранее (3.72). Этот интеграл определен корректно, т.к. не зависит от выбора системы координат на $\mathbb{U}_i$. Определим *интеграл по ориентируемому многообразию* от n -формы с компактным носителем следующей формулой

$$\int_{\mathbb{M}} A := \sum_i \int_{\mathbb{M}} f_i A. \quad (3.76)$$

Для любого заданного разбиения единицы правая часть равенства однозначно определена. $\square$

Нетрудно проверить, что данное определение не зависит от разбиения единицы. Требование компактности носителя формы A является достаточным для существования интеграла.

Замечание. При изменении ориентации системы координат в области $\mathbb{U}$ компонента n -формы $A_{1\dots n}$ изменит знак. Поэтому правая часть формулы (3.72) также изменит знак, что недопустимо. Поскольку на неориентируемом многообразии не существует возможности однозначно задать ориентацию, то правая часть равенства (3.76) будет зависеть от выбора ориентации областей и, следовательно, определена неоднозначно. Поэтому требование ориентируемости $\mathbb{M}$ является существенным. $\square$

Из свойств интеграла следует, что для любых двух n -форм A и B с компактным носителем справедливы равенства

$$\begin{aligned}\int_{\mathbb{M}} (A + B) &= \int_{\mathbb{M}} A + \int_{\mathbb{M}} B, \\ \int_{\mathbb{M}} aA &= a \int_{\mathbb{M}} A, \quad a \in \mathbb{R}.\end{aligned}$$

Поэтому интеграл является линейным функционалом на множестве n -форм с компактным носителем.

Если $\mathbb{M}$ – ориентируемое многообразие с краем $\partial\mathbb{M}$, то положим

$$\int_{\mathbb{M}} A := \int_{\mathbb{M} \setminus \partial\mathbb{M}} A.$$

Правая часть этого равенства определена выше, т.к. $\mathbb{M} \setminus \partial\mathbb{M}$ – ориентируемое многообразие без края.

Если ориентируемое многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, можно представить в виде объединения

$$\mathbb{M} = \left(\cup_i \mathbb{U}_i \right) \cup \left(\cup_A \mathbb{V}_A \right)$$

непересекающихся открытых подмногообразий $\mathbb{U}_i$ (и, следовательно, многообразий той же размерности $\dim \mathbb{U}_i = n$, $\forall i$) и произвольного набора подмногообразий меньшей размерности $\mathbb{V}_A$, $\dim \mathbb{V}_A < n$, $\forall A$, тогда

$$\int_{\mathbb{M}} A = \sum_i \int_{\mathbb{U}_i} A.$$

Это свойство можно использовать для вычисления интегралов, используя “глобальные” координаты, заданные на всем многообразии, за исключением подмногообразий меньшей размерности, которые не дают вклада в интеграл.

Пример 3.7.1. В задачах со сферической симметрией в $\mathbb{R}^3$ интегралы удобнее вычислять в сферических координатах, которые покрывают все многообразие, за исключением оси z . В этом случае евклидово пространство $\mathbb{R}^3$ представляется в виде

$$\mathbb{R}^3 = \mathbb{U} \cup \mathbb{V},$$

где $\mathbb{V} = \mathbb{R}$ – ось z , и $\mathbb{U} = \mathbb{R}^3 \setminus \mathbb{V}$ – неодносвязное открытое подмногообразие в $\mathbb{R}^3$, $\dim \mathbb{U} = 3$. □

3.7.1 Форма объема

Пусть на ориентируемом многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, задана метрика $dx^\alpha \otimes dx^\beta g_{\alpha\beta}$ ¹ произвольной сигнатуры. Рассмотрим карту $(\mathbb{U}, \varphi)$. В этой карте метрика имеет компоненты $g_{\alpha\beta}$, и ее определитель при преобразовании координат $x \mapsto y(x)$ (в областях пересечения двух карт с координатами x^α и y^α) преобразуется по правилу

$$\det g_{\alpha\beta}(y) = \det g_{\alpha\beta}(x) J^{-2},$$

¹Согласно теореме 4.1.1 на любом гладком многообразии можно задать риманову (положительно определенную) метрику.

где $J := \det(\partial y^\beta / \partial x^\alpha)$ – якобиан преобразования координат. Поэтому определитель метрики является тензорной плотностью степени -2 . Обозначим

$$g := \det g_{\alpha\beta}. \quad (3.77)$$

Тогда $\sqrt{|g|}$ при преобразованиях координат умножается на модуль якобиана,

$$\sqrt{|g|}(y) = \sqrt{|g|}(x)|J|^{-1}. \quad (3.78)$$

Мы ограничимся преобразованиями координат с положительными якобианами, что позволяет рассматривать $\sqrt{|g|}$ как тензорную плотность степени -1 . Если задано координатное покрытие многообразия $\mathbb{M} = \cup_i U_i$, то в каждой карте определена плотность $\sqrt{|g|}$. Для того, чтобы плотность $\sqrt{|g|}$ была определена глобально, необходимо существование такого координатного покрытия, в котором все якобианы преобразований координат были положительны, а это влечет за собой ориентируемость многообразия. Таким образом, плотность $\sqrt{|g|}$ определена только на ориентируемых многообразиях. Она всюду отлична от нуля, что следует из невырожденности метрики.

Определение. n -форма на ориентируемом многообразии $\mathbb{M}$ с заданной метрикой $g_{\alpha\beta}$

$$v := dx^1 \wedge \cdots \wedge dx^n \sqrt{|g|} = \frac{1}{n!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_n} \varepsilon_{\alpha_1 \dots \alpha_n}, \quad (3.79)$$

называется *формой объема* v , задаваемой метрикой $g_{\alpha\beta}$. $\square$

Замечание. Форма объема на ориентируемом многообразии $\mathbb{M}$ задается произвольной метрикой, независимо от ее сигнатуры. Ясно, что различные метрики могут определять одну и ту же форму объема из-за наличия знака модуля. Например, евклидова метрика и метрика Лоренца в евклидовом пространстве $\mathbb{R}^n$ задают одинаковую форму объема. $\square$

Замечание. Для обозначения формы объема мы используем греческую букву эпсилон v , которая по написанию очень похожа на латинскую букву v . $\square$

Форма объема определена глобально на ориентируемых многообразиях и всюду отлична от нуля, поэтому она задает ориентацию многообразия и выбор знака в форме (3.79) соответствует выбору ориентации. При замене координат с отрицательным якобианом ориентация формы объема меняется на противоположную, а интеграл меняет знак.

Метрика евклидова пространства $\mathbb{R}^n$ в декартовых координатах равна единичной матрице. Поэтому форма объема евклидова пространства $\mathbb{R}^n$ в декартовых координатах имеет канонический вид

$$v_0 = dx^1 \wedge \cdots \wedge dx^n. \quad (3.80)$$

Будем говорить, что данная форма объема задает *каноническую ориентацию*.

Замечание. Покажем, что форма объема v на многообразии $\mathbb{M}$ соответствует нашему обычному представлению об объеме. Для этого зафиксируем произвольную точку x риманова многообразия $\mathbb{M}$ с положительно определенной метрикой и диагонализуем в ней метрику, что всегда возможно в произвольной фиксированной точке. Тогда все базисные векторы в точке x будут ортогональны, а длина базисного вектора ∂_α будет равна $\sqrt{g_{\alpha\alpha}}$. В этом случае объем бесконечно малого параллелепипеда со сторонами ∂_α равен $\prod_\alpha \sqrt{g_{\alpha\alpha}} = \sqrt{g}$. $\square$

Интеграл (3.72) от произвольной n -формы по многообразию можно переписать, используя форму объема,

$$\int_{\mathbb{U}} A = \frac{1}{n!} \int_{\varphi(\mathbb{U})} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_n} A_{\alpha_1 \dots \alpha_n} = \frac{1}{n!} \int_{\varphi(\mathbb{U})} dx \sqrt{|g|} \varepsilon^{\alpha_1 \dots \alpha_n} A_{\alpha_1 \dots \alpha_n} \operatorname{sgn}, \quad (3.81)$$

где множитель sgn равен знаку определителя метрики (2.67).

Приведем два свойства формы объема. Нетрудно проверить, что форма объема дуальна к единице

$$v = \frac{1}{n!} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_n} \varepsilon_{\alpha_1 \dots \alpha_n} = *1.$$

Кроме того она является гармонической формой $\Delta_F v = 0$, т.к. форма объема точна $dv = 0$, как форма максимальной размерности, и коточна $\delta v = 0$, поскольку $(n+1)$ -форм на $\mathbb{R}^n$ не существует и $\tilde{\nabla}_\alpha \sqrt{|g|} = 0$.

Если на многообразии задан репер, $g_{\alpha\beta} = e_\alpha^a e_\beta^b \eta_{ab}$, где η_{ab} – евклидова или лоренцева метрика, то $\sqrt{|g|} = |\det e_\alpha^a|$. Тогда репер e_α^a однозначно определяет форму объема v . В обратную сторону утверждение неверно, т.к. форма объема определяет только определитель репера. Даже если метрика и форма объема на ориентируемом многообразии $\mathbb{M}$ существуют, то репер может не существовать. Локально репер всегда существует, однако глобально это не так.

Пример 3.7.2. В теореме 10.2.1 о невозможности причесать ежа доказано, что на четномерной сфере не существует непрерывных векторных полей, нигде не обращающихся в нуль. Поскольку репер представляет собой набор из n линейно независимых в каждой точке векторных полей на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, то на четномерной сфере не существует глобально определенного репера. В то же время метрика определена глобально. $\square$

Наличие формы объема на (псевдо)римановом ориентируемом многообразии позволяет определить инвариантным образом интеграл от произвольной функции $f(x) \in \mathcal{C}(\mathbb{M})$ с компактным носителем:

$$\int_{\mathbb{M}} v f = \int_{\mathbb{M}} dx^1 \wedge \dots \wedge dx^n \sqrt{|g|} f(x) = \int_{\mathbb{M}} dx \sqrt{|g|} f(x).$$

Используя определение скалярного произведения (3.55), этот интеграл можно переписать в виде

$$\int_{\mathbb{M}} v f = \langle f, 1 \rangle.$$

Замечание. Именно это свойство чаще всего используется для построения моделей математической физики. А именно, мы фиксируем некоторый набор полей $\{\varphi^a(x)\}$, $a = 1, 2, \dots$, на многообразии $\mathbb{M}$, среди которых содержится метрика пространства-времени, как в общей теории относительности. Затем строится лагранжева плотность $L(\varphi, \partial\varphi, \partial^2\varphi, \dots)$, которая является скалярным полем (функцией), зависящим от данного набора полей и их частных производных. Определяем инвариантное действие на многообразии

$$S := \int_{\mathbb{M}} v L.$$

Уравнения Эйлера–Лагранжа для данного действия принимаются в качестве уравнений движения (уравнений равновесия и т.д.) для рассматриваемой модели. По построению, действие инвариантно относительно общих преобразований координат и,

возможно, других преобразований симметрии. Соответствующие уравнения Эйлера–Лагранжа ковариантны, т.е. преобразуются, как тензорные поля, и из второй теоремы Нетер следует зависимость уравнений движения. В релятивистских моделях теории поля, не описывающих гравитационные взаимодействия, предполагается, что многообразие $\mathbb{M}$ является пространством-временем Минковского $\mathbb{R}^{1,3}$ с заданной лоренцевой метрикой, по которой варьирование не проводится. Соответствующие модели инвариантны относительно преобразований из группы Пуанкаре и, возможно, других групп симметрии, а инвариантность относительно общих преобразований координат в них отсутствует. В этом случае из первой теоремы Нетер следуют законы сохранения. $\square$

Приведем без доказательства два свойства формы объема v на ориентированном римановом многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$. Пусть $X_1, \dots, X_n$ и $Y_1, \dots, Y_n$ – два набора векторных полей на $\mathbb{M}$. Тогда

$$\begin{aligned} v(X_1, \dots, X_n) \cdot v(Y_1, \dots, Y_n) &= \det \{(X_i, Y_j)\}, \quad i, j = 1, \dots, n, \\ v(X_1, \dots, X_n) \cdot v &= \tilde{X}_1 \wedge \dots \wedge \tilde{X}_n, \end{aligned}$$

где $(X_i, Y_j) := X_i^\alpha Y_j^\beta g_{\alpha\beta}$ и $\tilde{X}_i := dx^\alpha X_{i\alpha}$ обозначает 1-форму, сопряженную в смысле римановой метрики к векторному полю $X_i = X_i^\alpha \partial_\alpha$: $X_{i\alpha} := X_i^\beta g_{\beta\alpha}$.

Определение. Интеграл от формы объема (3.79) по всему многообразию, если он существует, называется *объемом* ориентируемого многообразия $\mathbb{M}$

$$V_{\mathbb{M}} := \int_{\mathbb{M}} v = \int_{\mathbb{M}} dx \sqrt{|g|}. \quad \square \quad (3.82)$$

Объем многообразия определяется метрикой, входящей в определение формы объема, и всегда положителен, т.к. ориентация многообразия фиксирована.

3.7.2 Формула Стокса

Пусть $\mathbb{N}$, $\dim \mathbb{N} = n$ – ориентируемое многообразие, на котором задана произвольная r -форма $A \in \Lambda_r(\mathbb{N})$. Рассмотрим r -мерное ориентируемое подмногообразие $(f, \mathbb{M})$ в $\mathbb{N}$. Поскольку r -форма задана также на подмногообразии $\mathbb{M} \subset \mathbb{N}$, то ее можно по нему проинтегрировать. Пусть подмногообразие $\mathbb{M} \subset \mathbb{N}$ задано параметрически²

$$x^\alpha = x^\alpha(u^i), \quad \alpha = 1, \dots, n, \quad i = 1, \dots, r \leq n.$$

Тогда интеграл от произвольной r -формы A степени r по этому подмногообразию равен

$$\begin{aligned} \int_{f(\mathbb{M})} A &= \frac{1}{r!} \int_{f(\mathbb{M})} dx^{\alpha_1} \wedge \dots \wedge dx^{\alpha_r} A_{\alpha_1 \dots \alpha_r} = \\ &= \frac{1}{r!} \int_{\mathbb{M}} du^{i_1} \wedge \dots \wedge du^{i_r} \frac{dx^{\alpha_1}}{du^{i_1}} \dots \frac{dx^{\alpha_r}}{du^{i_r}} A_{\alpha_1 \dots \alpha_r}. \end{aligned} \quad (3.83)$$

Этот интеграл имеет явно ковариантный вид и не зависит ни от выбора системы координат x^α , ни от параметризации подмногообразия u^i .

²Чтобы не усложнять ситуацию склейкой карт, которую можно провести, будем считать, что каждое многообразие покрывается одной картой.

Пример 3.7.3. Рассмотрим интеграл от 1-формы $A = dx^\alpha A_\alpha$ вдоль кривой $x^\alpha(t)$, $t \in [a, b]$, соединяющей точки x_1 и x_2 в евклидовом пространстве $\mathbb{R}^n$. Тогда интеграл (3.83) принимает стандартный вид

$$\int_{x_1}^{x_2} dx^\alpha A_\alpha = \int_a^b dt \frac{dx^\alpha}{dt} A_\alpha.$$

Этот интеграл инвариантен относительно перепараметризации кривой и общего преобразования координат на многообразии. $\square$

Пример 3.7.4. Интеграл от 2-формы $A = \frac{1}{2} dx^\alpha \wedge dx^\beta A_{\alpha\beta}$ по поверхности в евклидовом пространстве $\mathbb{R}^n$, заданной параметрически $x^\alpha = x^\alpha(u, v)$, $(u, v) \in \mathbb{U}$, равен

$$\frac{1}{2} \int_{f(\mathbb{U})} dx^\alpha \wedge dx^\beta A_{\alpha\beta} = \frac{1}{2} \int_{\mathbb{U}} du \wedge dv \left(\frac{dx^\alpha}{du} \frac{dx^\beta}{dv} - \frac{dx^\alpha}{dv} \frac{dx^\beta}{du} \right) A_{\alpha\beta}. \quad (3.84)$$

В трехмерном евклидовом пространстве $\mathbb{R}^3$ с декартовыми координатами x^α этот интеграл можно преобразовать к виду, хорошо знакомому из курса математического анализа. Касательные векторы к поверхности имеют вид

$$\xi = \frac{dx^\alpha}{du} e_\alpha, \quad \eta = \frac{dx^\alpha}{dv} e_\alpha,$$

где e_α – ортонормированные базисные векторы декартовой системы координат. Их векторное произведение

$$n^\alpha := [\xi, \eta]^\alpha = \varepsilon^{\alpha\beta\gamma} \xi_\beta \eta_\gamma$$

нормально к поверхности. Длина нормального вектора равна $\sqrt{n^\alpha n_\alpha} = \sqrt{\xi^2 \eta^2 - (\xi\eta)^2}$, что совпадает с элементом площади $\sqrt{g}$, где $g := \det g_{ij}$ – определитель индуцированной метрики на поверхности

$$g_{ij} = \frac{dx^\alpha}{dw^i} \frac{dx^\beta}{dw^j} \delta_{\alpha\beta}, \quad \{w^i\} = \{u, v\}, \quad i, j = 1, 2.$$

Так как в трехмерном евклидовом пространстве 2-форма A взаимно однозначно определяется векторным полем $X = X^\alpha \partial_\alpha$, где $A_{\alpha\beta} = \varepsilon_{\alpha\beta\gamma} X^\gamma$, то интеграл (3.84) можно переписать в виде

$$\frac{1}{2} \int_{f(\mathbb{U})} dx^\alpha \wedge dx^\beta A_{\alpha\beta} = \int_{\mathbb{U}} du \wedge dv \sqrt{|g|} (X^\alpha N_\alpha), \quad (3.85)$$

где

$$N^\alpha := \frac{[\xi, \eta]^\alpha}{\sqrt{|g|}}$$

– единичный вектор нормали к поверхности. Таким образом интеграл от 2-формы по поверхности в трехмерном евклидовом пространстве $\mathbb{R}^3$ сводится к суммарному потоку дуального векторного поля через эту поверхность. $\square$

Сформулируем общую теорему, связывающую интеграл от $(n-1)$ -формы A по границе некоторой области $\partial\mathbb{U}$ с интегралом от n -формы dA по самой области $\mathbb{U}$.

Теорема 3.7.1 (Формула Стокса). Пусть $\mathbb{U} \subset \mathbb{M}$ – открытое подмногообразие в $\mathbb{M}$ такое, что его замыкание $\bar{\mathbb{U}}$ компактно. Пусть $\partial\mathbb{U} = \bar{\mathbb{U}} \setminus \mathbb{U}$ – его край, который предполагается кусочно гладким. Тогда справедлива формула Стокса

$$\int_{\partial\mathbb{U}} A = \int_{\mathbb{U}} dA. \quad (3.86)$$

где $A \in \Lambda_{n-1}(\mathbb{M})$ – произвольная гладкая дифференциальная $(n-1)$ -форма на $\mathbb{M}$. При этом предполагается, что ориентация границы задана каноническим образом.

Доказательство. Доказательство формулы Стокса довольно громоздко и приведено, например, в [49], теорема 4.2. Каноническая ориентация края определена в доказательстве теоремы 2.3.2. $\square$

Формулу Стокса можно переписать в другом виде. Пусть на многообразии задано векторное поле $X = X^\alpha \partial_\alpha$. Тогда ему можно взаимно однозначно поставить в соответствие $(n-1)$ -форму:

$$A := *X = \frac{1}{(n-1)!} dx^{\alpha_2} \wedge \cdots \wedge dx^{\alpha_n} \varepsilon_{\beta\alpha_2\ldots\alpha_n} X^\beta.$$

Внешний дифференциал этой формы равен

$$\begin{aligned} dA &= \frac{1}{(n-1)!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_n} \partial_{\alpha_1} (\varepsilon_{\beta\alpha_2\ldots\alpha_n} X^\beta) = \\ &= dx^1 \wedge \cdots \wedge dx^n \sqrt{|g|} \tilde{\nabla}_\alpha X^\alpha. \end{aligned}$$

Отсюда следует, что формулу Стокса можно записать в виде

$$\int_{\partial\mathbb{U}} X^\alpha ds_\alpha = \int_{\mathbb{U}} dx^1 \wedge \cdots \wedge dx^n \sqrt{|g|} \tilde{\nabla}_\alpha X^\alpha, \quad (3.87)$$

где

$$ds_\alpha := \frac{1}{(n-1)!} dx^{\alpha_2} \wedge \cdots \wedge dx^{\alpha_n} \varepsilon_{\alpha\alpha_2\ldots\alpha_n} \quad (3.88)$$

– ориентированный элемент площади гиперповерхности.

Рассмотрим частные случаи формулы Стокса.

Пример 3.7.5. Пусть многообразие $\mathbb{M}$ одномерно, $n = 1$, а форма $A = f(x)$ имеет нулевую степень. Тогда граница области $\mathbb{U} = (a, b)$ состоит из двух точек $\partial\mathbb{U} = \{a, b\}$, и формула Стокса принимает вид

$$\int_{\partial\mathbb{U}} f = f(b) - f(a) = \int_a^b df = \int_a^b dx \partial_x f.$$

Здесь под интегралом от функции (0-формы) по подмногообразию нулевой размерности, состоящему из отдельных точек, понимается алгебраическая сумма значений функции в этих точках. Таким образом, мы получили хорошо известную *формулу Ньютона–Лейбница*. $\square$

Пример 3.7.6. Пусть задана 1-форма $dx^\alpha A_\alpha$ на евклидовой плоскости с декартовыми координатами $\{x^\alpha\} = \{x, y\}$. Рассмотрим область $\mathbb{U}$, ограниченную замкнутой кусочно гладкой кривой $x^\alpha(t)$. Тогда

$$\int_{\partial\mathbb{U}} dt \frac{dx^\alpha}{dt} A_\alpha = \int_{\partial\mathbb{U}} (dx A_x + dy A_y) = \frac{1}{2} \int_{\mathbb{U}} dx \wedge dy \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right). \quad (3.89)$$

Таким образом, мы получили *формулу Грина*. $\square$

Пример 3.7.7. Пусть задана 1-форма $A = dx^\alpha A_\alpha$ в трехмерном евклидовом пространстве с декартовыми координатами x^α . Рассмотрим поверхность $\mathbb{U} \subset \mathbb{R}^3$, с кусочно гладкой границей $\partial\mathbb{U}$. Тогда общая формула Стокса (3.86) принимает вид

$$\oint_{\partial\mathbb{U}} dx^\alpha A_\alpha = \int_{\mathbb{U}} [dx^1 \wedge dx^2 (\partial_1 A_2 - \partial_2 A_1) + dx^1 \wedge dx^3 (\partial_1 A_3 - \partial_3 A_1) + dx^2 \wedge dx^3 (\partial_2 A_3 - \partial_3 A_2)].$$

Если поверхность задана параметрически $x^\alpha = x^\alpha(u, v)$, то последний интеграл можно преобразовать к виду

$$\oint_{\partial\mathbb{U}} dx^\alpha A_\alpha = \int_{\mathbb{U}} du \wedge dv \sqrt{|g|} (\text{rot } A, N), \quad (3.90)$$

где введен ротор (3.65) 1-формы: $[\text{rot } A]^\alpha := \varepsilon^{\alpha\beta\gamma} \partial_\beta A_\gamma$ и $N = N^\alpha \partial_\alpha$ – единичный вектор нормали. Таким образом, получена *формула Стокса*. $\square$

Пример 3.7.8. Пусть в трехмерном евклидовом пространстве $\mathbb{R}^3$ с декартовыми координатами x^α задана 2-форма $A = \frac{1}{2} dx^\alpha \wedge dx^\beta A_{\alpha\beta}$. Рассмотрим двумерную компактную поверхность $\mathbb{U} \hookrightarrow \mathbb{R}^3$ с кусочно гладкой границей $\partial\mathbb{U}$. Будем считать, что поверхность задана параметрически $x^\alpha = x^\alpha(u, v)$. Тогда общая формула Стокса (3.86) принимает вид

$$\frac{1}{2} \int_{\partial\mathbb{U}} dx^\alpha \wedge dx^\beta A_{\alpha\beta} = \int_{\mathbb{U}} dx^1 \wedge dx^2 \wedge dx^3 (\partial_1 A_{23} + \partial_2 A_{31} + \partial_3 A_{12}). \quad (3.91)$$

Как уже отмечалось 2-форма в трехмерном пространстве взаимно однозначно определяется векторным полем $A_{\alpha\beta} = \varepsilon_{\alpha\beta\gamma} X^\gamma$. Интеграл в левой части равенства сводится к интегралу по поверхности от скалярного произведения вектора X^α на единичную нормаль $N = N^\alpha \partial_\alpha$ к поверхности (3.85). Подынтегральное выражение в правой части (3.91) представляет собой дивергенцию от векторного поля

$$\partial_1 A_{23} + \partial_2 A_{31} + \partial_3 A_{12} = \partial_\alpha X^\alpha.$$

Таким образом, окончательно получаем равенство

$$\int_{\partial\mathbb{U}} du \wedge dv \sqrt{|g|} X^\alpha N_\alpha = \int_{\mathbb{U}} dx^1 \wedge dx^2 \wedge dx^3 \partial_\alpha X^\alpha, \quad (3.92)$$

где $\sqrt{|g|}$ – квадратный корень из индуцированной метрики (см. пример 3.7.4). Это есть *формула Гаусса–Остроградского* в трехмерном евклидовом пространстве. $\square$

Глава 4

Метрика

Одним из важнейших понятий в геометрии и физике является дифференциально геометрическая метрика или просто метрика. Трудно переоценить ту роль, которую метрика играет в физических приложениях. С ее помощью строятся инварианты, определяется форма объема. Симметрии метрики являются симметриями пространства-времени, которые определяют сохраняющиеся токи. В частности, фундаментальные законы сохранения энергии-импульса и момента количества движения связаны с наличием у метрики определенных симметрий (см. раздел 20.2).

4.1 Определение и свойства

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$.

Определение. Метрикой $g \in \mathcal{T}_2(\mathbb{M})$ на многообразии $\mathbb{M}$ называется достаточно гладкое ковариантное тензорное поле типа $(0, 2)$, которое является симметричным и невырожденным в каждой точке многообразия:

- 1) $g_{\alpha\beta} = g_{\beta\alpha}$ – симметричность,
- 2) $\det g_{\alpha\beta} \neq 0$ – невырожденность,

где $g = dx^\alpha \otimes dx^\beta g_{\alpha\beta}$ – выражение для метрики в координатном базисе. Метрика называется *римановой*, если матрица $g_{\alpha\beta}(x)$ положительно определена во всех точках $x \in \mathbb{M}$. В противном случае метрика называется *псевдоримановой*. $\square$

Замечание. Метрика в данном определении не совпадает с тем же термином (топологическая метрика), определенном в разделе 1.3.1. Слово метрика широко используется в научной литературе в обоих значениях. Как правило, смысл термина ясен из контекста. Там, где мы хотим подчеркнуть различие терминов, мы будем говорить дифференциально геометрическая метрика для метрики, которая определена в настоящем разделе, и топологическая метрика, определенная в разделе 1.3.1. Далее в подавляющем большинстве случаев термин метрика будет употребляться в дифференциально геометрическом смысле. $\square$

Метрика определяет билинейное невырожденное и симметричное отображение

$$g : \mathcal{X}(\mathbb{M}) \times \mathcal{X}(\mathbb{M}) \ni X, Y \mapsto (X, Y) \in \mathcal{C}^\infty(\mathbb{M}),$$

которое называется *скалярным произведением* векторных полей. В компонентах скалярное произведение векторных полей $X = X^\alpha \partial_\alpha$ и $Y = Y^\alpha \partial_\alpha$ задается сверткой индексов:

$$(X, Y) := X^\alpha Y^\beta g_{\alpha\beta}, \quad (4.1)$$

для которого мы также будем иногда употреблять обозначение $(X, Y) = g(X, Y)$, чтобы отметить метрику, с помощью которой проводится свертка. Скалярное произведение, очевидно, симметрично: $(X, Y) = (Y, X)$. Невырожденность метрики формулируется так: не существует отличного от нуля векторного поля $X \in \mathcal{X}(\mathbb{M})$ такого, что скалярное произведение $(X, Y) = 0$ для всех $Y \in \mathcal{X}(\mathbb{M})$ и во всех точках $x \in \mathbb{M}$. Для римановой метрики скалярное произведение является положительно определенным в каждой точке, т.е. $(X, X) \geq 0$, причем равенство $(X, X) = 0$ имеет место только для тривиального векторного поля $X = 0$. Скалярное произведение базисных векторных полей ∂_α равно компонентам метрики:

$$(\partial_\alpha, \partial_\beta) = g_{\alpha\beta}. \quad (4.2)$$

Поскольку метрика невырождена, то существует *обратная метрика*, т.е. симметричное невырожденное контравариантное тензорное поле $g^{\alpha\beta} \partial_\alpha \otimes \partial_\beta \in \mathcal{T}^2(\mathbb{M})$ типа $(2, 0)$, компоненты которого в каждой точке многообразия удовлетворяют условию

$$g^{\alpha\beta} g_{\beta\gamma} = g_{\gamma\beta} g^{\beta\alpha} = \delta_\gamma^\alpha. \quad (4.3)$$

При преобразовании координат $x^\alpha \rightarrow x^{\alpha'}(x)$ компоненты метрики и ее обратной преобразуются по обычным правилам для тензорных полей:

$$g_{\alpha'\beta'} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} \frac{\partial x^\beta}{\partial x^{\beta'}} g_{\alpha\beta}, \quad (4.4)$$

$$g^{\alpha'\beta'} = g^{\alpha\beta} \frac{\partial x^{\alpha'}}{\partial x^\alpha} \frac{\partial x^{\beta'}}{\partial x^\beta}. \quad (4.5)$$

Отсюда следует, что определитель метрики

$$g(x) := \det g_{\alpha\beta}(x) \quad (4.6)$$

преобразуется по правилу $g' = J^{-2}g$, где J – якобиан преобразования координат (1.69). То есть определитель метрики является скалярной плотностью степени -2 .

В общем случае, ввиду симметрии по индексам, метрика $g_{\alpha\beta}$ задается

$$[g_{\alpha\beta}] = \frac{n(n+1)}{2}$$

произвольными компонентами с единственным условием невырожденности. Здесь и в дальнейшем число компонент тензора мы будем обозначать квадратными скобками.

С помощью метрики и ее обратной можно изменить тип тензора путем опускания или подъема всех или части индексов произвольного тензорного поля.

Пример 4.1.1. Произвольному векторному полю $X \in \mathcal{X}(\mathbb{M})$ можно взаимно однозначно поставить в соответствие 1-форму и наоборот. В компонентах отображение задается следующими формулами:

$$X_\alpha := X^\beta g_{\beta\alpha}, \quad X^\alpha = g^{\alpha\beta} X_\beta. \quad (4.7)$$

Это отображение $\mathcal{X}(\mathbb{M}) \leftrightarrow \Lambda_1(\mathbb{M})$, очевидно, линейно, не зависит от выбора системы координат и, значит, является изоморфизмом $\mathcal{C}^\infty(\mathbb{M})$ -модулей. Операция подъема и опускания индексов естественным образом продолжается на тензоры произвольного типа. $\square$

Замечание. Если на многообразии задана метрика, то между контравариантными и ковариантными тензорными полями одного ранга r существует взаимно однозначное соответствие, точнее, изоморфизм $C^\infty(\mathbb{M})$ -модулей:

$$\mathcal{T}^r(\mathbb{M}) \simeq \mathcal{T}_r(\mathbb{M}) \simeq \mathcal{T}_q^p(\mathbb{M}), \quad \forall p, q : p + q = r.$$

По этой причине будем считать, что тензоры одного ранга, но с различным типом индексов, описывают один и тот же геометрический объект и будем обозначать их, как правило, одной буквой. Например,

$$T_{\alpha\beta} = T_\alpha{}^\gamma g_{\gamma\beta} = T^\gamma{}_\beta g_{\gamma\alpha} = T^{\gamma\delta} g_{\gamma\alpha} g_{\delta\beta},$$

где $T_{\alpha\beta}$ и $T^{\alpha\beta}$ – ковариантные и контравариантные компоненты тензора второго ранга. Компоненты $T_\alpha{}^\beta$ и $T^\alpha{}_\beta$ называются смешанными. Поскольку при наличии метрики появилась возможность опускать и поднимать индексы, то необходимо следить за порядком ковариантных и контравариантных индексов. В общем случае $T_\alpha{}^\beta \neq T^\beta{}_\alpha$. Поэтому не стоит писать контра- и ковариантные индексы один под другим. $\square$

Определение. Скалярное произведение вектора в каждой точке $x \in \mathbb{M}$ с самим собой называется *квадратом* вектора и обозначается

$$X^2 := (X, X). \quad (4.8)$$

Длиной вектора X назовем выражение

$$|X| := \sqrt{|X^2|}. \quad (4.9)$$

Знак модуля в этом определении необходим в том случае, если метрика не является положительно определенной. Пусть X, Y – два векторных поля на римановом многообразии $\mathbb{M}$, т.е. с положительно определенной метрикой. Если в точке $x \in \mathbb{M}$ они отличны от нуля, то определим угол φ между ними в данной точке с помощью следующего соотношения

$$\cos \varphi := \frac{(X, Y)}{\sqrt{X^2 Y^2}}. \quad \square \quad (4.10)$$

Определение угла (4.10) корректно. Действительно, для положительно определенной метрики правая часть соотношения (4.10) не превышает единицу, поскольку справедливо неравенство треугольника

$$(X, Y) \leq \sqrt{X^2 Y^2}.$$

Определение угла инвариантно относительно выбора координат, т.к. правая часть (4.10) содержит только инвариантные комбинации векторных полей.

Определение. Если $(X, Y) = 0$, то два вектора называются *ортгональными* или *перпендикулярными*, независимо от положительной определенности метрики. $\square$

Определение. Две метрики $g' = dx^\alpha \otimes dx^\beta g'_{\alpha\beta}$ и $g = dx^\alpha \otimes dx^\beta g_{\alpha\beta}$ на многообразии $\mathbb{M}$ называются *конформно эквивалентными* или связанными преобразованием Вейля, если они пропорциональны

$$g' = e^{2\phi} g, \quad (4.11)$$

где $\phi(x)$ – некоторая функция на $\mathbb{M}$. Преобразование (4.11) называют *конформным преобразованием* метрики или *преобразованием Вейля*. $\square$

Это определение не зависит от выбора системы координат и поэтому задает вейлевскую (конформную) эквивалентность глобально.

Угол между векторами определяется только метрикой и не изменится, если метрику умножить на произвольную функцию, отличную от нуля (преобразование Вейля).

Замечание. Термин конформная эквивалентность метрик широко используется, однако он не имеет отношения к конформным преобразованиям в комплексном анализе. Конформные преобразования комплексных переменных – это подгруппа общих преобразований координат, в то время как при преобразовании (4.11) координаты не меняются. Поэтому мы будем употреблять термин вейлевская инвариантность, т.к. в двумерных моделях математической физики термин конформная инвариантность употребляется в своем первоначальном значении, как в комплексном анализе. $\square$

Вейлевские преобразования метрики подробно рассмотрены в разделе 17.4.

Поскольку метрика невырождена, то у нее существует обратная метрика (4.3), которая естественным образом определяет скалярное произведение в пространстве 1-форм. Скалярное произведение двух 1-форм $A = dx^\alpha A_\alpha$ и $B = dx^\beta B_\beta$ по определению равно

$$(A, B) = g(A, B) := g^{\alpha\beta} A_\alpha B_\beta. \quad (4.12)$$

Это скалярное произведение согласовано с операцией подъема и опускания индексов (4.7):

$$g^{\alpha\beta} A_\alpha B_\beta = g_{\alpha\beta} A^\alpha B^\beta, \quad \text{где} \quad A^\alpha := g^{\alpha\beta} A_\beta, \quad B^\alpha := g^{\alpha\beta} B_\beta.$$

Скалярные произведения векторов (4.1) и 1-форм (4.12) естественным образом продолжаются на тензорные поля произвольного типа (r, s) .

Метрика на многообразии определяет инвариантную квадратичную форму дифференциалов, которая называется *интервалом*

$$ds^2 := g_{\alpha\beta} dx^\alpha dx^\beta. \quad (4.13)$$

Интервал задает расстояние между двумя бесконечно близкими точками с координатами x^α и $x^\alpha + dx^\alpha$. Это расстояние зависит от точки многообразия, но не от выбора системы координат. В случае псевдоримановой метрики выражение (4.13) не задает метрику (расстояние) в топологическом смысле, определенную в разделе 1.3.1, т.к. квадратичная форма (4.13) не является положительно определенной. С помощью положительно определенной римановой метрики топологическое расстояние между двумя точками $p, q \in \mathbb{M}$ можно определить как точную нижнюю грань интеграла

$$l(p, q) := \inf_p^q \int ds, \quad ds := \sqrt{ds^2} = dt \sqrt{g_{\alpha\beta} \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt}}, \quad (4.14)$$

где интегрирование ведется по всем кусочно дифференцируемым кривым $x(t)$, соединяющим точки p и q . Доказательство свойств расстояния при этом существенно опирается на положительную определенность римановой метрики. Топология, определяемая топологической метрикой (4.14), совпадает с топологией многообразия $\mathbb{M}$.

В заключение раздела обсудим вопрос о существовании римановых метрик.

Теорема 4.1.1. *На любом n -мерном дифференцируемом многообразии $\mathbb{M}$ существует риманова метрика.*

Доказательство. Выберем локально конечный атлас $\{\mathbb{U}_i\}$ на $\mathbb{M}$. Такой атлас всегда существует в силу теоремы 2.2.2. Обозначим координаты на каждой карте через x_i^α , $\alpha = 1, \dots, n$. Пусть $\{f_i\}$ – разбиение единицы, подчиненное данному атласу, такое, что $\text{supp } f_i \subset \mathbb{U}_i$. Выберем евклидову метрику на каждой карте и склеим карты с помощью разбиения единицы

$$ds_i^2 := \delta_{\alpha\beta} dx_i^\alpha dx_i^\beta, \quad (4.15)$$

$$ds^2 := \sum_i f_i ds_i^2, \quad (4.16)$$

где выражение $f_i ds_i^2$ определено

$$(f_i ds_i^2)(x) = \begin{cases} f_i(x) ds_i^2, & x \in \mathbb{U}_i, \\ 0, & x \notin \mathbb{U}_i. \end{cases}$$

Уравнения (4.15), (4.16) определяют гладкое симметричное ковариантное тензорное поле второго ранга на $\mathbb{M}$. Поскольку правая часть (4.16) содержит конечное число слагаемых в каждой точке $x \in \mathbb{M}$, то сумма корректно определена. Выберем координатную окрестность $\mathbb{U}$ с координатами x^α такую, что замыкание области $\bar{\mathbb{U}}$ является компактным. Тогда область $\mathbb{U}$ пересекается с конечным числом карт $\mathbb{U}_{i_1}, \dots, \mathbb{U}_{i_r}$, т.к. атлас является локально конечным покрытием. Поэтому ограничение (4.16) на $\mathbb{U}$ можно записать в виде

$$ds^2 = \sum_{\lambda=1}^r f_{i_\lambda} ds_{i_\lambda}^2 = g_{\alpha\beta} dx^\alpha dx^\beta,$$

где

$$g_{\alpha\beta} := \sum_{\lambda=1}^r f_{i_\lambda} \delta_{\gamma\delta} \frac{\partial x_{i_\lambda}^\gamma}{\partial x^\alpha} \frac{\partial x_{i_\lambda}^\delta}{\partial x^\beta}.$$

Поскольку $0 \leq f_i \leq 1$ и $\sum_i f_i = 1$, то существует такой индекс j , что $f_j(x) > 0$. Поэтому

$$ds^2(x) \geq f_j ds_j^2.$$

Таким образом, метрика ds^2 является положительно определенной на $\mathbb{M}$. $\square$

Замечание. Доказательство этой теоремы свелось к переносу евклидовой метрики δ_{ij} из евклидова пространства $\mathbb{R}^n$ на многообразие $\mathbb{M}$ с помощью возврата отображения, которое фигурирует в определении многообразия, как гомеоморфизм. При этом положительная определенность евклидовой метрики существенна. Доказательство не проходит, если таким же образом попытаться перенести лоренцеву метрику из $\mathbb{R}^{1,n-1}$ на $\mathbb{M}$. $\square$

4.2 Метрика на лоренцевых многообразиях

За счет выбора системы координат метрику всегда можно привести к диагональному виду в любой наперед заданной точке $x \in \mathbb{M}$. Действительно, при замене координат компоненты метрики преобразуются по правилу (4.4). При этом матрица Якоби преобразования координат $\partial x^\alpha / \partial x^{\alpha'}$ в фиксированной точке многообразия может быть выбрана произвольным образом.

Пример 4.2.1. Однородное линейное преобразование координат $x^\alpha = x^{\alpha'} M_{\alpha'}^\alpha$ с постоянной невырожденной матрицей $M_{\alpha'}^\alpha = \text{const}$ дает

$$g_{\alpha'\beta'} = M_{\alpha'}^\alpha M_{\beta'}^\beta g_{\alpha\beta}.$$

□

Замечание. В общем случае метрику можно привести к диагональному виду в фиксированной точке, но не в окрестности. Это связано с тем, что n функций преобразований координат, которыми можно воспользоваться, недостаточно для фиксирования $n(n-1)/2$ функций, параметризующих недиагональные элементы метрики. Исключения составляют многообразия двух и трех измерений. В двумерном случае метрика имеет только одну недиагональную компоненту и ее можно привести к диагональному виду в окрестности произвольной точки. Более того, ее можно преобразовать к конформно плоскому виду, теорема 17.5.1. На трехмерном многообразии метрика имеет три недиагональные компоненты, что равно числу произвольных функций, параметризующих преобразования координат. Можно показать, что ее можно привести к диагональному виду не только в заданной точке, но и в некоторой окрестности этой точки. □

Привести метрику к диагональному виду в точке можно многими способами. При этом сигнатура не зависит от выбора системы координат в которой метрика диагональна. Действительно, если в точке $p \in \mathbb{M}$ метрика диагональна в двух системах координат и при этом имеет различную сигнатуру, то эти системы координат не могут быть связаны никаким преобразованием координат. В различных системах координат диагональные компоненты метрики могут иметь различные (ненулевые) значения, может меняться последовательность положительных и отрицательных компонент, однако число положительных и отрицательных компонент остается неизменным.

Определение. *Сигнатурой метрики*, заданной на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, называется пара натуральных чисел (p, q) таких, что $p + q = n$, где p и q – количество, соответственно, положительных и отрицательных чисел, стоящих на диагонали метрики после ее диагонализации в какой-либо точке многообразия $x \in \mathbb{M}$. □

Если $g_{\alpha\beta}$ – матрица, составленная из компонент метрики в некоторой системе координат, то числа p и q равны, соответственно, числу положительных и отрицательных собственных значений. При общих преобразованиях координат собственные числа могут менять свою величину, но не знак. Нулевых собственных значений быть не может, т.к. в этом случае метрика была бы вырожденной.

Предложение 4.2.1. *Сигнатура метрики не зависит от точки линейно связного многообразия.*

Доказательство. Допустим, что в некоторых точках $x_1, x_2 \in \mathbb{M}$ метрика имеет разную сигнатуру, и соединим эти точки произвольной кривой. Тогда из непрерывности метрики следует, что ее определитель обратился бы в нуль в некоторой точке кривой, что недопустимо. □

Замечание. В зависимости от сигнатуры метрики скалярное произведение векторов может быть положительно определено или нет. Риманова метрика на многообразии размерности n имеет сигнатуру $(n, 0)$. Отрицательно определенная метрика имеет сигнатуру $(0, n)$. В римановой геометрии $(\mathbb{M}, g)$ положительно и отрицательно

определенные метрики, по существу, можно не различать, т.к. все геометрические характеристики многообразий связаны простым преобразованием – изменением знака метрики. Поэтому многообразия с отрицательно определенной метрикой мы также будем называть римановыми. В то же время многие модели математической физики инвариантны относительно изменения знака метрики

$$g_{\alpha\beta} \mapsto -g_{\alpha\beta}, \quad (4.17)$$

поскольку действие, которым описывается модель, помимо метрики содержит также другие поля. Поэтому решения уравнений Эйлера–Лагранжа с положительно и отрицательно определенной метрикой, вообще говоря, не эквивалентны. $\square$

Сигнатура метрики инвариантна, не зависит от выбора системы координат и точки линейно связного многообразия. Будем считать, что после диагонализации номера координат выбраны таким образом, что сначала идут все положительные, а затем – отрицательные собственные значения. Будем писать

$$\text{sign } g_{\alpha\beta} = (\underbrace{+\cdots+}_p \underbrace{-\cdots-}_q). \quad (4.18)$$

Определение. Пара $(\mathbb{M}, g)$ называется *римановым многообразием*, если метрика является знакоопределенной $p = n$ или $q = n$, и *псевдоримановым многообразием*, если метрика не является знакоопределенной, $p \neq 0$ и $q \neq 0$. Если положительный элемент на диагонали один,

$$\text{sign } g_{\alpha\beta} = (+ - \cdots -),$$

то говорят, что метрика имеет *лоренцеву сигнатуру*. Если на многообразии задана метрика лоренцевой сигнатуры, то будем говорить, что многообразие *лоренцево*. $\square$

Подчеркнем, что задание метрики на многообразии может быть произвольным. В частности, на одном многообразии можно задать несколько метрик одновременно, причем разной сигнатуры, если такие существуют.

Если на четырехмерном лоренцевом многообразии рассматриваются спинорные поля, то общий знак метрики может играть существенную роль. В разделе 26.2 будет показано, что алгебры Клиффорда $\text{cl}(1, 3, \mathbb{R})$ и $\text{cl}(3, 1, \mathbb{R})$, которые лежат в основе определения спинорных полей, не изоморфны.

В настоящей монографии псевдоримановы многообразия размерности $n \geq 4$ с $p \geq 2$ и $q \geq 2$ не рассматриваются, т.к. они недостаточно хорошо изучены и не имеют широкого применения в математической физике.

Для лоренцевой метрики скалярное произведение двух векторов может быть положительно, отрицательно или равно нулю, а из условия $X^2 = 0$ не следует, что $X = 0$.

Определение. Назовем *пространством-временем* псевдориманово многообразие $\mathbb{M}$ с заданной метрикой $g = dx^\alpha \otimes dx^\beta g_{\alpha\beta}$ лоренцевой сигнатуры (лоренцево многообразие). В пространстве-времени векторное поле X в точке $x \in \mathbb{M}$ называется:

$$\begin{array}{lll} \text{времениподобным,} & \text{если} & (X, X) > 0, \\ \text{светоподобным (изотропным, нулевым),} & \text{если} & (X, X) = 0, \\ \text{пространственноподобным,} & \text{если} & (X, X) < 0. \end{array} \quad (4.19)$$

Это определение распространяется на область $\mathbb{U} \subset \mathbb{M}$, если во всех точках $x \in \mathbb{U}$ выполнены соответствующие соотношения. $\square$

Определение (4.19) инвариантно относительно замены координат. В общем случае гладкое векторное поле может иметь различный тип в различных областях связного многообразия.

Определение. Так же, как и в римановом пространстве, два вектора в пространстве-времени называются *ортогональными* (*перпендикулярными*), если их скалярное произведение равно нулю. В частности, любой изотропный вектор перпендикулярен самому себе.

Определим тип координатного векторного поля ∂_0 . Из (4.2) следует, что $(\partial_0, \partial_0) = g_{00}$. Если в данной системе координат $g_{00} > 0$, то векторное поле ∂_0 времениподобно. В этом случае назовем координату $x^0 = t$ *времениподобной* или *временем*. Другими словами, временем называется любой параметр вдоль интегральной кривой произвольного времениподобного векторного поля. Противоположно направленное векторное поле $-\partial_0$ также времениподобно и связано с ∂_0 преобразованием координат $(x^0, x^1, \dots, x^{n-1}) \mapsto (-x^0, x^1, \dots, x^{n-1})$, которое назовем *обращением времени*. Выберем (произвольно) ориентацию координаты x^0 и будем говорить, что векторное поле ∂_0 направлено в будущее, а $-\partial_0$ – в прошлое. Тогда на времениподобных векторных полях можно ввести ориентацию. Произвольное времениподобное поле X направлено в будущее, если $(X, \partial_0) = X^0 > 0$. В противном случае, когда $X^0 < 0$, будем говорить, что времениподобное векторное поле направлено в прошлое. $\square$

Замечание. В общем случае метрика, индуцированная на сечениях $t = x^0 = \text{const}$ лоренцева многообразия, может быть отрицательно определена или быть знаконеопределенной. То есть сечения пространства-времени, соответствующие постоянному времени, совсем не обязательно пространственноподобны (= все касательные векторы пространственноподобны). Это зависит от выбора остальных координат. В конце настоящего раздела мы рассмотрим простой пример. $\square$

Можно доказать, что каждая точка многообразия M лоренцевой сигнатуры $(1, n-1)$ имеет такую координатную окрестность с координатами $(t, u^1, \dots, u^{n-1})$, что в этих координатах метрика имеет блочно диагональный вид

$$ds^2 = dt^2 + g_{\mu\nu}(t, u) du^\mu du^\nu, \quad \mu, \nu = 1 \dots n-1. \quad (4.20)$$

При этом метрика $g_{\mu\nu}$, индуцированная на сечениях $t = \text{const}$, является отрицательно определенной. Другими словами, все сечения $t = \text{const}$ для метрики вида (4.20) являются пространственноподобными подмногообразиями, т.е. все касательные векторы к сечениям пространственноподобны. Эти сечения являются вложенными римановыми многообразиями с локальной системой координат $\{u^\mu\}$ и отрицательно определенной метрикой $g_{\mu\nu}$. Отметим также, что все векторы, касательные к этому подмногообразию ортогональны времениподобному векторному полю ∂_0 .

Замечание. Здесь и в дальнейшем мы будем использовать следующие обозначения координат на лоренцевом многообразии. Греческие буквы из начала алфавита $\alpha, \beta, \dots$ будут использоваться для нумерации всех координат, а буквы из середины алфавита $\mu, \nu, \dots$ – для нумерации остальных координат, $\{x^\alpha\} = \{x^0, x^\mu\}$. Это правило легко запомнить по следующим включениям:

$$\{1, \dots, n-1\} \subset \{0, 1, \dots, n-1\}, \quad \{\mu, \nu, \dots\} \subset \{\alpha, \beta, \dots\}.$$

Как правило, мы будем считать, что координата x^0 является временем, и все остальные координатные линии x^μ – пространственноподобны. Поэтому координаты x^μ будем называть пространственноподобными. $\square$

В общем случае метрика на лоренцевом многообразии

$$g_{\alpha\beta} = \begin{pmatrix} g_{00} & g_{0\nu} \\ g_{\mu 0} & g_{\mu\nu} \end{pmatrix} \quad (4.21)$$

не имеет блочно диагонального вида (4.20). Приведем критерий того, что метрика (4.21), заданная на некотором многообразии $\mathbb{M}$, имеет лоренцеву сигнатуру.

Предложение 4.2.2. Пусть в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ задана метрика (4.21) такая, что $g_{00} > 0$. Эта метрика имеет лоренцеву сигнатуру тогда и только тогда, когда матрица

$$g_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}} \quad (4.22)$$

отрицательно определена в каждой точке $\mathbb{U}$.

Доказательство. Достаточно рассмотреть произвольную точку из $\mathbb{U}$. Пусть на $\mathbb{U}$ задана метрика (4.21) для которой $g_{00} > 0$. Интервал в окрестности $\mathbb{U} \subset \mathbb{M}$ имеет вид

$$ds^2 = g_{00}dx^0dx^0 + 2g_{0\mu}dx^0dx^\mu + g_{\mu\nu}dx^\mu dx^\nu, \quad g_{00} > 0.$$

Зафиксируем произвольную точку $x \in \mathbb{U}$. Введем вместо x^0 новую координату $\tilde{x}^0$, для которой в точке $x \in \mathbb{U}$ выполнено соотношение

$$dx^0 = d\tilde{x}^0 - \frac{dx^\mu g_{0\mu}}{g_{00}}.$$

В фиксированной точке этого всегда можно добиться линейным преобразованием координат. Тогда интервал примет вид

$$ds^2 = g_{00}d\tilde{x}^0d\tilde{x}^0 + \left(g_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}}\right)dx^\mu dx^\nu. \quad (4.23)$$

Если метрика (4.21) имеет лоренцеву сигнатуру, то существует такая система координат, что в точке $x \in \mathbb{M}$ метрика диагональна, причем $g_{00} > 0$, а все остальные диагональные компоненты $g_{\mu\mu}$ отрицательны. Поскольку метрика (4.23) связана с диагональной метрикой также невырожденным преобразованием координат, то матрица (4.22) отрицательно определена.

Обратно. Если матрица (4.22) отрицательно определена, то дальнейшим линейным преобразованием координат x^μ ее всегда можно преобразовать к диагональному виду в фиксированной точке, причем на диагонали будут стоять отрицательные числа. Следовательно, метрика имеет лоренцеву сигнатуру. $\square$

Замечание. В разделе 24.2 мы докажем, что отрицательная определенность матрицы (4.22) эквивалентна отрицательной определенности “пространственного блока” $g^{\mu\nu}$ обратной метрики

$$g^{\alpha\beta} = \begin{pmatrix} g^{00} & g^{0\nu} \\ g^{\mu 0} & g^{\mu\nu} \end{pmatrix}. \quad \square$$

Матрица (4.22) на лоренцевом многообразии симметрична и невырождена, т.к. отрицательно определена. Она имеет следующий геометрический смысл. Вдоль координатных линий времени $t = x^0$ всегда можно определить единичное векторное поле

$$n := \frac{1}{\sqrt{g_{00}}}\partial_0, \quad (n, n) = 1.$$

Рассмотрим произвольное векторное поле $X = X^0 \partial_0 + X^\mu \partial_\mu$. У него есть составляющая, перпендикулярная времениподобному векторному полю n ,

$$X_\perp = X - (X, n)n = X^\mu \partial_\mu - X^\mu g_{0\mu} \partial_0.$$

Если задано два произвольных векторных поля X и Y , то скалярное произведение их перпендикулярных составляющих равно

$$(X_\perp, Y_\perp) = \left(g_{\mu\nu} - \frac{g_{0\mu} g_{0\nu}}{g_{00}} \right) X^\mu Y^\nu.$$

Таким образом, матрица (4.22) играет роль метрики для перпендикулярных составляющих векторных полей. Из теоремы 4.2.2 следует, что, если $\mathbb{M}$ – лоренцево многообразие и x^0 – время, то составляющие векторных полей, перпендикулярные n , всегда пространственноподобны.

Замечание. Не следует думать, что матрицу (4.22) можно рассматривать, как метрику на некотором $(n - 1)$ -мерном подмногообразии, касательные векторы к которому всюду перпендикулярны времениподобному векторному полю n . Дело в том, что такие подмногообразия могут не существовать. Можно проверить, что коммутатор двух перпендикулярных векторных полей $[X_\perp, Y_\perp]$ в общем случае не будет ортогонален n . Для этого достаточно, не ограничивая общности, положить $g_{00} = 1$ и проверить, что коммутатор $[X_\perp, Y_\perp]$ имеет нетривиальную составляющую вдоль n . Тем самым множество всех векторных полей $X_\perp$ не находится в инволюции, и, согласно теореме Фробениуса, подмногообразия, касательные векторы к которому всюду перпендикулярны n , не существует. $\square$

Приведем еще один критерий того, что метрика на многообразии является лоренцевой.

Теорема 4.2.1. Для того, чтобы метрика (4.21) с $g_{00} > 0$ в точке $x \in \mathbb{M}$ имела и достаточно, чтобы выполнялись неравенства:

$$\det \begin{pmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{pmatrix} < 0, \quad \det \begin{pmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{pmatrix} > 0, \quad \dots, \quad \det(g_{\alpha\beta}) \begin{cases} > 0, & n - \text{нечетно}, \\ < 0, & n - \text{четно}. \end{cases} \quad (4.24)$$

Доказательство. См., например, [11], часть II, глава 1, теорема 7. $\square$

Пример 4.2.2. Рассмотрим плоскость Минковского $\mathbb{R}^{1,1}$ с декартовой системой координат $\{x^0, x^1\}$ и метрикой

$$\eta_{\alpha\beta} := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Введем новую систему координат $\{\tilde{x}^0, \tilde{x}^1\} = \{x^0 - \sqrt{2}x^1, x^1\}$. Тогда координатные базисные векторы преобразуются по правилу (см. рис. 4.1)

$$\{\partial_0, \partial_1\} \mapsto \{\tilde{\partial}_0, \tilde{\partial}_1\} = \{\partial_0, \sqrt{2}\partial_0 + \partial_1\}.$$

В новой системе координат сечения $\tilde{x}^0 = \text{const}$ являются времениподобными прямыми, а метрика имеет вид

$$g_{\alpha\beta} = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 1 \end{pmatrix}.$$

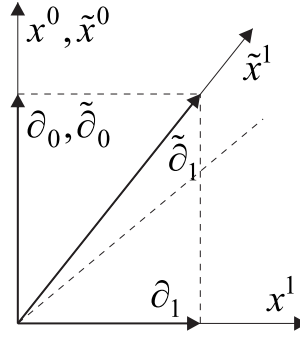


Рис. 4.1: Пример времениподобного сечения $\tilde{x}^0 = \text{const}$.

Матрица, соответствующая выражению (4.22), состоит из одного элемента $g_{11} - g_{01}^2/g_{00} = -1$.

Обратная метрика равна

$$g^{\alpha\beta} = \begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & -1 \end{pmatrix}.$$

Отметим, что в этих координатах обе компоненты обратной метрики на диагонали g^{00} и g^{11} отрицательны. $\square$

Рассмотренный пример показывает, что “пространственные” компоненты метрики $g_{\mu\nu}$ в общем случае не образуют отрицательно определенной матрицы. Роль пространственной части метрики играет квадратичная форма (4.22), и она должна быть отрицательно определена на лоренцевом многообразии. Кроме того, из примера следует, что, даже если координата x^0 на лоренцевом многообразии является временем, тем не менее временная компонента обратной метрики g^{00} может быть отрицательна.

В дальнейшем мы всегда предполагаем, что координаты на лоренцевом многообразии выбраны таким образом, что координата x^0 является временем, и все сечения $x^0 = \text{const}$ пространственноподобны. Такой выбор координат удобен, например, при постановке задачи Коши в различных моделях математической физики. На языке компонент метрики эти условия означают, что $g_{00} > 0$, и матрица $g_{\mu\nu}$ отрицательно определена. В этом случае нетрудно доказать, что временная компонента обратной метрики также положительна $g^{00} > 0$. В обратную сторону утверждение неверно: условий положительности $g_{00} > 0$ и $g^{00} > 0$ недостаточно для отрицательной определенности матрицы $g_{\mu\nu}$.

4.3 Векторные поля и вложения

В разделе 2.6.6 было установлено, что отличная от нуля 1-форма $A = dx^\alpha A_\alpha$ определяет в касательном расслоении $\mathbb{T}(\mathbb{M})$ распределение $(n - 1)$ -подпространств, натянутое на линейно независимые векторы с компонентами Y^α , удовлетворяющими уравнению $Y^\alpha A_\alpha = 0$. Наличие метрики позволяет каждой 1-форме поставить в соответствие векторное поле $X^\alpha = g^{\alpha\beta} A_\beta$. По построению, это векторное поле перпендикулярно любому вектору из $(n - 1)$ -мерного подпространства в касательном пространстве, т.к. $X^\alpha Y^\beta g_{\alpha\beta} = 0$. Таким образом, при наличии метрики задание ненулевого векторного поля эквивалентно заданию $(n - 1)$ -мерного распределения на многообразии. В общем случае это распределение не является инволютивным, и поэтому у него не существует интегральных подмногообразий. То есть векторное поле

не определяет ортогональные интегральные подмногообразия. Обратное утверждение верно. Каждое $(n - 1)$ -мерное подмногообразие (гиперповерхность) определяет нормальное векторное поле, заданное на подмногообразии.

Определение. Пусть $\mathbb{M}$ – риманово многообразие. Тогда *углом* между двумя пересекающимися гиперповерхностями в точках их пересечения называется угол между соответствующими нормальными векторами в данных точках. $\square$

Определение. Если на многообразии задана метрика и векторное поле, то определим проекционные операторы на заданное векторное поле и перпендикулярные направления следующим образом. Пусть X^α – компоненты произвольного векторного поля, такого, что $X^2 \neq 0$. Тогда операторы

$$\Pi_\alpha^{\text{L}\beta} := \frac{X_\alpha X^\beta}{X^2}, \quad \Pi_\alpha^{\text{T}\beta} := \delta_\alpha^\beta - \frac{X_\alpha X^\beta}{X^2}, \quad (4.25)$$

являются *проекционными операторами* на направление векторного поля и перпендикулярные направления, соответственно. При этом выполнены равенства, определяющие набор проекционных операторов:

$$(\Pi^{\text{L}})^2 = \Pi^{\text{L}}, \quad (\Pi^{\text{T}})^2 = \Pi^{\text{T}}, \quad \Pi^{\text{L}}\Pi^{\text{T}} = 0, \quad \Pi^{\text{L}} + \Pi^{\text{T}} = 1. \quad \square$$

Пример 4.3.1. Векторное поле $Y = Y^\alpha \partial_\alpha$ имеет следующее разложение

$$Y = Y^{\text{L}} + Y^{\text{T}},$$

где его компоненты определены следующими выражениями:

$$Y^{\text{L}\alpha} := Y^\beta \Pi_\beta^{\text{L}\alpha} = \frac{(Y, X) X^\alpha}{X^2},$$

$$Y^{\text{T}\alpha} := Y^\beta \Pi_\beta^{\text{T}\alpha} = Y^\alpha - \frac{(Y, X) X^\alpha}{X^2}.$$

Аналогично раскладывается 1-форма $A = dx^\alpha A_\alpha$:

$$A = A^{\text{L}} + A^{\text{T}},$$

где

$$A_\alpha^{\text{L}} := \Pi_\alpha^{\text{L}\beta} A_\beta = \frac{A(X) X_\alpha}{X^2},$$

$$A_\alpha^{\text{T}} := \Pi_\alpha^{\text{T}\beta} A_\beta = A_\alpha - \frac{A(X) X_\alpha}{X^2}.$$

Для того, чтобы разложить тензоры более высокого ранга, необходимо произвести разложение по каждому индексу. Например, контравариантный тензор второго ранга T имеет следующее разложение

$$T = T^{\text{LL}} + T^{\text{LT}} + T^{\text{TL}} + T^{\text{TT}},$$

где

$$T^{\text{LL}\alpha\beta} := T^{\gamma\delta} \Pi_\gamma^{\text{L}\alpha} \Pi_\delta^{\text{L}\beta} = \frac{T^{\gamma\delta} X_\gamma X_\delta X^\alpha X^\beta}{X^4},$$

$$T^{\text{LT}\alpha\beta} := T^{\gamma\delta} \Pi_\gamma^{\text{L}\alpha} \Pi_\delta^{\text{T}\beta} = \frac{T^{\gamma\beta} X_\gamma X^\alpha}{X^2} - \frac{T^{\gamma\delta} X_\gamma X_\delta X^\alpha X^\beta}{X^4},$$

$$T^{\text{TL}\alpha\beta} := T^{\gamma\delta} \Pi_\gamma^{\text{T}\alpha} \Pi_\delta^{\text{L}\beta} = \frac{T^{\alpha\gamma} X_\gamma X^\beta}{X^2} - \frac{T^{\gamma\delta} X_\gamma X_\delta X^\alpha X^\beta}{X^4},$$

$$T^{\text{TT}\alpha\beta} := T^{\gamma\delta} \Pi_\gamma^{\text{T}\alpha} \Pi_\delta^{\text{T}\beta} = T^{\alpha\beta} - \frac{T^{\alpha\gamma} X_\gamma X^\beta}{X^2} - \frac{T^{\gamma\beta} X_\gamma X^\alpha}{X^2} + \frac{T^{\gamma\delta} X_\gamma X_\delta X^\alpha X^\beta}{X^4}.$$

Аналогично раскладываются тензоры с произвольным числом контра- и ковариантных индексов. $\square$

Поскольку нетривиальные ковекторы с компонентами

$$V_\alpha - X_\alpha \frac{(V, X)}{X^2} \quad \text{и} \quad X_\alpha$$

являются собственными векторами проекционных операторов Π^L и Π^T с нулевыми собственными значениями, то проекционные операторы вырождены:

$$\det \Pi^L = \det \Pi^T = 0.$$

Рассмотрим, как меняется метрика при отображениях. Пусть многообразие $\mathbb{M}$ вложено¹ в $\mathbb{N}$, $\dim \mathbb{M} \leq \dim \mathbb{N}$, и на пространстве мишени $\mathbb{N}$ задана метрика g . Тогда возврат отображения $f : \mathbb{M} \hookrightarrow \mathbb{N}$ задает метрику на $\mathbb{M}$. Бескоординатное определение индуцированной метрики f^*g имеет вид

$$(f^*g)(X, Y) := g(f_*X, f_*Y), \quad \forall X, Y \in \mathcal{X}(\mathbb{M}),$$

где f_* – дифференциал отображения, отображающий векторные поля X, Y на $\mathbb{M}$ в векторные поля f_*X, f_*Y на $\mathbb{N}$. Скалярные произведения на $\mathbb{M}$ и $\mathbb{N}$ заданы, соответственно, метриками f^*g и g . Это значит, что при вложении длины векторов и углы между ними сохраняются. Обозначив координаты на $\mathbb{M}$ и $\mathbb{N}$ через x^α и y^A , получим явное выражение для компонент индуцированной метрики f^*g :

$$(f^*g)_{\alpha\beta}(x) = \frac{\partial y^A}{\partial x^\alpha} \frac{\partial y^B}{\partial x^\beta} g_{AB}(y). \quad (4.26)$$

Определение. Пусть на многообразии $\mathbb{M}$ задана некоторая метрика h . Если при вложении $\mathbb{M}$ в многообразие $\mathbb{N}$ с метрикой g исходная метрика совпадает с индуцированной, $h = f^*g$, то такое вложение называется *изометрическим*. $\square$

Представляет большой интерес задача о нахождении изометрических вложений заданного многообразия с определенной на нем метрикой в евклидово пространство.

Пример 4.3.2. Пусть двумерная сфера $\mathbb{S}_r^2$ радиуса r вложена в трехмерное евклидово пространство $\mathbb{R}^3$. Вложение можно записать в параметрическом виде

$$\begin{aligned} x &= r \sin \theta \cos \varphi, \\ y &= r \sin \theta \sin \varphi, \\ z &= r \cos \theta, \end{aligned}$$

где $0 \leq \theta \leq \pi$ и $0 \leq \varphi < 2\pi$ – координаты на сфере и $r = \text{const} > 0$. Тогда метрика, индуцированная на сфере, примет вид

$$ds^2 = dx^2 + dy^2 + dz^2 = r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2. \quad (4.27)$$

То есть в выражение для евклидова интервала вместо дифференциалов dx , dy и dz необходимо просто подставить их выражения через дифференциалы полярного и азимутального углов. $\square$

¹На самом деле достаточно погружения.

Пример 4.3.3. Пусть двумерный тор $\mathbb{T}^2$ радиуса r с направляющей окружностью радиуса $R > r$ вложен в трехмерное евклидово пространство $\mathbb{R}^3$ (см. рис. 4.2):

$$\begin{aligned}x &= (R + r \cos \theta) \cos \varphi, \\y &= (R + r \cos \theta) \sin \varphi, \\z &= r \sin \theta,\end{aligned}$$

где $0 \leq \theta < 2\pi$ и $0 \leq \varphi < 2\pi$ – координаты на торе. Тогда индуцированная метрика

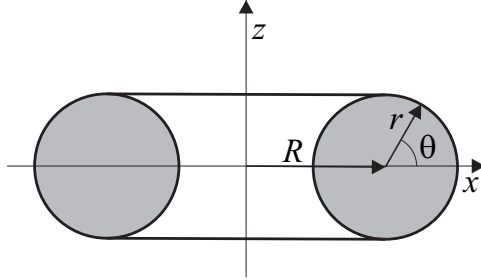


Рис. 4.2: Сечение двумерного тора, вложенного в трехмерное евклидово пространство.

имеет вид

$$ds^2 = dx^2 + dy^2 + dz^2 = r^2 d\theta^2 + (R + r \cos \theta)^2 d\varphi^2. \quad \square \quad (4.28)$$

4.4 Выбор системы координат

Если на многообразии задана метрика, которую преобразованием координат нельзя привести к евклидовой или лоренцевой форме, тогда понятие декартовой системы координат отсутствует, и это создает определенные трудности для наглядного представления многообразия. С практической точки зрения удобнее сначала совершить преобразование координат, после которого метрика примет какой-либо относительно простой вид, связанный, например, с симметрией задачи, и уже потом дать физическую интерпретацию пространству или пространству-времени.

Определение. Риманова метрика называется *конформно евклидовой* или *вейлевски евклидовой* в некоторой области $\mathbb{U} \subset \mathbb{M}$, если существует такая система координат на $\mathbb{U}$, в которой метрика имеет вид

$$g_{\alpha\beta} = e^{2\phi} \delta_{\alpha\beta}. \quad (4.29)$$

Здесь $\phi(x)$ – некоторая достаточно гладкая функция. Функция $e^{2\phi}$ называется *конформным множителем*. В таком виде метрика невырождена при $\phi \neq \pm\infty$. Если у каждой точки $x \in \mathbb{M}$ существует окрестность, на которой метрика является конформно евклидовой, то метрика называется *локально конформно евклидовой* или *локально вейлевски евклидовой*. $\square$

Метрика, обратная к метрике (4.29) имеет вид

$$g^{\alpha\beta} = e^{-2\phi} \delta^{\alpha\beta}.$$

Замечание. Подчеркнем, что $\phi(x)$ не является скалярным полем, хотя и не имеет ни одного индекса, т.к. в другой системе координат конформно евклидова метрика может не иметь вида (4.29), и, значит, функция ϕ не определена вовсе. $\square$

Если метрика записана в конформно евклидовом виде, то в этой системе координат угол (4.10) между двумя произвольными векторами тот же, что и в евклидовом пространстве. Это обстоятельство позволяет более наглядно представить себе свойства многообразия.

Конформно евклидова метрика является частным случаем метрики, и далеко не каждую метрику можно привести к такому виду путем преобразования системы координат. Это видно уже из того, что в общем случае конформно евклидова метрика зависит от $n+1$ независимой функции (здесь n функций соответствует выбору системы координат и одна функция – конформному множителю), в то время как метрика общего вида зависит от $n(n+1)/2$ функций. Исключение представляет только случай двух измерений.

Теорема 4.4.1. *Если риманова метрика принадлежит классу $C^3(\mathbb{M})$ на двумерном многообразии $\mathbb{M}$ (поверхности), то для любой точки $x \in \mathbb{M}$ существует окрестность $\mathbb{U}_x$, в которой можно выбрать конформно евклидовы координаты.*

Доказательство. См., например, [54], теорема 2.5.14. □

Локально конформно евклидовы координаты x^1, x^2 в теореме 4.4.1 определены на некоторой окрестности $\mathbb{U}_x$ с точностью до конформных преобразований комплексных координат, которые вводятся на $\mathbb{U}_x$ следующим образом: $z := x^1 + ix^2$, что оправдывает название “конформно евклидова” для метрики.

Определение. Если в пространстве-времени метрика имеет вид (4.29), где символ Кронекера $\delta_{\alpha\beta}$ заменен на лоренцеву метрику $\eta_{\alpha\beta}$, то пространство-время называется *конформно минковским*. □

В общем случае, для того, чтобы задать метрику в некоторой окрестности, необходимо задать $n(n+1)/2$ функций от n переменных (компоненты метрики). Во многих случаях, например, в общей теории относительности, метрика ищется как решение некоторой системы уравнений, которая ковариантна относительно выбора системы координат. То есть система уравнений движения выглядит по-разному в различных системах отсчета, однако пространства решений находятся во взаимно однозначном соответствии. При этом соответствие устанавливается преобразованием координат. При решении ковариантной системы уравнений удобно выбрать ту или иную систему отсчета с тем, чтобы уменьшить число неизвестных функций. Поскольку переход между системами координат характеризуется n функциями от n переменных, то максимальное число функций, которое можно зафиксировать в метрике общего вида, равно n . В результате метрика будет определяться $n(n-1)/2$ функциями. В этом случае говорят, что выбрана система координат или *зафиксирована калибровка* для метрики². Как правило, координаты выбираются таким образом, чтобы система уравнений имела наиболее простой вид.

Выбор системы координат не означает, что n функций, входящих в метрику, можно зафиксировать произвольным образом. Для того, чтобы определить, существует ли система отсчета, где метрика имеет заданный вид, необходимо проанализировать систему уравнений для функций перехода к такой системе координат. Если эта система уравнений имеет решение, то такая калибровка называется *допустимой*. Часто система уравнений на функции перехода, имеет только локальные решения.

²Это название пришло из электродинамики, инвариантной относительно локальных преобразований, которые называются калибровочными (см. раздел 25.3).

Это означает, что соответствующая система координат может быть выбрана только локально.

Различные удобные выборы систем координат в общей теории относительности рассмотрены в разделе 23.13.

Глава 5

Связность на векторном расслоении и расслоении реперов

Большинство моделей математической физики сводится к решению некоторых дифференциальных уравнений, которыми являются уравнения движения или уравнения равновесия. Мы предполагаем, что дифференциальные уравнения являются ковариантными объектами, т.е. при преобразовании координат преобразуются по тензорным правилам. В этом случае физические следствия не будут зависеть от выбора системы отсчета. Чтобы строить такие модели используется понятие связности и соответствующей ковариантной производной, т.к. обычные частные производные от тензорных полей на многообразии не приводят к тензорным объектам. Это относится не только к уравнениям общей теории относительности, но и к другим моделям математической физики. По сути дела ковариантные производные появляются и в плоском пространстве при переходе к криволинейным системам координат.

5.1 Векторные расслоения

В разделе (2.4) были определены расслоения общего вида, когда типичным слоем является произвольное многообразие. Ниже мы рассмотрим частный случай расслоений, типичными слоями которых являются векторные пространства, и введем понятие связности. Рассмотренные ранее касательное $T(M)$, кокасательное $T^*(M)$ и тензорные $T_s^r(M)$ расслоения с базой M представляет собой частные примеры векторных расслоений, а понятие связности приводит к оператору ковариантного дифференцирования произвольных тензорных полей – исключительно важному инструменту для построения моделей математической физики.

Определение. Дифференцируемое многообразие $E(M, \pi, V)$ называется *векторным расслоением* с базой M , проекцией π и типичным слоем V , где M – дифференцируемое многообразие и $V \simeq \mathbb{R}^n$ – векторное пространство. Проекция π является дифференцируемым сюръективным отображением $E \xrightarrow{\pi} M$. При этом требуется, чтобы для любого атласа $M = \bigcup_i U_i$ существовали отображения $\{\chi_i\}$, удовлетворяющие следующим условиям:

- 1) Отображение χ_i есть диффеоморфизм $\chi_i : \pi^{-1}(U_i) \rightarrow U_i \times V$ такой, что $\pi \circ \chi_i^{-1}(x, v) = x$ для всех точек $x \in U_i$ и $v \in V$.

- 2) Сужение отображения χ_i на каждый слой $\chi_{i,x} : \pi^{-1}(x) \rightarrow \mathbb{V}$ есть гомоморфизм векторных пространств для всех $x \in \mathbb{U}_i$. При этом, если $x \in \mathbb{U}_i \cap \mathbb{U}_j$, то отображение

$$f_{ji}(x) := \chi_{j,x} \chi_{i,x}^{-1} : \mathbb{V} \rightarrow \mathbb{V},$$

является автоморфизмом, т.е. $f_{ij}(x) \in \mathrm{GL}(n, \mathbb{R})$.

- 3) Если $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$, то отображение $f_{ij}(x) : \mathbb{U}_i \cap \mathbb{U}_j \rightarrow \mathrm{GL}(n, \mathbb{R})$ является достаточно гладким.

Отображения $\{f_{ij}\}$ называются *функциями перехода*. □

Пусть $p \in \pi^{-1}(x)$ – некоторая точка векторного расслоения $\mathbb{E}$ из слоя над $x \in \mathbb{U}_i \cap \mathbb{U}_j$. Тогда $\chi_i(p) = (x, v_i)$ и $\chi_j(p) = (x, v_j)$. Следовательно,

$$p = \chi_i^{-1}(x, v_i) = \chi_j^{-1}(x, v_j). \quad (5.1)$$

Зафиксируем базис $\hat{e}_a$, $a = 1, \dots, n$ в векторном пространстве $\mathbb{V}$ и разложим по нему векторы: $v_i = v_i^a \hat{e}_a$ и $v_j = v_j^a \hat{e}_a$. Тогда из условия 2) следует равенство

$$v_j^a = v_i^b f_{ij}^a{}_b,$$

где введена невырожденная матрица $f_{ij}(x) = \{f_{ij}^a{}_b(x)\} \in \mathrm{GL}(n, \mathbb{R})$. Эта матрица достаточно гладко зависит от точки x в силу условия 3 из определения расслоения.

Замечание. В данном определении оба свойства из определения расслоения общего вида в разделе 2.4 объединены в одно свойство 1). Свойство 2) предполагает, что слой векторного расслоения $\pi^{-1}(x)$ в каждой точке $x \in \mathbb{M}$, сам является векторным пространством, *гомоморфным* типичному слою, а в областях пересечения карт допускаются автоморфизмы, т.е. линейные отображения $\mathbb{V}$ на себя. С учетом свойства 1) получаем, что каждый слой изоморфен векторному пространству, $\pi^{-1}(x) \simeq \mathbb{V}$. В условии 2) гомоморфизм можно заменить на гомеоморфизм и не требовать изначально наличия структуры векторного пространства в слое $\pi^{-1}(x)$. Гомеоморфизма достаточно для переноса структуры векторного пространства из $\mathbb{V}$ на слой $\pi^{-1}(x)$, при этом каждый слой будет, конечно, изоморфен $\mathbb{V}$. Свойство 3) говорит о том, что автоморфизмы задаются матрицей $\mathrm{GL}(n, \mathbb{R})$, элементы которой достаточно гладко зависят от координат. Поэтому в определении векторного расслоения был выбран атлас вместо окрестности точки $x \in \mathbb{M}$ в определении расслоения общего вида. □

Если в векторном пространстве $\mathbb{V}$, $\dim \mathbb{V} = n$, зафиксирован базис $\{\hat{e}_a\}$, $a = 1, \dots, n$, то его можно отождествить с евклидовым пространством $\mathbb{R}^n$. После этого топология и дифференцируемая структура евклидова пространства с помощью отображений $\{\chi_i\}$ естественным образом переносятся на пространство расслоения $\mathbb{E}$, превращая его в многообразие. Фактически, требование дифференцируемости отображений χ_i означает, что дифференцируемая структура на пространстве расслоения $\mathbb{E}$ согласована с дифференцируемой структурой, индуцированной отображениями χ_i .

В общем случае размерности базы $\mathbb{M}$ и векторного пространства $\mathbb{V}$ (типичного слоя) могут не совпадать. Если $\dim \mathbb{M} = m$ и $\dim \mathbb{V} = n$, то $\dim \mathbb{E} = m + n$.

В силу условия 1) в определении расслоения многообразие $\pi^{-1}(\mathbb{U}_i)$ является тривиальным расслоением над каждой областью $\mathbb{U}_i$.

Пример 5.1.1. Построим двумерные расслоения $\mathbb{E}(\mathbb{S}^1, \pi, \mathbb{R})$, базой которых является окружность $\mathbb{S}^1$, а типичным слоем – вещественная прямая $\mathbb{R}$, которая рассматривается, как векторное пространство. Отождествим каждую точку окружности с

полярным углом $\mathbb{S}^1 \supset \alpha \in [0, 2\pi)$. Покроем окружность двумя картами (см. рис.5.1): $\mathbb{S}^1 \supset \mathbb{U}_1 = (2\pi - \delta, 2\pi) \cup [0, \pi + \delta)$ и $\mathbb{S}^1 \supset \mathbb{U}_2 = (\pi - \delta, 2\pi) \cup [0, \delta)$, где $0 < \delta < \pi/2$, которые пересекаются по двум областям $\mathbb{U}_1 \cap \mathbb{U}_2 = \mathbb{U}_{12} \cup \mathbb{U}_{21}$, где $\mathbb{U}_{12} = (\pi - \delta, \pi + \delta)$ и $\mathbb{U}_{21} = (2\pi - \delta, 2\pi) \cup [0, \delta)$, при этом координатные функции определяют диффеоморфизмы:

$$\begin{aligned}\varphi_1(\mathbb{U}_1) &= (-\delta, \pi + \delta) \in \mathbb{R}, \\ \varphi_2(\mathbb{U}_2) &= (\pi - \delta, 2\pi + \delta) \in \mathbb{R}.\end{aligned}$$

Определим векторное расслоение $\mathbb{E}$ его локальным вложением в $\mathbb{R}^3$:

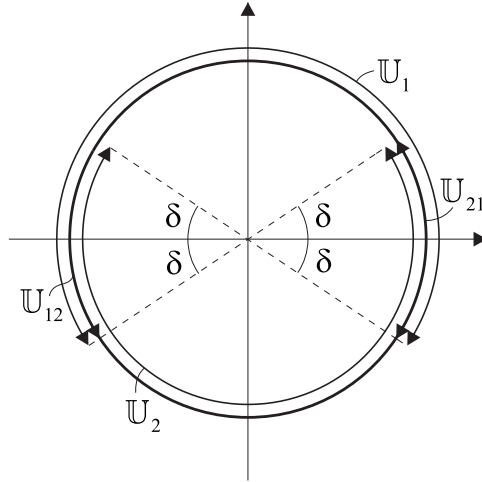


Рис. 5.1: Покрытие окружности двумя картами $\mathbb{S}^1 = \mathbb{U}_1 \cup \mathbb{U}_2$. $\mathbb{U}_{12}$ и $\mathbb{U}_{21}$ – области пересечения карт.

$$\begin{aligned}\pi^{-1}(\mathbb{U}_1) &= \{p = (x, v^1, v^2) \in \mathbb{R}^3 : x \in \varphi_1(\mathbb{U}_1), v^2 \cos(kx/2) = v^1 \sin(kx/2)\}, \\ \pi^{-1}(\mathbb{U}_2) &= \{p = (x, v^1, v^2) \in \mathbb{R}^3 : x \in \varphi_2(\mathbb{U}_2), v^2 \cos(kx/2) = v^1 \sin(kx/2)\},\end{aligned}$$

где $k = 0, 1, 2, \dots$ и v^1, v^2 – декартовы координаты на плоскости $\mathbb{R}^2$. Для большей ясности мы обозначили точки окружности $\alpha \in \mathbb{S}^1$ и их координаты $x = \varphi(\alpha) \in \mathbb{R}$ разными буквами. Словами, при движении точки по окружности $\mathbb{S}^1$ вещественная прямая $\mathbb{R} \subset \mathbb{R}^2$ поворачивается в плоскости $\mathbb{R}^2$ на угол $kx/2$. Если $v \in \mathbb{R}$ – точка на прямой, то ей соответствует точка $\mathbf{v} = (v^1, v^2) = (v \cos(kx/2), v \sin(kx/2)) \in \mathbb{R}^2$ на плоскости. Направляющим вектором прямой является $\mathbf{n} = (\cos(kx/2), \sin(kx/2)) \in \mathbb{R}^2$. Теперь определим отображения:

$$\begin{aligned}\chi_1 : \pi^{-1}(\mathbb{U}_1) \ni p &\mapsto (\alpha, v) \in \mathbb{S}^1 \times \mathbb{R}, \\ \chi_2 : \pi^{-1}(\mathbb{U}_2) \ni p &\mapsto (\alpha, v) \in \mathbb{S}^1 \times \mathbb{R},\end{aligned}$$

где $v := (\mathbf{n}, \mathbf{v}) = v^1 \cos(kx/2) + v^2 \sin(kx/2)$. Для определения гомоморфизмов f_{ij} необходимо учесть, что точки x и $x + 2\pi$ в пересечении $\mathbb{U}_{21}$ соответствуют одной и той же точке окружности, и

$$\cos(k(2\pi + x)/2) = (-1)^k \cos(kx/2), \quad \sin(k(2\pi + x)/2) = (-1)^k \sin(kx/2).$$

Поэтому $f_{12} = \text{id}(\mathbb{R})$ и $f_{21} = (-1)^k \text{id}(\mathbb{R})$. Отсюда следует, что все расслоения для четных k диффеоморфны между собой и диффеоморфны тривиальному расслоению $\mathbb{E} = \mathbb{S}^1 \times \mathbb{R}$, которое получается при $k = 0$. Расслоения при нечетных k также диффеоморфны между собой и диффеоморфны листу Мебиуса, соответствующего $k = 1$. $\square$

При задании расслоения функции перехода должны удовлетворять условиям совместности:

- 1) для всех $x \in \mathbb{U}_i$, $f_{ii} = \text{id}$;
- 2) для всех $x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k \neq \emptyset$, $f_{ij}f_{jk}f_{ki} = \text{id}$.

Эти условия совместности являются, очевидно, необходимыми. Второе свойство означает коммутативность диаграммы

$$\begin{array}{ccc} \mathbb{U}_i & \xrightarrow{f_{ij}} & \mathbb{U}_j \\ & \searrow f_{ik} & \downarrow f_{jk} \\ & & \mathbb{U}_k \end{array}$$

где $f_{ik} = f_{ki}^{-1}$. При задании расслоения с помощью функций перехода необходимо требовать выполнения данных свойств.

Теорема 5.1.1. Пусть $\{\mathbb{U}_i\}$ – открытое координатное покрытие (атлас) дифференцируемого многообразия $\mathbb{M}$, и $\mathbb{V}$ – векторное пространство. Если для всех пар пересекающихся координатных областей $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$ заданы достаточно гладкие отображения $f_{ij} : \mathbb{U}_i \cap \mathbb{U}_j \rightarrow \text{GL}(\mathbb{V})$, которые удовлетворяют условиям совместности 1) и 2), тогда существует векторное расслоение $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ с функциями перехода f_{ij} .

Доказательство. См., например, [55], теорема 3.2. □

Поскольку слоями векторных расслоений являются векторные пространства, то из заданных векторных расслоений с одной и той же базой можно строить новые векторные расслоения, производя со слоями операции, которые допустимы для векторных пространств, поточечно.

Пример 5.1.2 (Прямая сумма $\mathbb{E}_1 \oplus \mathbb{E}_2$ и тензорное произведение $\mathbb{E}_1 \otimes \mathbb{E}_2$ расслоений.). Пусть $\mathbb{E}_1(\mathbb{M}, \pi_1, \mathbb{V}_1)$ и $\mathbb{E}_2(\mathbb{M}, \pi_2, \mathbb{V}_2)$ – два векторных расслоения с одной и той же базой $\mathbb{M}$. Определим прямую сумму векторных расслоений

$$\mathbb{E}_1 \oplus \mathbb{E}_2(\mathbb{M}, \pi, \mathbb{V}_1 \oplus \mathbb{V}_2) := \bigcup_{x \in \mathbb{M}} \pi_1^{-1}(x) \oplus \pi_2^{-1}(x). \quad (5.2)$$

Если точка $p \in \mathbb{E}_1 \oplus \mathbb{E}_2$, то она имеет вид $p = p_1 \oplus p_2$, где $p_1 \in \pi_1^{-1}(x)$, $p_2 \in \pi_2^{-1}(x)$ для некоторого $x \in \mathbb{M}$. Тем самым определена проекция $\pi(p) = x$. Типичным слоем прямой суммы векторных расслоений является векторное пространство $\mathbb{V}_1 \oplus \mathbb{V}_2$. Пусть $\{\mathbb{U}_i\}$ – координатное покрытие $\mathbb{M}$ и $\{f_{ij}^{(1)}\}$ и $\{f_{ij}^{(2)}\}$ – функции перехода для расслоений $\mathbb{E}_1$ и $\mathbb{E}_2$, соответственно. Функции перехода для прямой суммы векторных расслоений имеют вид

$$h_{ij} := \begin{pmatrix} f_{ij}^{(1)} & 0 \\ 0 & f_{ij}^{(2)} \end{pmatrix},$$

которые задают автоморфизм в $\mathbb{V}_1 \oplus \mathbb{V}_2$.

Аналогично определяется тензорное произведение векторных расслоений:

$$\mathbb{E}_1 \otimes \mathbb{E}_2(\mathbb{M}, \pi, \mathbb{V}_1 \otimes \mathbb{V}_2) := \bigcup_{x \in \mathbb{M}} \pi_1^{-1}(x) \otimes \pi_2^{-1}(x). \quad (5.3)$$

Функции перехода для тензорного произведения имеют вид

$$h_{ij} := f_{ij}^{(1)} \otimes f_{ij}^{(2)},$$

которые действуют в тензорном произведении $\mathbb{V}_1 \otimes \mathbb{V}_2$. $\square$

Размерности прямой суммы и тензорного произведения расслоений равны:

$$\begin{aligned} \dim(\mathbb{E}_1 \oplus \mathbb{E}_2) &= \dim \mathbb{M} + \dim \mathbb{V}_1 + \dim \mathbb{V}_2, \\ \dim(\mathbb{E}_1 \otimes \mathbb{E}_2) &= \dim \mathbb{M} + \dim \mathbb{V}_1 \times \dim \mathbb{V}_2. \end{aligned}$$

В заключение данного раздела обсудим касательное расслоение, построенное в разделе 2.6.2.

Теорема 5.1.2. *Касательное расслоение $\mathbb{T}(\mathbb{M})$ является векторным расслоением $\mathbb{E}(\mathbb{M}, \pi, \mathbb{R}^n)$ с базой $\mathbb{M}$, $\dim \mathbb{M} = n$, типичным слоем $\mathbb{R}^n$ и проекцией $\pi : (x, X) \rightarrow x$, где $X \in \mathbb{T}_x(\mathbb{M})$.*

Доказательство. Проверка свойств 1)–3) в определении векторного расслоения. $\square$

Для касательного расслоения размерности базы и типичного слоя совпадают:

$$\dim \mathbb{M} = \dim \mathbb{R}^n = n. \quad \square$$

Касательное расслоение имеет свою специфику. Вообще говоря, в определении векторного расслоения ничего не говорится о том, какое именно линейное преобразование векторного пространства происходит в областях пересечения карт. Говорится лишь о том, что оно возможно. Для касательного расслоения реализован естественный способ гладко сопоставить каждому преобразованию координат автоморфизм касательного пространства. А именно, каждому преобразованию координат соответствует тождественное преобразование касательного пространства, $f_{ij} = \text{id}(\mathbb{T}_x(\mathbb{M}))$ для всех $x \in \mathbb{U}_i \cap \mathbb{U}_j$. Если в касательном пространстве в точке $x \in \mathbb{M}$ (слое над x) выбран координатный (голономный) базис $\{\partial_\alpha\}$, то при преобразовании координат в базе компоненты вектора умножаются на матрицу Якоби, а голономный базис касательного пространства – на обратную матрицу Якоби (2.9). Поэтому сам вектор остается без изменения. Если вектор касательного пространства рассматривается, как набор компонент $\{X^\alpha\}$, считая, что базис векторного пространства фиксирован, то при преобразовании координат компоненты вектора умножаются на матрицу Якоби $\partial_\alpha x^{\alpha'}$ справа. Тогда функциями перехода являются матрицы Якоби. Обе точки зрения на функции перехода допустимы. Это зависит от того, как рассматривать базис векторного пространства $\mathbb{V}$. В общем случае базис касательного пространства e_a , $a = 1, \dots, n$, можно зафиксировать произвольным образом (неголономный базис). Тогда мы считаем, что при преобразовании координат базис и координаты векторов относительно этого базиса, $X = X^a e_a$, не меняются. Такой подход часто бывает удобнее и соответствует использованию репера $e_a = e^a_\alpha \partial_\alpha$, который образует базис касательного пространства.

Дифференцируемое сечение векторного расслоения $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$, которое всюду отлочно от нуля, существует не всегда. Существование таких сечений отражает определенные топологические свойства базы $\mathbb{M}$. Множество сечений векторного расслоения $\mathbb{E}$ будем обозначать $\mathcal{V}_{\mathbb{E}}(\mathbb{M})$.

Понятие (псевдо)риманова многообразия естественным образом обобщается на произвольные векторные расслоения.

Определение. Если в каждой точке $x \in \mathbb{M}$ задана невырожденная симметричная билинейная форма g на слое $\pi^{-1}(x)$, и значение $g(V_1, V_2)$ является достаточно гладкой функцией на $\mathbb{M}$ для произвольных достаточно гладких сечений $V_1, V_2 \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$, то $\mathbb{E}$ называется (псевдо)римановым векторным расслоением. $\square$

Пример 5.1.3. Тензорное расслоение на римановом многообразии является римановым векторным расслоением. Например, для тензорных полей $X = \{X_{\alpha}^{\beta}\}$ и $\{Y_{\alpha}^{\beta}\}$ типа $(1, 1)$ квадратичная форма задается римановой метрикой:

$$g(X, Y) = g^{\alpha\beta} X_{\alpha}^{\gamma} Y_{\beta}^{\delta} g_{\gamma\delta}. \quad \square$$

5.2 Связность на векторном расслоении

В настоящем разделе мы определим связность на векторном расслоении в инвариантной форме. Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = m$, и векторное пространство $\mathbb{V}$, $\dim \mathbb{V} = n$. Пусть $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ – векторное расслоение. Сечениями этого расслоения являются векторные поля $V(x) \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$. На множестве сечений векторного расслоения поточечно вводится сложение и умножение на гладкие функции. Таким образом, множество сечений $\mathcal{V}_{\mathbb{E}}(\mathbb{M})$ является $\mathcal{C}^{\infty}(\mathbb{M})$ -модулем.

Все дальнейшие конструкции мы будем иллюстрировать записью соответствующих выражений в компонентах, поскольку это, во-первых, наглядно и, во-вторых, необходимо при проведении вычислений. Выберем в каждом слое $\pi^{-1}(x)$, $x \in \mathbb{U} \subset \mathbb{M}$, базис $e_a(x) \in \mathcal{V}_{\mathbb{E}}(\mathbb{U})$, $a = 1, \dots, n$. Этот базис называется *репером*. Для каждой точки x существует окрестность $\mathbb{U} \subset \mathbb{M}$, в которой репер задается достаточно гладкими функциями от x . Вообще говоря, такой репер глобально существует не для всех многообразий, и это зависит от их топологических свойств. В компонентах сечение векторного расслоения имеет вид $V = V^a(x)e_a$.

Кокасательное расслоение $\mathbb{T}^*(\mathbb{M})$ имеет ту же базу, что и $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$, поэтому можно построить тензорное произведение расслоений $\mathbb{T}^*(\mathbb{M}) \otimes \mathbb{E}$. Пусть $\mathcal{V}_{\mathbb{T}^* \otimes \mathbb{E}}(\mathbb{M})$ – множество сечений тензорного произведения расслоений. В компонентах сечение этого расслоения задается векторным полем с дополнительным ковариантным индексом $dx^{\alpha} \otimes W_{\alpha}^a(x)e_a$, где dx^{α} , $\alpha = 1, \dots, m$, – координатный базис 1-форм и $W_{\alpha}^a(x)$ – компоненты сечения.

Определение. Связностью ∇ на векторном расслоении $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ называется отображение

$$\nabla : \mathcal{V}_{\mathbb{E}}(\mathbb{M}) \rightarrow \mathcal{V}_{\mathbb{T}^* \otimes \mathbb{E}}(\mathbb{M}), \quad (5.4)$$

которое удовлетворяет двум условиям:

1) для любых двух сечений $V_1, V_2 \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$

$$\nabla(V_1 + V_2) = \nabla V_1 + \nabla V_2;$$

2) для произвольного сечения $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$ и произвольной функции $f \in \mathcal{C}^{\infty}(\mathbb{M})$

$$\nabla(fV) = df \otimes V + f \nabla V. \quad \square \quad (5.5)$$

Если в определении связности положить $f = \text{const}$, то из условий 1) и 2) следует линейность отображения ∇ .

Определение. Сечение $\nabla V \in \Lambda_1(\mathbb{M}) \otimes \mathbb{V}$ представляет собой 1-форму на $\mathbb{M}$ со значениями в векторном пространстве $\mathbb{V}$. Рассмотрим касательное векторное поле к базе $X \in \mathcal{X}(\mathbb{M})$. Тогда определено значение 1-формы ∇V на векторном поле X

$$\nabla_X V := \nabla V(X), \quad (5.6)$$

которое называется *ковариантной производной* векторного поля $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$ вдоль касательного векторного поля $X \in \mathcal{X}(\mathbb{M})$. $\square$

Ковариантная производная вдоль векторного поля ∇_X является отображением $\mathcal{V}_{\mathbb{E}}(\mathbb{M}) \rightarrow \mathcal{V}_{\mathbb{E}}(\mathbb{M})$. Пусть $X, Y \in \mathcal{X}(\mathbb{M})$ – два произвольных векторных поля на $\mathbb{M}$, и $f \in \mathcal{C}^\infty(\mathbb{M})$ – гладкая функция. Тогда ковариантная производная имеет следующие свойства:

$$\begin{aligned} \nabla_{X+Y} V &= \nabla_X V + \nabla_Y V, \\ \nabla_{fX} V &= f \nabla_X V, \\ \nabla_X (V_1 + V_2) &= \nabla_X V_1 + \nabla_X V_2, \\ \nabla_X (fV) &= (Xf)V + f \nabla_X V. \end{aligned}$$

Построим выражение для ковариантной производной в компонентах. Запишем отображение (5.4) для векторов репера:

$$\nabla e_a = dx^\alpha \otimes \omega_{\alpha a}{}^b e_b, \quad (5.7)$$

где $\omega_{\alpha a}{}^b(x)$ – некоторые функции на координатной окрестности $\mathbb{U}$. Введем 1-формы

$$\omega_a{}^b := dx^\alpha \omega_{\alpha a}{}^b. \quad (5.8)$$

Тогда соотношение (5.7) можно переписать в виде

$$\nabla e_a = \omega_a{}^b \otimes e_b. \quad (5.9)$$

Рассмотрим, как меняются компоненты $\omega_{\alpha a}{}^b$ при локальном вращении репера

$$e'_a := S_a{}^b e_b, \quad (5.10)$$

где $S_a{}^b(x) \in \mathbb{GL}(n, \mathbb{R})$ – некоторая матрица преобразования, элементы которой могут зависеть от точки $x \in \mathbb{U}$. В новом базисе справедливо равенство

$$\nabla e'_a = \omega'_a{}^b \otimes e'_b.$$

С другой стороны, из определения связности (5.5) следует, что

$$\nabla(S_a{}^b e_b) = dS_a{}^b \otimes e_b + S_a{}^b \nabla e_b.$$

Сравнивая эти выражения, получаем правило преобразования:

$$\omega'_a{}^b = S_a{}^c \omega_c{}^d S^{-1}{}_d{}^b + dS_a{}^c S^{-1}{}_c{}^b, \quad (5.11)$$

или, в компонентах,

$$\omega'_{\alpha a}{}^b = S_a{}^c \omega_{\alpha c}{}^d S^{-1}{}_d{}^b + \partial_\alpha S_a{}^c S^{-1}{}_c{}^b. \quad (5.12)$$

Замечание. Это – одна из самых важных формул дифференциальной геометрии, которая широко используется в калибровочных моделях математической физики. Подчеркнем, что преобразование 1-форм ω_a^b содержит неоднородное слагаемое и поэтому не является тензорным. Следовательно, компоненты ω_{aa}^b не определяют никакого тензора. Они определяют связность. $\square$

Теорема 5.2.1. На любом векторном расслоении $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ существует связность.

Доказательство. Доказательство проводится путем явного построения связности с использованием разложения единицы [49], глава 4, теорема 1.1. $\square$

Пусть на многообразии $\mathbb{M}$ задан атлас $\{\mathbb{U}_i\}$. Тогда связность в каждой карте задается компонентами $\omega_{\alpha a}^b$, которые в областях пересечения карт связаны преобразованием (5.12). Верно также и обратное утверждение. Набор компонент $\omega_{\alpha a}^b$, связанных преобразованием (5.12) в областях пересечения карт, однозначно определяет связность на векторном расслоении.

Определение. Компоненты $dx^\alpha \omega_{\alpha a}^b$ называются *локальной формой связности*.

Определив действие связности на репер, можно выписать явное выражение для действия связности на произвольное векторное поле $V^a e_a \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$:

$$\nabla V = (dV^a + V^b \omega_b^a) \otimes e_a = dx^\alpha (\partial_\alpha V^a + V^b \omega_{\alpha b}^a) \otimes e_a, \quad (5.13)$$

где мы воспользовались определением (5.5). Это выражение называется *ковариантной производной* векторного поля. Выражение в скобках,

$$\nabla_\alpha V^a = \partial_\alpha V^a + V^b \omega_{\alpha b}^a, \quad (5.14)$$

называется *ковариантной производной компонент* векторного поля. $\square$

Замечательным свойством ковариантной производной является то, что при локальном изменении базиса (5.10) ковариантная производная меняется по тензорному закону. Действительно, при изменении базиса (репера) (5.10) компоненты вектора $V = V^a e_a = V'^a e'_a$ преобразуются по векторному представлению группы $\mathbb{GL}(n, \mathbb{R}^n)$:

$$V'^a = V^b S_b^{-1a}, \quad S_b^a(x) \in \mathbb{GL}(n, \mathbb{R}^n). \quad (5.15)$$

Используя преобразование связности (5.12), нетрудно проверить справедливость следующей формулы преобразования ковариантной производной

$$\nabla'_\alpha V'^a = \partial_\alpha V'^a + V'^b \omega'_{\alpha b}{}^a = (\nabla_\alpha V^b) S_b^{-1a},$$

где ковариантная производная со штрихом ∇'_α берется с локальной формой связности $\omega'_{\alpha b}{}^a$. Тем самым ковариантная производная компонент векторного поля преобразуется по векторному представлению группы $\mathbb{GL}(n, \mathbb{R}^n)$.

Ковариантная производная от компонент векторного поля V вдоль касательного векторного поля X (5.6) в компонентах имеет вид

$$\nabla_X V^a = X^\alpha \nabla_\alpha V^a.$$

Рассмотрим некоторые свойства связности. Легко доказывается

Предложение 5.2.1. Пусть на векторном расслоении $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ задана связность ∇ и зафиксирована точка базы $x \in \mathbb{M}$. Тогда в окрестности точки x существует такой репер, что компоненты локальной формы связности в этой точке обращаются в нуль $\omega_{\alpha a}{}^b(x) = 0$.

Подчеркнем, что в данном предложении речь идет о фиксированной точке многообразия, а не об окрестности.

Продолжим изучение локальной формы связности. Перепишем соотношение (5.11) в виде

$$\omega' S = S\omega + dS, \quad (5.16)$$

где, для краткости, мы опустили векторные индексы, предполагая всюду суммирование “с десяти до четырех” (имеется ввиду расположение чисел на циферблате часов). Внешняя производная от этого равенства имеет вид

$$d\omega' S - \omega' \wedge dS = dS \wedge \omega + Sd\omega.$$

Подстановка выражения для dS из (5.16) приводит к равенству

$$d\omega' - \omega' \wedge \omega' = S(d\omega - \omega \wedge \omega)S^{-1}, \quad (5.17)$$

которое приводит к следующему важному понятию.

Определение. 2-форма на $\mathbb{M}$

$$R_a{}^b := d\omega_a{}^b - \omega_a{}^c \wedge \omega_c{}^b \quad (5.18)$$

называется *локальной формой кривизны* связности ∇ .

В компонентах локальная форма кривизны имеет вид:

$$R_a{}^b = \frac{1}{2} dx^\alpha \wedge dx^\beta R_{\alpha\beta a}{}^b,$$

где

$$R_{\alpha\beta a}{}^b := \partial_\alpha \omega_{\beta a}{}^b - \partial_\beta \omega_{\alpha a}{}^b - \omega_{\alpha a}{}^c \omega_{\beta c}{}^b + \omega_{\beta a}{}^c \omega_{\alpha c}{}^b. \quad \square \quad (5.19)$$

Локальная форма кривизны – это 2-форма на многообразии $\mathbb{M}$ со значениями в тензорном произведении $\mathbb{V}^* \otimes \mathbb{V}$, где $\mathbb{V}^*$ – векторное пространство, сопряженное к $\mathbb{V}$.

Формула (5.17) показывает, что при локальном преобразовании репера форма кривизны преобразуется по тензорному закону, т.е. однородно, несмотря на то, что преобразование формы связности (5.11) содержит неоднородное слагаемое.

Пусть на многообразии $\mathbb{M}$ задано два векторных поля $X, Y \in \mathcal{X}(\mathbb{M})$. Тогда значение формы кривизны на этих полях $R(X, Y)_a{}^b$ определяет линейное преобразование из $\mathcal{V}_{\mathbb{E}}(\mathbb{M})$ в $\mathcal{V}_{\mathbb{E}}(\mathbb{M})$ (эндоморфизм). Если $V = V^a e_a \in \mathcal{V}_{\mathbb{E}}(\mathbb{E})$, то

$$R(X, Y) : V^a \rightarrow V^b R(X, Y)_b{}^a.$$

Это отображение обладает очевидными свойствами:

- 1) $R(X, Y) = -R(Y, X)$;
- 2) $R(fX, Y) = fR(X, Y)$;
- 3) $R(X, Y)(fV) = fR(X, Y)V$;

где $f \in \mathcal{C}^\infty(\mathbb{M})$.

Предложение 5.2.2. Пусть на векторном расслоении $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ задана связность ∇ с локальной формой кривизны R . Тогда для любых векторных полей $X, Y \in \mathcal{X}(\mathbb{M})$ справедлива формула

$$R(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X, Y]}. \quad (5.20)$$

Доказательство. Прямая проверка в компонентах. $\square$

Теорема 5.2.2 (Тождества Бианки). Форма кривизны R удовлетворяет тождествам Бианки:

$$dR_a{}^b + R_a{}^c \wedge \omega_c{}^b - \omega_a{}^c \wedge R_c{}^b = 0. \quad (5.21)$$

Доказательство. Внешняя производная от определения кривизны (5.18) равна

$$\begin{aligned} dR &= -d\omega \wedge \omega + \omega \wedge d\omega = \\ &= -(R + \omega \wedge \omega) \wedge \omega + \omega \wedge (R + \omega \wedge \omega) = \\ &= -R \wedge \omega + \omega \wedge R. \end{aligned} \quad \square$$

В компонентах тождества Бианки имеют вид:

$$dx^\alpha \wedge dx^\beta \wedge dx^\gamma \partial_\alpha R_{\beta\gamma a}{}^b = dx^\alpha \wedge dx^\beta \wedge dx^\gamma (\omega_{\alpha a}{}^c R_{\beta\gamma c}{}^b - R_{\alpha\beta a}{}^c \omega_{\gamma c}{}^b)$$

или

$$\nabla_\alpha R_{\beta\gamma a}{}^b + \nabla_\beta R_{\gamma\alpha a}{}^b + \nabla_\gamma R_{\alpha\beta a}{}^b = 0, \quad (5.22)$$

где

$$\nabla_\alpha R_{\beta\gamma a}{}^b := \partial_\alpha R_{\beta\gamma a}{}^b + R_{\beta\gamma a}{}^c \omega_{\alpha c}{}^b - \omega_{\alpha a}{}^c R_{\beta\gamma c}{}^b.$$

– ковариантная производная от тензора (2-формы) кривизны.

Замечание. Построенная конструкция имеет важное приложение в калибровочных моделях теории поля. Если матрицы $S_a{}^b(x) \in \mathbb{GL}(n, \mathbb{R})$ в (5.15) образуют представление ρ некоторой полупростой группы Ли $\mathbb{G}$, то эта группа называется калибровочной. В этом случае локальная форма связности $\omega_{\alpha a}{}^b$ называется *полем Янга–Миллса*, а локальная форма кривизны $R_{\alpha\beta a}{}^b$ – *напряженностью* поля Янга–Миллса. Векторные поля $V = V^a e_a$ называются *полями материи*, преобразующимися по представлению ρ калибровочной группы. Эти поля называются скалярными полями, т.к. представляют собой набор функций на многообразии, не меняющихся при преобразовании координат, $V^a \in \mathcal{C}^\infty(\mathbb{M})$ для всех значений индекса a . Конечно, необходимо также постулировать соответствующие уравнения движения. $\square$

Определение. Рассмотрим сечение векторного расслоения $\mathbb{E}$, т.е. векторное поле $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$. Сечение V называется *параллельным или горизонтальным* относительно связности ∇ , если ковариантная производная от него равна нулю

$$\nabla V = 0. \quad \square \quad (5.23)$$

Нулевое сечение, которое каждой точке базы сопоставляет нулевой вектор, очевидно, всегда является параллельным. В то же время отличное от нуля параллельное сечение в общем случае может отсутствовать. Рассмотрим этот вопрос подробно. Разложим векторное поле по локальному базису: $V = V^a e_a$. Тогда уравнение (5.23) эквивалентно пфаффовым уравнениям на компоненты:

$$dV^a + V^b \omega_b{}^a = 0. \quad (5.24)$$

Для анализа этой системы уравнений введем обозначение для левой части

$$A^a := dV^a + V^b \omega_b^a.$$

Внешняя производная от этой 1-формы равна

$$dA^a = A^b \wedge \omega_b^a + V^b R_b^a.$$

Из этого выражения и теоремы Фробениуса 2.11.4, записанной в терминах дифференциальных форм, следует, что, если кривизна связности равна нулю, то пфафова система уравнений (5.23) является вполне интегрируемой. В этом случае существует n линейно независимых параллельных сечений векторного расслоения. Мы видим, что для существования отличных от нуля локальных решений системы уравнений (5.24) локальная форма кривизны данной связности должна обращаться в нуль. Другими словами, равенство нулю формы кривизны является необходимым и достаточным условием разрешимости системы уравнений в частных производных (5.24).

Определение. Пусть на многообразии $\mathbb{M}$ задана дифференцируемая кривая $\gamma = \{x^\alpha(t)\}$, $\alpha = 1, \dots, m$. Обозначим касательный вектор к кривой (вектор скорости) через u . Тогда сечение векторного расслоения $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$ называется *параллельным вдоль кривой γ* , если ковариантная производная от него вдоль u равна нулю

$$\nabla_u V = 0. \quad \square \quad (5.25)$$

Касательный вектор к кривой (вектор скорости) имеет вид $u := \dot{x}^\alpha \partial_\alpha$. В компонентах условие параллельности сечения вдоль кривой γ сводится к системе уравнений на компоненты сечения

$$\dot{V}^a + \dot{x}^\alpha V^b \omega_{\alpha b}^a = 0. \quad (5.26)$$

Поскольку это – система обыкновенных дифференциальных уравнений, то локально у нее существует единственное решение при заданных начальных данных. Таким образом, если задан вектор V_p в некоторой точке p на кривой γ , то он однозначно определяет векторное поле вдоль кривой γ , которое называется *параллельным переносом* вектора V_p вдоль кривой γ . Очевидно, что параллельный перенос задает изоморфизм слоев векторного расслоения $\mathbb{E}$ в различных точках кривой γ .

Связность ∇ на векторном расслоении $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ индуцирует связность на сопряженном векторном расслоении $\mathbb{E}^*(\mathbb{M}, \pi, \mathbb{V}^*)$, которую мы также обозначим ∇ . Это делается следующим образом. Пусть заданы сечения $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$ и $V^* \in \mathcal{V}_{\mathbb{E}^*}(\mathbb{M})$. Тогда каждое сечение $V^*(x)$ задает поточечное линейное отображение

$$V^*(x) : \mathcal{V}_{\mathbb{E}}(\mathbb{M}) \ni V(x) \mapsto (V^*, V) \in \mathcal{C}^\infty(\mathbb{M}).$$

Определим индуцированную связность на сопряженном расслоении $\mathbb{E}^*$ следующей формулой

$$d(V^*, V) = (V^*, \nabla V) + (\nabla V^*, V), \quad \forall V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M}). \quad (5.27)$$

Найдем выражение для индуцированной связности в компонентах. Пусть e^a – базис сопряженного векторного пространства $\mathbb{V}^*$: $(e^a, e_b) = \delta_b^a$. Этот базис называется *корепером*. Тогда подстановка корепера e^a и репера e_b в определение (5.27) приводит к равенству

$$\nabla e^a = -e^b \otimes \omega_b^a. \quad (5.28)$$

Отсюда следует выражение для ковариантной производной сечения $V^* = e^a V_a$ сопряженного расслоения

$$\nabla V^* = e^a \otimes (dV_a - \omega_a^b V_b). \quad (5.29)$$

Эта ковариантная производная отличается знаком у слагаемого со связностью от соответствующего выражения для сечения векторного расслоения (5.13).

Определение. Пусть на векторных расслоениях $\mathbb{E}_1(\mathbb{M}, \pi_1, \mathbb{V}_1)$ и $\mathbb{E}_2(\mathbb{M}, \pi_2, \mathbb{V}_2)$ с одинаковой базой $\mathbb{M}$ независимо заданы связности ∇_1 и ∇_2 . Пусть заданы сечения $V_1 \in \mathcal{V}_{\mathbb{E}_1}(\mathbb{M})$ и $V_2 \in \mathcal{V}_{\mathbb{E}_2}(\mathbb{M})$. Тогда на прямой сумме $\mathbb{E}_1 \oplus \mathbb{E}_2$ и тензорном произведении $\mathbb{E}_1 \otimes \mathbb{E}_2$ расслоений определена связность

$$\begin{aligned} \nabla(V_1 \oplus V_2) &= \nabla_1 V_1 \oplus \nabla_2 V_2, \\ \nabla(V_1 \otimes V_2) &= \nabla_1 V_1 \otimes V_2 + V_1 \otimes \nabla_2 V_2, \end{aligned} \quad (5.30)$$

которая называется *индуцированной*. □

5.3 Аффинная связность

Касательное расслоение $\mathbb{T}(\mathbb{M})$, $\dim \mathbb{M} = n$, является векторным расслоением, типичным слоем которого является векторное пространство $\mathbb{R}^n$. Поэтому определение связности на векторном расслоении без изменений переносится на касательные расслоения. Связность на касательном расслоении называется *аффинной связностью*. Если в касательном пространстве и сопряженном к нему пространстве 1-форм выбрать координатный базис ∂_α и dx^α , то формулы для компонент связности (5.7) и (5.28) определяют аффинную связность

$$\begin{aligned} \nabla(\partial_\alpha) &= dx^\beta \otimes \Gamma_{\beta\alpha}^\gamma \partial_\gamma, \\ \nabla(dx^\alpha) &= -dx^\gamma \otimes dx^\beta \Gamma_{\beta\gamma}^\alpha. \end{aligned} \quad (5.31)$$

У компонент аффинной связности $\Gamma_{\alpha\beta}^\gamma$ все индексы являются координатными, и для них принято обозначение Γ вместо ω . Формула преобразования компонент аффинной связности при преобразовании координат получается из формулы преобразования компонент связности на векторном расслоении (5.12) после замены матрицы $S_b^a(x)$ на матрицу Якоби $\partial_\alpha x^{\alpha'}$. При этом необходимо учесть, что при преобразовании координат преобразуется также базис 1-форм, что приводит к дополнительному общему множителю:

$$\Gamma_{\beta'\gamma'}^{\alpha'} = (\partial_{\beta'} x^\beta \partial_{\gamma'} x^\gamma \Gamma_{\beta\gamma}^\alpha + \partial_{\beta'\gamma'}^2 x^\alpha) \partial_\alpha x^{\alpha'}. \quad (5.32)$$

Определение. Функции $\Gamma_{\alpha\beta}^\gamma$, преобразующиеся по правилу (5.32) при преобразовании координат, называются *компонентами локальной формы аффинной связности* или, короче, аффинной связностью на $\mathbb{M}$. □

Замечание. Аффинная связность, так же как и связность на произвольном векторном расслоении, определенная в предыдущем разделе, не имеет ничего общего с метрикой. Ее можно задать независимо от метрики, которая может вообще отсутствовать на многообразии. □

Аффинная связность ∇ осуществляет отображение

$$\nabla : \mathcal{X}(\mathbb{M}) \times \mathcal{X}(\mathbb{M}) \ni X, Y \mapsto \nabla_X Y \in \mathcal{X}(\mathbb{M}), \quad (5.33)$$

где $\nabla_X Y$ – ковариантная производная от векторного поля Y вдоль векторного поля X . Как и в разделе 5.2, можно убедиться, что это отображение обладает следующими свойствами:

$$\begin{aligned}\nabla_{X+Y} Z &= \nabla_X Z + \nabla_Y Z, \\ \nabla_{fX} Y &= f \nabla_X Y, \\ \nabla_X (Y + Z) &= \nabla_X Y + \nabla_X Z, \\ \nabla_X (fY) &= (Xf)Y + f \nabla_X Y,\end{aligned}\tag{5.34}$$

где $f \in C^\infty(\mathbb{M})$ и $X, Y, Z \in \mathcal{X}(\mathbb{M})$. Верно также обратное утверждение: отображение ∇ со свойствами (5.34) определяет аффинную связность на касательном расслоении. Иногда его принимают в качестве определения аффинной связности.

Связность, определенная на касательном и кокасательном расслоении, индуцирует связность на произвольных тензорных расслоениях с помощью правил (5.30). В компонентах ковариантная производная от тензорных полей будет выписана в следующей главе. Индуцированная ковариантная производная линейна,

$$\nabla(aY \oplus bZ) = a\nabla Y \oplus b\nabla Z, \quad a, b \in \mathbb{R},$$

где $Y, Z \in \mathcal{T}(\mathbb{M})$ – произвольные тензорные поля. Справедливо правило Лейбница

$$\nabla(Y \otimes Z) = (\nabla Y) \otimes Z + Y \otimes (\nabla Z).\tag{5.35}$$

Нетрудно проверить, что ковариантная производная перестановочна с каждым свертыванием: $\nabla C = C\nabla$, где C – оператор свертки. Кроме того, ковариантная производная вдоль векторного поля ∇_X сохраняет тип тензорных полей $\nabla_X \mathcal{T}_s^r(\mathbb{M}) \subset \mathcal{T}_s^r(\mathbb{M})$. Таким образом, ковариантная производная вдоль векторного поля ∇_X является дифференцированием в тензорной алгебре, которое было определено в разделе 2.14. В силу теоремы 2.14.1 ковариантная производная вдоль векторного поля ∇_X отличается от производной Ли L_X на дифференцирование, порожденное тензорным полем типа $(1, 1)$.

Тензор кривизны (5.19) для аффинной связности (5.31) в компонентах принимает вид

$$R_{\alpha\beta\gamma}{}^\delta = \partial_\alpha \Gamma_{\beta\gamma}{}^\delta - \partial_\beta \Gamma_{\alpha\gamma}{}^\delta - \Gamma_{\alpha\gamma}{}^\epsilon \Gamma_{\beta\epsilon}{}^\delta + \Gamma_{\beta\gamma}{}^\epsilon \Gamma_{\alpha\epsilon}{}^\delta.\tag{5.36}$$

Рассмотрим тензор кривизны (5.36), как отображение касательных пространств

$$\mathcal{X}(\mathbb{M}) \times \mathcal{X}(\mathbb{M}) \times \mathcal{X}(\mathbb{M}) \rightarrow \mathcal{X}(\mathbb{M})$$

и обозначим

$$R(X, Y, Z) := X^\alpha Y^\beta Z^\gamma R_{\alpha\beta\gamma}{}^\delta \partial_\delta.$$

Тогда нетрудно проверить справедливость следующего тождества:

$$R(X, Y, Z) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z,$$

которое является частным случаем предложения 5.2.2. Это соотношение иногда принимают в качестве определения кривизны R аффинной связности ∇ .

Определение. 2-форма со значениями в касательном пространстве,

$$T = \frac{1}{2} dx^\alpha \wedge dx^\beta T_{\alpha\beta}{}^\gamma \partial_\gamma,$$

определенная равенством:

$$T(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y], \quad \forall X, Y \in \mathcal{X}(\mathbb{M}).\tag{5.37}$$

называется *тензором кручения*. □

Компоненты тензора кручения равны антисимметричной части компонент аффинной связности:

$$T_{\alpha\beta}{}^\gamma = \Gamma_{\alpha\beta}{}^\gamma - \Gamma_{\beta\alpha}{}^\gamma. \quad (5.38)$$

Поскольку при преобразовании координат неоднородное слагаемое в (5.32) симметрично по индексам β', γ' , то компоненты кручения действительно преобразуются по тензорному закону. Таким образом, аффинная связность определяет два тензора: тензор кривизны и тензор кручения.

5.4 Связность на расслоении реперов

Существует тесная взаимосвязь между касательным расслоением и расслоением реперов. Начнем с конструктивного построения расслоения реперов. Предположим, что задано дифференцируемое многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$.

Определение. *Репером* в точке $x \in \mathbb{M}$ называется упорядоченный набор объектов $p = (x, e_1, \dots, e_n) = \{x, e_a\}$, $a = 1, \dots, n$, где $\{e_a\}$ – набор линейно независимых касательных векторов в точке x . Множество всех реперов на $\mathbb{M}$ обозначим через $\mathbb{L}(\mathbb{M})$. Ниже мы построим дифференцируемую структуру на $\mathbb{L}(\mathbb{M})$, тем самым превратив множество всех реперов в дифференцируемое многообразие размерности $\dim \mathbb{L}(\mathbb{M}) = n + n^2$ такое, что естественная проекция

$$\pi(x, e_1, \dots, e_n) := x, \quad (5.39)$$

является гладким отображением $\mathbb{L}(\mathbb{M}) \rightarrow \mathbb{M}$. Тогда $\mathbb{L}(\mathbb{M})$ называется *расслоением реперов* на $\mathbb{M}$. $\square$

Замечание. Забегая вперед, скажем, что расслоение реперов представляет собой главное расслоение $\mathbb{L}(\mathbb{M}) = \mathbb{P}(\mathbb{M}, \pi, \mathbb{GL}(n, \mathbb{R}))$ с базой $\mathbb{M}$, проекцией π и структурной группой $\mathbb{GL}(n, \mathbb{R})$, с которым ассоциировано касательное расслоение $\mathbb{T}(\mathbb{M})$. $\square$

Процедура построения дифференцируемой структуры на $\mathbb{L}(\mathbb{M})$ аналогична построению дифференцируемой структуры на тензорных расслоениях. Пусть $\mathbb{U} \subset \mathbb{M}$ – некоторая карта с координатами x^α , $\alpha = 1, \dots, n$. Координатный базис касательного расслоения ∂_α образует координатный репер в каждой точке. Тогда произвольный репер можно разложить по координатному базису:

$$e_a = e^\alpha{}_a \partial_\alpha. \quad (5.40)$$

По определению, совокупность чисел $e^\alpha{}_a$ образует невырожденную $n \times n$ матрицу, т.е. представляет собой элемент группы $\mathbb{GL}(n, \mathbb{R})$ в каждой точке $x \in \mathbb{M}$. Поэтому можно определить отображение $\chi : \pi^{-1}(\mathbb{U}) \rightarrow \mathbb{U} \times \mathbb{GL}(n, \mathbb{R})$ такое, что для каждой точки $(x, e_1, \dots, e_n) \in \mathbb{L}(\mathbb{M})$ и $x \in \mathbb{U}$, $e^\alpha{}_a \in \mathbb{GL}(n, \mathbb{R})$ имеем равенство

$$\chi(x, e_1, \dots, e_n) = (x, e^\alpha{}_a),$$

где репер e_a задан равенством (5.40). Очевидно, что отображение χ является взаимно однозначным.

Теперь выберем атлас $\{\mathbb{U}_i\}$ на $\mathbb{M}$ с построенными выше отображениями $\{\chi_i\}$. Прообразы всех открытых подмножеств топологических произведений $\mathbb{U}_i \times \mathbb{GL}(n, \mathbb{R})$

для отображений χ_i образует базу топологии на $\mathbb{L}(\mathbb{M})$. При этом все отображения χ_i будут гомеоморфизмами.

Благодаря отображению χ_i , прообраз проекции $\pi^{-1}(\mathbb{U}_i)$ становится координатной окрестностью расслоения реперов $\mathbb{L}(\mathbb{M})$. Пусть в пересекающихся областях $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$ заданы координаты x^α и $x^{\alpha'}$, соответственно. Тогда координатами репера $(x, e_1, \dots, e_n)$ в пересечении $\mathbb{U}_i \cap \mathbb{U}_j$ будут наборы из $n + n^2$ чисел $\{x^\alpha, e^\beta_a\}$ и $\{x^{\alpha'}, e^{\beta'}_a\}$, которые связаны преобразованием

$$\begin{aligned} x^{\alpha'} &= x^{\alpha'}(x), \\ e^{\alpha'}_a &= e^\alpha_a \partial_\alpha x^{\alpha'}. \end{aligned} \quad (5.41)$$

Если преобразование координат на $\mathbb{M}$ класса $\mathcal{C}^\infty$, то мы получаем бесконечно дифференцируемую структуру на $\mathbb{L}(\mathbb{M})$. Таким образом расслоение реперов становится дифференцируемым $(n + n^2)$ -мерным многообразием класса $\mathcal{C}^\infty$. При этом естественная проекция $\pi : \mathbb{L}(\mathbb{M}) \rightarrow \mathbb{M}$ является гладким сюръективным отображением.

Сечение расслоения реперов $\mathbb{L}(\mathbb{M})$ мы будем обозначать $e_a(x)$. Его также можно разложить по координатному базису $e_a(x) = e^\alpha_a(x) \partial_\alpha$. Мы предполагаем, что все компоненты этого разложения являются гладкими функциями: $e^\alpha_a(x) \in \mathcal{C}^\infty(\mathbb{U})$. В разделе 5.2 под репером понималось именно сечение расслоения реперов.

Замечание. В формуле (5.40) компоненты разложения e^α_a рассматриваются, как координаты (числа) на пространстве главного расслоения $\mathbb{L}(\mathbb{M})$, а не как функции на базе $e^\alpha_a(x)$. Формулы (5.41) – это также преобразование координат на $\mathbb{L}(\mathbb{M})$. В физической литературе сечение расслоения реперов $e_a(x)$ принято называть просто репером. Для краткости мы также будем употреблять термин репер для обозначения сечения расслоения реперов там, где это не вызывает двусмысленности. $\square$

Отображение χ задает локально структуру прямого произведения на расслоении реперов $\mathbb{L}(\mathbb{M})$. То есть прообраз $\pi^{-1}(\mathbb{U})$ диффеоморфен прямому произведению $\mathbb{U} \times \mathbb{GL}(n, \mathbb{R})$.

Сужение отображения χ на каждый слой $\chi_x : \pi^{-1}(x) \rightarrow \mathbb{GL}(n, \mathbb{R})$ является гомоморфизмом групп. Если $x \in \mathbb{U}_i \cap \mathbb{U}_j$ тогда отображения $f_{ij}(x) = \chi_{j,x} \circ \chi_{i,x}^{-1}$ (функции перехода) представляют собой автоморфизмы группы $\mathbb{GL}(n, \mathbb{R})$. Из (5.41) следует, что преобразование $\chi_{j,x} \circ \chi_{i,x}^{-1}$ является умножением справа в группе $\mathbb{GL}(n, \mathbb{R})$ на матрицу Якоби $J_{ij}^{\alpha'} := \partial x^{\alpha'}_{(j)} / \partial x^\alpha_{(i)}$. Это значит, что набор матриц Якоби J_{ij} представляет собой совокупность функций склейки на расслоении реперов.

Замечание. Типичным слоем расслоения реперов является группа $\mathbb{GL}(n, \mathbb{R})$, на которой нет структуры векторного пространства. Поэтому расслоение реперов не является векторным расслоением.

Поскольку каждый слой $\pi^{-1}(x)$ диффеоморфен и гомоморфен $\mathbb{GL}(n, \mathbb{R})$, то он изоморфен группе Ли: $\pi^{-1}(x) \simeq \mathbb{GL}(n, \mathbb{R})$. $\square$

Структурная группа $\mathbb{GL}(n, \mathbb{R})$ естественным образом действует на расслоении реперов и представляет собой группу преобразований $\mathbb{L}(\mathbb{M})$. Пусть $S_a^b \in \mathbb{GL}(n, \mathbb{R})$. Тогда правое действие r_S структурной группы на $\mathbb{L}(\mathbb{M})$ определим следующим образом

$$r_S(x, e_a) = (x, e'_a), \quad \text{где} \quad e'_a := S_a^b e_b.$$

Замечание. Мы записываем правое действие структурной группы в виде матричного умножения слева. В этом нет неоднозначности, т.к. суммирование всегда проводится по одному нижнему и одному верхнему индексу, $S_a^b e_b = e_b S_a^b$. Наши обозначения

вызваны принятым правилом суммирования “с десяти до четырех”. Это правило является следствием общепринятой записи $dx^\alpha \partial_\alpha$. Другая запись $\partial_\alpha dx^\alpha$ может вызвать недоразумения. Кроме того, векторы репера принято нумеровать нижним индексом, поскольку координатные векторы ∂_α образуют координатный репер. $\square$

Очевидно, что каждое преобразование r_S задает диффеоморфизм $\mathbb{L}(\mathbb{M})$ и сохраняет слои: $\pi \circ r_S = \pi$. Действие r_S называется *правым действием структурной группы*. Если $S, T \in \mathbb{GL}(n, \mathbb{R})$, то

$$r_{ST} = r_S \circ r_T.$$

Действие группы преобразований $\mathbb{GL}(n, \mathbb{R})$ свободно, т.е. любой элемент группы, отличный от единичного, перемещает все точки многообразия $\mathbb{L}(\mathbb{M})$.

Обозначим матрицу, обратную к реперу e^α_a , через e_β^b :

$$e^\alpha_a e_\alpha^b = \delta_a^b, \quad e_\alpha^a e_\beta^b = \delta_\alpha^\beta.$$

Тогда при преобразовании координат (5.41) на пространстве расслоения $\mathbb{L}(\mathbb{M})$ она преобразуется, как ковекторное поле

$$e_{\alpha'}^a = \partial_{\alpha'} x^\alpha e_\alpha^a.$$

Отсюда следует, что 1-формы (корепер)

$$e^a = dx^\alpha e_\alpha^a \quad (5.42)$$

не зависят от выбора координат на $\mathbb{M}$. Это значит, что e^a являются дифференциальными 1-формами, заданными в каждой координатной окрестности согласованным образом и, следовательно, определяющими некоторую 1-форму на всем пространстве расслоения $\mathbb{L}(\mathbb{M})$. Напомним, что числа e_α^a в (5.42) рассматриваются в качестве координат на $\mathbb{L}(\mathbb{M})$. Поэтому e_α^a представляют собой набор хорошо определенных функций на $\mathbb{L}(\mathbb{M})$, поскольку $\det e_\alpha^a \neq 0$.

Пфафова система уравнений

$$e^a = 0 \quad (5.43)$$

определяет n^2 -мерное касательное подпространство $\mathbb{V}_p(\mathbb{L})$ в касательном пространстве $\mathbb{T}_p(\mathbb{L})$ к каждой точке расслоения $p = (x, e) \in \mathbb{L}(\mathbb{M})$, которое называется *вертикальным подпространством*. Система уравнений (5.43) в каждой координатной окрестности эквивалентна системе уравнений

$$dx^\alpha = 0.$$

Эта система уравнений является вполне интегрируемой, и ее максимальное интегральное многообразие задается уравнениями

$$x^\alpha = \text{const},$$

т.е. является слоем $\pi^{-1}(x)$ в точке $x \in \mathbb{M}$. Это значит, что вертикальные подпространства являются касательными пространствами к слоям.

Векторное поле на пространстве расслоения $X \in \mathcal{X}(\mathbb{L})$ в компонентах имеет вид

$$X = X^\alpha \frac{\partial}{\partial x^\alpha} + X^a_a \frac{\partial}{\partial e^a_a}.$$

Поэтому векторное поле Y принадлежит вертикальному подпространству $Y \in \mathbb{V}_p(\mathbb{L})$ в каждой точке $p = (x, e) \in \mathbb{L}(\mathbb{M})$ тогда и только тогда, когда оно имеет компоненты только вдоль координат слоя:

$$Y = Y^\alpha{}_a \frac{\partial}{\partial e^\alpha{}_a}$$

Дифференциал π_* проекции (5.39) отображает касательные пространства $\mathbb{T}_p(\mathbb{L}) \rightarrow \mathbb{T}_x(\mathbb{M})$. Очевидно, что все векторы из вертикального подпространства $\mathbb{V}_p(\mathbb{L})$ отображаются в нулевой вектор $\mathbb{T}_x(\mathbb{M})$. Верно и обратное утверждение. Если какой-либо вектор из $\mathbb{T}_p(\mathbb{L})$ отображается в нулевой вектор, то он принадлежит вертикальному подпространству. Это значит, что вертикальные подпространства являются ядром дифференциала проекции π_* .

Пусть на базе $\mathbb{M}$ задана аффинная связность Γ или, что эквивалентно, 1-формы $\Gamma_\alpha{}^\beta = dx^\gamma \Gamma_{\gamma\alpha}{}^\beta$. Тогда ковариантный дифференциал векторного поля $e_a(x)$ равен

$$De_a = (de^\alpha{}_a + e^\beta{}_a \Gamma_\beta{}^\alpha) \otimes \partial_\alpha.$$

Если рассматривать $e^\alpha{}_a$, как координаты на $\mathbb{L}(\mathbb{M})$ (что мы и делаем), то выражение

$$De^\alpha{}_a = de^\alpha{}_a + e^\beta{}_a \Gamma_\beta{}^\alpha$$

является 1-формой на координатной окрестности $\pi^{-1}(\mathbb{U})$ расслоения реперов $\mathbb{L}(\mathbb{M})$. Поскольку 1-форма $De^\alpha{}_a$ при преобразовании координат является вектором по индексу α , то 1-форма

$$\omega_a{}^b = e_a{}^b De^\alpha{}_a \quad (5.44)$$

не зависит от выбора системы координат и, значит, задает дифференциальную 1-форму на всем расслоении реперов $\mathbb{L}(\mathbb{M})$. 1-форма $\omega_a{}^b$ называется *формой линейной связности* на расслоении реперов. Так как $(x^\alpha, e^\beta{}_a)$ являются координатами на $\mathbb{L}(\mathbb{U})$, то $(dx^\alpha, de^\beta{}_a)$ представляют собой координатный базис кокасательного пространства к $\mathbb{L}(\mathbb{U})$. Поэтому $(x^\alpha, e^\beta{}_a, dx^\alpha, de^\beta{}_a)$ задают систему координат кокасательного расслоения $\mathbb{T}^*(\mathbb{L})$ к $\mathbb{L}(\mathbb{U})$. Теперь e^a и $\omega_a{}^b$ задают $n + n^2$ дифференциальных 1-форм на расслоении реперов $\mathbb{L}(\mathbb{U})$. В каждой координатной окрестности $\pi^{-1}(\mathbb{U})$ их можно выразить в виде линейной комбинации 1-форм dx^α и $de^\alpha{}_a$, и наоборот. Поскольку 1-формы e^a и $\omega_a{}^b$ линейно независимы, то они задают поле корепера на всем пространстве $\mathbb{L}(\mathbb{M})$. Дуальный к нему базис определяет поле репера на пространстве расслоения $\mathbb{L}(\mathbb{M})$ глобально.

Замечание. В общем случае сечение расслоения реперов $e_a(x)$ может не существовать глобально. Например, на двумерной сфере $\mathbb{S}^2$ не существует непрерывного векторного поля, которое не обращается в нуль ни в одной точке (теорема 10.2.1). Следовательно, глобального сечения $\mathbb{L}(\mathbb{S}^2)$ не существует. В то же время, поскольку на произвольном многообразии $\mathbb{M}$ существует аффинная связность, то на расслоении реперов $\mathbb{L}(\mathbb{M})$ всегда существует корепер e^a глобально. В этом смысле многообразие расслоения реперов $\mathbb{L}(\mathbb{M})$ устроено проще, чем база $\mathbb{M}$. $\square$

В локальной системе координат из определения 1-форм e^a и $\omega_a{}^b$ следуют равенства

$$\begin{aligned} dx^\alpha &= e^a e^\alpha{}_a, \\ de^\alpha{}_a &= -e^\beta{}_a \Gamma_\beta{}^\alpha + \omega_a{}^b e^\alpha{}_b. \end{aligned}$$

Здесь, по-прежнему, все формы рассматриваются на пространстве главного расслоения $\mathbb{L}(\mathbb{M})$. Взятие внешней производной от равенств (5.42) и (5.44) приводит к *структурным уравнениям* для аффинной связности:

$$de^a - e^b \wedge \omega_b^a = \frac{1}{2} e^b \wedge e^c T_{bc}^a, \quad (5.45)$$

$$d\omega_a^b - \omega_a^c \wedge \omega_c^b = \frac{1}{2} e^c \wedge e^d R_{cda}^b, \quad (5.46)$$

где $T_{bc}^a := e^\beta_b e^\gamma_c T_{\beta\gamma}^\alpha e_\alpha^a$ и $R_{cda}^b := e^\gamma_c e^\delta_d e^\alpha_a R_{\gamma\delta\alpha}^\beta e_\beta^b$ – компоненты тензоров кручения и кривизны. Компоненты тензоров кручения $T_{\beta\gamma}^\alpha$ и кривизны $R_{\gamma\delta\alpha}^\beta$ в координатном базисе были определены ранее (5.38) и (5.36). Введем обозначения для 2-форм кручения и кривизны:

$$\begin{aligned} T^a &:= \frac{1}{2} e^b \wedge e^c T_{bc}^a, \\ R_a^b &:= \frac{1}{2} e^c \wedge e^d R_{cda}^b. \end{aligned} \quad (5.47)$$

Взятие внешней производной от уравнений структуры (5.45), (5.46) приводит к *тождествам Бианки* для аффинной связности:

$$\begin{aligned} dT^a + T^b \wedge \omega_b^a &= e^b \wedge R_b^a, \\ dR_a^b + R_a^c \wedge \omega_c^b - \omega_a^c \wedge R_c^b &= 0. \end{aligned} \quad (5.48)$$

Второе из этих тождеств совпадает с полученной ранее формулой (5.21). В компонентах тождества Бианки будут подробно рассмотрены в разделе 6.10.

Замечание. Структурные уравнения (5.45) и (5.46) часто принимают за определение тензоров кручения и кривизны. $\square$

В координатном базисе $e^a \mapsto dx^\alpha$, $\omega_a^b \mapsto \Gamma_\alpha^\beta := dx^\gamma \Gamma_{\gamma\alpha}^\beta$, и структурные уравнения принимают вид

$$\begin{aligned} -dx^\beta \wedge \Gamma_\beta^\alpha &= \frac{1}{2} dx^\beta \wedge dx^\gamma T_{\beta\gamma}^\alpha, \\ d\Gamma_\alpha^\beta - \Gamma_\alpha^\gamma \wedge \Gamma_\gamma^\beta &= \frac{1}{2} dx^\gamma \wedge dx^\delta R_{\gamma\delta\alpha}^\beta, \end{aligned} \quad (5.49)$$

где в правых частях стоят тензоры кручения $T_{\beta\gamma}^\alpha$ и кривизны $R_{\gamma\delta\alpha}^\beta$ и учтено равенство $ddx^\alpha = 0$.

Теорема 5.4.1. Пусть на расслоении реперов $\mathbb{L}(\mathbb{M})$ заданы n^2 дифференциальных 1-форм ω_a^b , которые вместе с корепером e^a удовлетворяют структурным уравнениям (5.45), (5.46), где T_{bc}^a и R_{cda}^b – некоторые функции, определенные на $\mathbb{L}(\mathbb{M})$. Тогда существует аффинная связность ∇ такая, что выполнено равенство (5.44).

Доказательство. См., например, [49], глава 4, теорема 3.1. $\square$

Замечание. От компонент, определяющих кручение и кривизну, нет надобности требовать выполнения тождеств Бианки, т.к. они являются следствием структурных уравнений, выполнение которых является условием теоремы. $\square$

Ранее было отмечено, что пфафхова система уравнений $e^a = 0$ определяет распределение вертикальных подпространств $\mathbb{V}_p(\mathbb{L}) \subset \mathbb{T}_p(\mathbb{L})$ на расслоении реперов $\mathbb{L}(\mathbb{M})$. В каждой точке $(x, e) \in \mathbb{L}(\mathbb{M})$ система пфаффовых уравнений

$$\omega_a{}^b = 0 \quad (5.50)$$

определяет n -мерное касательное подпространство $\mathbb{H}_p(\mathbb{L}) \subset \mathbb{T}_p(\mathbb{L})$, которое называется *горизонтальным*. Таким образом, мы имеем распределение горизонтальных подпространств $\mathbb{H}(\mathbb{L})$, однозначно определенное аффинной связностью ∇ на расслоении реперов $\mathbb{L}(\mathbb{M})$. Оно имеет следующие свойства.

Теорема 5.4.2. *Пусть задано расслоение реперов $\mathbb{L}(\mathbb{M})$ и на базе $\mathbb{M}$ определена аффинная связность ∇ . Тогда определены распределения вертикальных $\mathbb{V}_p(\mathbb{L})$ и горизонтальных $\mathbb{H}_p(\mathbb{L})$ подпространств, которые удовлетворяют свойствам:*

- 1) *В каждой точке $(x, e) \in \mathbb{L}(\mathbb{M})$ касательное пространство $\mathbb{T}_p(\mathbb{L})$ разлагается в прямую сумму*

$$\mathbb{T}_p(\mathbb{L}) = \mathbb{V}_p(\mathbb{L}) \oplus \mathbb{H}_p(\mathbb{L}). \quad (5.51)$$

При этом образ горизонтального пространства $\mathbb{H}_p(\mathbb{L})$ при проекции π изоморфен касательному пространству к базе $\mathbb{T}_x(\mathbb{M})$.

- 2) *Горизонтальные пространства инвариантны относительно правого действия группы преобразований r_S , $S \in \mathbb{GL}(n, \mathbb{R})$. То есть*

$$(r_S)_* \mathbb{H}_p(\mathbb{L}) = \mathbb{H}_{r_S(x,p)}(\mathbb{L}).$$

Доказательство. Докажем свойство 1). Поскольку сумма размерностей горизонтального $\mathbb{H}_p(\mathbb{L})$ и вертикального $\mathbb{V}_p(\mathbb{L})$ подпространств равна размерности касательного пространства $\mathbb{T}_p(\mathbb{L})$, то достаточно доказать, что $\mathbb{H}_p(\mathbb{L}) \cap \mathbb{V}_p(\mathbb{L}) = \emptyset$. От противного. Пусть $X \in \mathbb{H}_p(\mathbb{L}) \cap \mathbb{V}_p(\mathbb{L})$. Тогда, по определению, выполнены уравнения

$$e^a(X) = 0, \quad \omega_a{}^b(X) = 0.$$

Поскольку 1-формы $\{e^a, \omega_b{}^c\}$ образуют базис кореперов на $\mathbb{L}(\mathbb{M})$, то $X = 0$. Это доказывает разложение (5.51). Поскольку проекция π является гладким и сюръективным отображением, то ее дифференциал π_* является сюръективным гомоморфизмом. Поскольку $\pi_*(\mathbb{V}_p(\mathbb{L})) = 0$, то отображение $\pi_* : \mathbb{H}_p(\mathbb{L}) \rightarrow \mathbb{T}_x(\mathbb{M})$ является изоморфизмом.

Для доказательства свойства 2) необходимо проверить, что правое действие группы преобразований на 1-формы $\omega_a{}^b$ имеет вид

$$(r_S)^* \omega_a{}^b = S_a{}^c \omega_c{}^d S_d{}^{-1b}.$$

Действительно, это следует из определения (5.44) и равенства

$$de'^\alpha{}_a = \frac{\partial e'^\alpha{}_a}{\partial e^\beta{}_b} de^\beta{}_b = S_a{}^b de^\alpha{}_b.$$

Поскольку распределение горизонтальных подпространств $\mathbb{H}(\mathbb{L})$ является аннигилятором подпространства $\omega_a{}^b$, то из правила действия левых преобразований следует свойство 2). $\square$

Верно также обратное утверждение. Если в касательном расслоении $\mathbb{T}(\mathbb{L})$ задано дифференцируемое n -мерное распределение горизонтальных подпространств $\mathbb{H}$, удовлетворяющих свойствам 1) и 2), то на $\mathbb{M}$ существует аффинная связность ∇ такая, что $\mathbb{H}$ является распределением горизонтальных подпространств расслоения реперов $\mathbb{L}(\mathbb{M})$ для связности ∇ . Поэтому, с точки зрения расслоения реперов, задание аффинной связности эквивалентно заданию распределения горизонтальных подпространств со свойствами 1), 2). Далее эти свойства лягут в основу определения связности на главных расслоениях общего вида.

Замечание. Поскольку аффинная связность ∇ зависит дифференцируемо от $p \in \mathbb{L}(\mathbb{M})$, то распределение горизонтальных подпространств $\mathbb{H}_p(\mathbb{L})$ также дифференцируемо зависит от точки p . $\square$

В формулах (5.45)–(5.48) все формы рассматривались на пространстве расслоения $\mathbb{L}(\mathbb{M})$. В частности, компоненты тензоров кручения T_{bc}^a и кривизны R_{cda}^b в структурных уравнениях (5.45) и (5.46) явно зависят от координат слоя e^α_a . Эти формы можно спустить на базу с помощью возврата отображения.

Определение. Рассмотрим сечение $\sigma : \mathbb{M} \rightarrow \mathbb{L}(\mathbb{M})$, которому в локальных координатах соответствует репер $e^\alpha_a(x)$. Тогда дифференциал сечения σ_* отображает касательные векторы к базе в касательные векторы к сечению:

$$\mathcal{X}(\mathbb{M}) \ni X = X^\alpha \partial_\alpha \mapsto \sigma_* X = X^\alpha \partial_\alpha + X^\beta \partial_\beta e^\alpha_a \frac{\partial}{\partial e^\alpha_a} \in \mathcal{X}(\sigma(\mathbb{M})).$$

Следовательно, возврат отображения σ^* спускает форму связности для каждого сечения на базу:

$$\sigma^* \omega_a^b := dx^\alpha \omega_{\alpha a}^b = dx^\alpha e^\gamma_b (\partial_\alpha e^\gamma_a + e^\beta_a \Gamma_{\alpha\beta}^\gamma). \quad (5.52)$$

Форма $dx^\alpha \omega_{\alpha a}^b$ называется *локальной формой связности* на расслоении реперов. Она получается из формы связности (5.44) простой заменой: $e^\alpha_a \mapsto e^\alpha_a(x)$ и $de^\alpha_a \mapsto dx^\beta \partial_\beta e^\alpha_a$. Аналогично спускаются на базу 2-формы кручения и кривизны (5.47):

$$T^a = \frac{1}{2} dx^\alpha \wedge dx^\beta T_{\alpha\beta}^a = de^a - e^b \wedge \omega_b^a, \quad (5.53)$$

$$R_a^b = \frac{1}{2} dx^\alpha \wedge dx^\beta R_{\alpha\beta a}^b = d\omega_a^b - \omega_a^c \wedge \omega_c^b, \quad (5.54)$$

где введены 1-формы локальной связности и корепера:

$$\omega_a^b = dx^\alpha \omega_{\alpha a}^b, \quad e^a = dx^\alpha e_\alpha^a, \quad (5.55)$$

и

$$T_{\alpha\beta}^a = \partial_\alpha e_\beta^a - e_\alpha^b \omega_{\beta b}^a - (\alpha \leftrightarrow \beta), \quad (5.56)$$

$$R_{\alpha\beta a}^b = \partial_\alpha \omega_{\beta a}^b - \omega_{\alpha a}^c \omega_{\beta c}^b - (\alpha \leftrightarrow \beta). \quad (5.57)$$

2-формы (5.53) и (5.54) называются *локальными формами* кручения и кривизны. $\square$

В приложениях выражения (5.56) и (5.57) часто называют тензорами кручения и кривизны.

Замечание. Для локальных форм мы сохранили прежние обозначения, чтобы не вводить новых. Фактически, в приложениях при проведении вычислений используются локальные формы, заданные на базе $\mathbb{M}$, а не на расслоении реперов $\mathbb{L}(\mathbb{M})$. $\square$

Тождества Бианки (5.48) после спуска на базу имеют точно такой же вид, только вместо компонент корепера e_α^a надо рассматривать сечение $e_\alpha^a(x)$, вместо формы связности ω_a^b , определенную в (5.44), – локальную форму связности $dx^\alpha \omega_{\alpha a}^b(x)$ и вместо форм кручения и кривизны – соответствующие им локальные формы.

Определение. Поля $e_\alpha^a(x)$ и $\omega_{\alpha a}^b(x)$, заданные на $\mathbb{M}$, в аффинной геометрии называются *переменными Картана*. $\square$

Переменные Картана часто используются в приложениях, т.к. позволяют упростить вычисления. Формулы (5.56) и (5.54) дают выражение для тензоров кручения и кривизны в переменных Картана.

Нетрудно проверить, что при локальных лоренцевых вращениях репера (5.10) и лоренцевой связности (5.12) компоненты локальных форм кривизны и кручения преобразуются по тензорному закону:

$$R'_{\alpha\beta a}{}^b = S_a{}^c R_{\alpha\beta c}{}^d S_d^{-1b}, \quad T'_{\alpha\beta}{}^a = T_{\alpha\beta}{}^b S_b^{-1a}. \quad (5.58)$$

В инфинитезимальной форме матрицы вращений имеют вид

$$S_a{}^b \approx \delta_a^b + \omega_a^b, \quad S_a^{-1b} \approx \delta_a^b - \omega_a^b, \quad |\omega_a^b| \ll 1. \quad (5.59)$$

Это приводит к следующему правилу преобразования компонент репера и лоренцевой связности в линейном приближении:

$$\delta e_\alpha^a = \omega_a^b e_\alpha^b, \quad \delta e_\alpha^a = -e_\alpha^b \omega_b^a, \quad \delta \omega_{\alpha a}^b = \omega_a^c \omega_{\alpha c}^b - \omega_{\alpha a}^c \omega_c^b + \partial_\alpha \omega_a^b. \quad (5.60)$$

В свою очередь, преобразование репера и формы связности приводит к преобразованию компонент локальных форм кривизны и кручения, которое в линейном приближении имеет вид

$$\delta R_{\alpha\beta a}{}^b = \omega_a^c R_{\alpha\beta c}{}^b - R_{\alpha\beta a}{}^c \omega_c^b, \quad \delta T_{\alpha\beta}{}^a = T_{\alpha\beta}{}^b \omega_b^a. \quad (5.61)$$

В заключение раздела приведем правила преобразования компонент кривизны и кручения при инфинитезимальных преобразованиях координат (2.96). Преобразования компонент репера и лоренцевой связности уже были выписаны (2.113) и (2.114). Соответствующее преобразование компонент локальных форм кривизны и кручения в линейном приближении дается формулами:

$$\begin{aligned} \delta R_{\alpha\beta a}{}^b &= -\partial_\alpha u^\gamma R_{\gamma\beta a}{}^b - \partial_\beta u^\gamma R_{\alpha\gamma a}{}^b - u^\gamma \partial_\gamma R_{\alpha\beta a}{}^b = \\ &= -\nabla_\alpha u^\gamma R_{\gamma\beta a}{}^b - \nabla_\beta u^\gamma R_{\alpha\gamma a}{}^b - u^\gamma \nabla_\gamma R_{\alpha\beta a}{}^b + \\ &\quad + u^\gamma T_{\alpha\gamma}{}^\delta R_{\delta\beta a}{}^b + u^\gamma T_{\beta\gamma}{}^\delta R_{\alpha\delta a}{}^b - u^\gamma \omega_{\gamma a}^c R_{\alpha\beta c}{}^b + u^\gamma \omega_{\gamma c}^b R_{\alpha\beta a}{}^c, \\ \delta T_{\alpha\beta}{}^a &= -\partial_\alpha u^\gamma T_{\gamma\beta}{}^a - \partial_\beta u^\gamma T_{\alpha\gamma}{}^a - u^\gamma \partial_\gamma T_{\alpha\beta}{}^a = \\ &= -\nabla_\alpha u^\gamma T_{\gamma\beta}{}^a - \nabla_\beta u^\gamma T_{\alpha\gamma}{}^a - u^\gamma \nabla_\gamma T_{\alpha\beta}{}^a + \\ &\quad + u^\gamma T_{\alpha\gamma}{}^\delta T_{\delta\beta}{}^a + u^\gamma T_{\beta\gamma}{}^\delta T_{\alpha\delta}{}^a + u^\gamma \omega_{\gamma b}^a T_{\alpha\beta}{}^b. \end{aligned} \quad (5.62)$$

Появление слагаемых с лоренцевой связностью в правых частях вызвано тем, что она входит в ковариантную производную тензоров кривизны и кручения. Напомним, что при преобразованиях координат компоненты $\omega_{\alpha a}^b$ ведут себя как компоненты ковектора.

5.5 Критерий локальной тривиальности

Тензоры кривизны и кручения играют исключительно важную роль в аффинной геометрии. Ниже мы покажем, что их обращение в нуль является критерием локальной тривиальности связности и существования такой системы координат, в которой репер совпадает с координатным базисом касательного пространства.

Рассмотрим расслоение реперов $\mathbb{L}(\mathbb{M})$ с заданной связностью ∇ и координатную окрестность $\mathbb{U}$ произвольной точки базы $x \in \mathbb{M}$. Пусть задано некоторое локальное сечение $\sigma : \mathbb{M} \supset \mathbb{U} \rightarrow \mathbb{L}$. Тогда в области $\mathbb{U}$ определен репер $e_\alpha^a(x)$ и компоненты локальной формы связности $\omega_{\alpha a}^b(x)$. Кроме того, определены тензоры кривизны $R_{\alpha\beta a}^b$ и кручения $T_{\alpha\beta}^a$, которые заданы формулами (5.56) и (5.57).

При локальном вращении репера (5.10) с матрицей $S(x) \in \mathbb{GL}(n, \mathbb{R})$ компоненты линейной связности преобразуются по правилу (5.11). Поставим следующий вопрос: “Существует ли такая матрица вращений S , что после преобразования компоненты локальной формы связности $\omega'_{\alpha a}^b$ обратятся в нуль в некоторой окрестности точки x ?”. Ответ на этот вопрос дает следующая

Теорема 5.5.1. *Пусть в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ задана локальная форма аффинной связности. Тогда равенство нулю тензора кривизны $R_{\alpha\beta a}^b = 0$ в некоторой (односвязной) окрестности $\mathbb{U}$ произвольной точки $x \in \mathbb{M}$ является необходимым и достаточным условием существования такой матрицы $S(x)$, что после локального вращения репера (5.10) компоненты локальной формы связности обратятся в нуль в, возможно, меньшей окрестности точки x .*

Доказательство. Если после вращения $\omega'_{\alpha a}^b = 0$, то равенство (5.11) после умножения справа на S приводит к уравнению на компоненты матрицы вращений:

$$\partial_\alpha S_a^b = -S_a^c \omega_{\alpha c}^b. \quad (5.63)$$

Для того, чтобы получить критерий локальной разрешимости этой системы уравнений, ее нужно продифференцировать по x^β и антисимметризовать по индексам α и β . После исключения первых производных от матрицы S в правой части с помощью исходного уравнения, получим равенство

$$(\partial_\beta \partial_\alpha - \partial_\alpha \partial_\beta)S = S(\partial_\alpha \omega_\beta - \partial_\beta \omega_\alpha - \omega_\alpha \omega_\beta + \omega_\beta \omega_\alpha),$$

где мы, для краткости, опустили матричные индексы. Сравнивая это равенство с определением компонент локальной формы кривизны (5.56), находим искомый критерий. $\square$

На этом этапе можно забыть о существовании репера. Достаточно считать, что на $\mathbb{U}$ заданы только компоненты локальной формы связности $\omega_{\alpha a}^b(x)$ с соответствующим правилом преобразования. Про метрику можно вообще не вспоминать: она просто отсутствует на многообразии $\mathbb{M}$.

При нулевой кривизне зафиксируем матрицу вращений таким образом, что $\omega_{\alpha a}^b = 0$. Тогда после локального поворота компоненты локальной формы связности станут нетривиальными в соответствии с правилом преобразования (5.11). Это доказывает следующее утверждение.

Следствие. Пусть тензор кривизны для некоторой линейной связности обращается в нуль в некоторой (односвязной) области $\mathbb{U} \subset \mathbb{M}$. Тогда, возможно, в меньшей

окрестности существует такая матрица вращений $S(x) \in \mathbb{GL}(n, \mathbb{R})$, что компоненты связности представимы в следующем виде

$$\omega_{\alpha a}{}^b = \partial_\alpha S_a{}^c S^{-1}{}_c{}^b. \quad \square \quad (5.64)$$

Такая связность называется чистой калибровкой и является плоской, т.к. ее тензор кривизны тождественно равен нулю.

Теперь вернемся к вопросу, поставленному в начале раздела. Рассмотрим односвязную область $U \subset M$ с координатами x^α . Если в этой области существует такая система координат $y^a(x)$, что соответствующий координатный базис касательного пространства совпадает с заданным репером, то должно выполняться равенство

$$\partial_\alpha y^a = e_\alpha{}^a. \quad (5.65)$$

Это – система уравнений на функции перехода к новой системе координат $y^a(x)$. Критерием ее разрешимости являются условия

$$\partial_\alpha e_\beta{}^a - \partial_\beta e_\alpha{}^a = 0. \quad (5.66)$$

Левая часть этого равенства совпадает с тензором кручения (5.57) при нулевой связности $\omega_{\alpha a}{}^b = 0$. Поэтому справедлива

Теорема 5.5.2. *Пусть на M заданы репер и локальная форма аффинной связности. Тогда одновременное обращение в нуль тензоров кривизны и кручения в некоторой односвязной области $U \subset M$ является необходимым и достаточным условием существования в, возможно, меньшей области такой матрицы $S(x) \in \mathbb{GL}(n, \mathbb{R})$ и системы координат $y^a(x)$, что компоненты связности обратятся в нуль, $\omega_{\alpha a}{}^b = 0$, и репер совпадет с координатным базисом касательного пространства, $e_\alpha{}^a = \partial_\alpha y^a$.*

Подчеркнем, что на многообразии M не предполагается наличие какой-либо метрики.

В дальнейшем мы покажем, что последняя теорема лежит в основе геометрической теории дефектов в упругой среде.

Теперь предположим дополнительно, что на M задана метрика с компонентами $g_{\alpha\beta}$. Для определенности, предположим, что метрика положительно определена, т.е. риманова. В такой ситуации в приложениях рассматривают, как правило, не всю совокупность реперов, а их подмножество, которое определяется уравнением

$$g_{\alpha\beta} = e_\alpha{}^a e_\beta{}^b \delta_{ab}, \quad (5.67)$$

где $\delta_{ab} = \text{diag}(+ \dots +)$ – евклидова метрика. Решения данного квадратного уравнения относительно репера существуют и определены с точностью до локальных $\mathbb{O}(n)$ вращений. То есть, если репер $e_\alpha{}^a$ удовлетворяет уравнению (5.67), то повернутый репер $e'_\alpha{}^a = e_\alpha{}^b S_b{}^a$, где $S \in \mathbb{O}(n)$ – произвольная матрица ортогональных вращений из полной группы $\mathbb{O}(n)$, также удовлетворяет данному уравнению. Поскольку произвол в выборе репера теперь сужен до подгруппы вращений $\mathbb{O}(n) \subset \mathbb{GL}(n, \mathbb{R})$, то естественно рассматривать не полную линейную связность, а $\mathbb{SO}(n)$ связность, которая удовлетворяет условию

$$\nabla_\alpha \delta_{ab} = -\omega_{\alpha a}{}^c \delta_{cb} - \omega_{\alpha b}{}^c \delta_{ac} = 0. \quad (5.68)$$

То есть компоненты связности $\omega_{\alpha a}{}^b$ принимают значения в алгебре вращений $\mathfrak{so}(n)$. Эта связность является метрической, поскольку ковариантная производная от метрики (5.68) равна нулю, а соответствующая геометрия – геометрией Римана–Картана.

Напомним, что алгебры Ли собственных вращений $\mathfrak{so}(n)$ и полной группы вращений $\mathfrak{o}(n)$ изоморфны. Поэтому $\mathbb{SO}(n)$ - и $\mathbb{O}(n)$ связности – это одно и то же.

Согласно теореме 5.5.2 при нулевой кривизне и кручении существует такая матрица вращений и система координат y^a , что будет выполнено равенство

$$g_{\alpha\beta} = \partial_\alpha y^a \partial_\beta y^b \delta_{ab}.$$

Это означает, что существует такая система координат в которой метрика является локально евклидовой. Тем самым доказана

Теорема 5.5.3. Пусть на многообразии $\mathbb{M}$ задана геометрия Римана–Картана, т.е. в произвольной области определены компоненты репера e_α^a , метрики $g_{\alpha\beta} = e_\alpha^a e_\beta^b \delta_{ab}$ и $\mathbb{SO}(n)$ связности $\omega_{\alpha a}^b$. Если тензоры кривизны и кручения в некоторой (односвязной) области $\mathbb{U} \subset \mathbb{M}$ обращаются в нуль, то, возможно, в меньшей области существует такая система координат $y^a(x)$, в которой метрика является евклидовой, $g_{ab} = \delta_{ab}$. При этом локальное лоренцево вращение можно подобрать таким образом, что $e_\alpha^a = \delta_\alpha^a$ и $\omega_{\alpha a}^b = 0$.

Если выбрать произвольный репер для заданной метрики $g_{\alpha\beta}$, то условия интегрируемости (5.66) в общем случае выполнены не будут. Однако равенство нулю тензора кривизны гарантирует существование такого вращения, что после поворота репера условия интегрируемости окажутся выполнены. Этот повернутый репер и определит функции перехода $y^a(x)$ к нужной системе координат.

В настоящем разделе мы подчеркивали, что рассматриваются только односвязные области. В дальнейшем мы увидим, что для неодносвязных областей приведенные выше утверждения в общем случае не имеют места.

В римановой геометрии кручение равно нулю с самого начала. И, конечно, теорема о приведении метрики к локально евклидову виду верна и в этом случае. Достаточно потребовать равенства нулю только тензора кривизны.

В доказательстве теоремы 5.5.3 сигнатура метрики никак не использовалась. Поэтому сформулированное утверждение верно для метрик произвольной сигнатуры. Например, в равенстве (5.67) евклидову метрику можно заменить на метрику Лоренца η_{ab} . Тогда вместо $\mathbb{SO}(n)$ связности следует рассматривать лоренцеву $\mathbb{SO}(1, n-1)$ связность, и все предыдущее построение остается в силе.

Глава 6

Аффинная геометрия. Локальное рассмотрение

В современной дифференциальной геометрии принят бескоординатный язык. Это оправдано тем, что все геометрические объекты, например, тензорные поля, являются понятиями, которые не зависят от выбора системы координат. Этот язык удобен для определений и формулировки утверждений. Однако он абстрактен, и требуется немалое время для его усвоения. Кроме того, при проведении вычислений, особенно в моделях математической физики, используется координатный подход, в котором геометрические объекты отождествляются с набором своих компонент. Этот язык действительно необходим для проведения вычислений и является более наглядным. В настоящем разделе мы дадим определения и опишем основные свойства объектов аффинной геометрии в компонентах. По сути дела это означает, что мы рассматриваем топологически тривиальное многообразие $\mathbb{M}$, которое диффеоморфно евклидову пространству $\mathbb{R}^n$ и покрывается одной картой. Это не значит, что координатный подход является менее строгим. Как только установлено, что некоторый объект или утверждение не зависят от выбора системы координат в $\mathbb{R}^n$, то все сказанное сразу переносится на нетривиальные многообразия.

6.1 Локальное определение аффинной связности

Ввиду важности понятия аффинной связности покажем, как она вводится локально без обращения к общей теории связностей в векторных расслоениях, которая была рассмотрена в предыдущей главе. Пусть на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, в некоторой карте заданы: функция $f \in \mathcal{C}^1(\mathbb{M})$, дифференцируемое векторное поле $X = X^\alpha \partial_\alpha \in \mathcal{X}(\mathbb{M})$ и дифференцируемая 1-форма $A = dx^\alpha A_\alpha \in \Lambda_1(\mathbb{M})$. Пусть на этом многообразии задана также аффинная связность ∇ . Тогда в компонентах ковариантные производные имеют вид:

$$\nabla_\alpha f = \partial_\alpha f \quad (6.1)$$

$$\nabla_\alpha X^\beta = \partial_\alpha X^\beta + X^\gamma \Gamma_{\alpha\gamma}^\beta, \quad (6.2)$$

$$\nabla_\alpha A_\beta = \partial_\alpha A_\beta - \Gamma_{\alpha\beta}^\gamma A_\gamma, \quad (6.3)$$

где $\Gamma_{\alpha\beta}^\gamma(x)$ – компоненты локальной формы аффинной связности, которые, мы предполагаем, являются достаточно гладкими функциями от координат x^α . То есть ковариантная производная для скалярного поля совпадает с обычной частной производной, а для векторного поля и 1-формы появляются дополнительные слагаемые,

линейные по компонентам. Отметим, что дополнительные слагаемые в (6.2) и (6.3) имеют разные знаки.

В дальнейшем мы часто будем обозначать аффинную связность ∇ буквой Γ как это принято в римановой геометрии.

Покажем, что ковариантная производная от компонент векторного поля при подходящем преобразовании аффинной связности преобразуется как тензор. Рассмотрим преобразование координат (1.66). Тогда компоненты векторного поля преобразуются по правилу (2.11). Дифференцируя соотношение (2.11) по $x^{\beta'}$, получаем два слагаемых

$$\partial_{\beta'} X^{\alpha'} = \partial_{\beta'} x^{\beta} \partial_{\beta} X^{\alpha} \partial_{\alpha} x^{\alpha'} + \partial_{\beta'} x^{\beta} X^{\alpha} \partial_{\beta\alpha}^2 x^{\alpha'}, \quad (6.4)$$

где $\partial_{\beta'} = \partial_{\beta'} x^{\beta} \partial_{\beta}$. Первое слагаемое в правой части соответствует тензорному закону преобразования для производной от векторного поля, в то время как второе слагаемое этот закон нарушает.

Чтобы получить тензорный закон преобразования в общем случае вводится понятие ковариантной производной, которая содержит дополнительное слагаемое в выражении (6.2). Если потребовать, чтобы ковариантная производная после преобразования координат имела тот же вид с некоторыми новыми компонентами $\Gamma_{\beta'\gamma'}^{\alpha'}$,

$$\partial_{\beta'} X^{\alpha'} + \Gamma_{\beta'\gamma'}^{\alpha'} X^{\gamma'} = \partial_{\beta'} x^{\beta} (\partial_{\beta} X^{\alpha} + \Gamma_{\beta\gamma}^{\alpha} X^{\gamma}) \partial_{\alpha} x^{\alpha'},$$

то с учетом уравнения (6.4) получим следующий закон преобразования компонент аффинной связности:

$$\Gamma_{\beta'\gamma'}^{\alpha'} = \partial_{\beta'} x^{\beta} \partial_{\gamma'} x^{\gamma} \Gamma_{\beta\gamma}^{\alpha} \partial_{\alpha} x^{\alpha'} - \partial_{\beta'} x^{\beta} \partial_{\gamma'} x^{\gamma} \partial_{\beta\gamma}^2 x^{\alpha'}, \quad (6.5)$$

С учетом тождества

$$\partial_{\alpha} (\partial_{\beta} x^{\beta'} \partial_{\beta'} x^{\gamma}) = \partial_{\alpha} \delta_{\gamma}^{\beta} = 0 \quad \Leftrightarrow \quad \partial_{\alpha\beta}^2 x^{\beta'} = -\partial_{\alpha} x^{\alpha'} \partial_{\beta} x^{\beta'} \partial_{\alpha'\beta'}^2 x^{\gamma} \partial_{\gamma} x^{\beta'}$$

закон преобразования (6.5) можно переписать в эквивалентной форме (5.32), отличающейся вторым слагаемым.

Закон преобразования компонент аффинной связности отличается от тензорного закона наличием неоднородных слагаемых в (5.32) и (6.5), которые содержат вторые производные от функций перехода. Если ограничить класс допустимых преобразований координат аффинными (линейными неоднородными), то аффинная связность будет преобразовываться как тензор.

Сами по себе компоненты аффинной связности не являются компонентами тензора, однако они позволяют строить новые тензорные поля из заданных с помощью ковариантного дифференцирования. Собственно, название “ковариантное” и отражает это обстоятельство. Нетрудно проверить, что ковариантная производная от 1-формы (6.3) является тензорным полем второго ранга типа (0, 2).

Ковариантное дифференцирование было продолжено на тензоры произвольного ранга с помощью формул (5.30). В компонентах это выглядит следующим образом. Помимо обычной частной производной в ковариантную производную входят слагаемые с компонентами аффинной связности со знаком плюс для каждого контравариантного и минус для каждого ковариантного индекса:

$$\begin{aligned} \nabla_{\alpha} K_{\beta_1 \dots \beta_s}^{\gamma_1 \dots \gamma_r} &= \partial_{\alpha} K_{\beta_1 \dots \beta_s}^{\gamma_1 \dots \gamma_r} - \Gamma_{\alpha\beta_1}^{\delta} K_{\delta\beta_2 \dots \beta_s}^{\gamma_1 \dots \gamma_r} - \dots - \Gamma_{\alpha\beta_s}^{\delta} K_{\beta_1 \dots \beta_{s-1} \delta}^{\gamma_1 \dots \gamma_r} + \\ &+ K_{\beta_1 \dots \beta_s}^{\delta\gamma_2 \dots \gamma_r} \Gamma_{\alpha\delta}^{\gamma_1} + \dots + K_{\beta_1 \dots \beta_s}^{\gamma_1 \dots \gamma_{r-1} \delta} \Gamma_{\alpha\delta}^{\gamma_r}. \end{aligned}$$

Пример 6.1.1. Для тензора второго ранга типа $(1, 1)$ ковариантная производная имеет вид

$$\nabla_\alpha K_\beta^\gamma = \partial_\alpha K_\beta^\gamma - \Gamma_{\alpha\beta}^\delta K_\delta^\gamma + \Gamma_{\alpha\delta}^\gamma K_\beta^\delta. \quad (6.6)$$

В частном случае, ковариантная производная от символа Кронекера (2.57) тождественно равна нулю

$$\nabla_\alpha \delta_\beta^\gamma = \partial_\alpha \delta_\beta^\gamma + \delta_\beta^\delta \Gamma_{\alpha\delta}^\gamma - \Gamma_{\alpha\beta}^\delta \delta_\delta^\gamma = 0. \quad (6.7)$$

То есть символ Кронекера ковариантно постоянен (параллелен) относительно произвольной аффинной связности. $\square$

Можно показать с помощью прямых вычислений (если доказательство в общем виде, которое было дано ранее, кого то не убеждает), что определенная таким образом ковариантная производная от тензора произвольного типа (r, s) действительно дает тензор типа $(r, s + 1)$.

Поскольку произвольный тензор типа (r, s) является полилинейным отображением (2.59), то вид ковариантной производной от тензорного поля в компонентах однозначно определяется четырьмя условиями: видом ковариантной производной от функции, векторного и ковекторного полей (6.1)–(6.3) и правилом Лейбница (5.35).

Пример 6.1.2. Найдем вид ковариантной производной от тензорного поля второго ранга K_α^β . Ковариантная производная от функции $K_\gamma^\beta X^\gamma A_\beta \in \mathcal{C}^\infty(\mathbb{M})$, где $X^\gamma \in \mathcal{X}(\mathbb{M})$ и $A_\beta \in \Lambda_1(\mathbb{M})$ – произвольные векторное и ковекторное поля, совпадает с обычной:

$$\nabla_\alpha (K_\gamma^\beta X^\gamma A_\beta) = \partial_\alpha (K_\gamma^\beta X^\gamma A_\beta).$$

Воспользовавшись правилом Лейбница и видом ковариантных производных (6.2), (6.3), получим формулу для ковариантной производной от тензора второго ранга (6.6). $\square$

Ковариантная производная коммутирует с произвольным свертыванием:

$$\nabla_\alpha (\delta_{\alpha_i}^{\beta_j} K_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}) = \delta_{\alpha_i}^{\beta_j} \nabla_\alpha K_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}, \quad 1 \leq i \leq r, \quad 1 \leq j \leq s.$$

что следует из ковариантного постоянства символа Кронекера (6.7). В частном случае,

$$\partial_\alpha (X^\beta A_\beta) = \nabla_\alpha (X^\beta A_\beta) = (\nabla_\alpha X^\beta) A_\beta + X^\beta (\nabla_\alpha A_\beta).$$

Определение. Произвольному векторному полю можно поставить в соответствие бесконечно малую величину

$$DX^\alpha := dx^\beta \nabla_\beta X^\alpha = dx^\beta (\partial_\beta X^\alpha + \Gamma_{\beta\gamma}^\alpha X^\gamma), \quad (6.8)$$

которая называется *ковариантным дифференциалом*. Аналогично определяется ковариантный дифференциал для произвольного тензорного поля. $\square$

Для функции ковариантный дифференциал совпадает с обычным.

Определение. Рассмотрим произвольную дифференцируемую кривую $\gamma = \{x^\alpha(t)\}$ в римановом пространстве с положительно определенной метрикой. Тогда длина дуги кривой отлична от нуля

$$ds := dt \sqrt{\dot{x}^2} \neq 0, \quad \text{где} \quad \dot{x}^2 := \dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}. \quad (6.9)$$

Выражение

$$\frac{DX^\alpha}{ds} := u^\beta \nabla_\beta X^\alpha, \quad (6.10)$$

где

$$u^\alpha := \frac{\dot{x}^\alpha}{\sqrt{\dot{x}^2}}, \quad u^2 = 1, \quad (6.11)$$

– единичный касательный вектор к кривой, называется *ковариантной производной вектора X вдоль кривой γ* . $\square$

Если метрика имеет лоренцеву сигнатуру, то ковариантную производную можно определить либо вдоль времениподобных, либо вдоль пространственноподобных кривых. В последнем случае $\dot{x}^2$ заменяется на $-\dot{x}^2$ под знаком корня.

Аналогично определяется ковариантная производная вдоль кривой от произвольного тензорного поля.

Пусть на многообразии задано векторное поле $X \in \mathcal{X}(\mathbb{M})$. Тогда *ковариантная производная вдоль векторного поля X* от тензорного поля $K_{\alpha_1 \dots \alpha_r}{}^{\beta_1 \dots \beta_s}$ типа (r, s) дает тензорное поле того же типа с компонентами

$$(\nabla_X K)_{\alpha_1 \dots \alpha_r}{}^{\beta_1 \dots \beta_s} := X^\gamma \nabla_\gamma K_{\alpha_1 \dots \alpha_r}{}^{\beta_1 \dots \beta_s}. \quad (6.12)$$

Пример 6.1.3. В частности, для двух векторных полей справедливо равенство

$$(\nabla_X Y)^\alpha = X^\beta \nabla_\beta Y^\alpha. \quad \square \quad (6.13)$$

Рассмотренная выше ковариантная производная от тензора вдоль кривой является частным случаем ковариантной производной вдоль единичного векторного поля, которое касается кривой. Ковариантная производная вдоль кривой определена только в тех точках многообразия $\mathbb{M}$, через которые проходит кривая.

Наличие неоднородного слагаемого в законе преобразования компонент аффинной связности (6.5) или (5.32) не позволяет их складывать как тензорные поля и умножать даже на числа. Например, если на многообразии задано две аффинные связности с компонентами $\Gamma_{1\alpha\beta}{}^\gamma$ и $\Gamma_{2\alpha\beta}{}^\gamma$, то их сумма $\Gamma_1 + \Gamma_2$ в общем случае связностью не является. В то же время нетрудно проверить, что сумма

$$\Gamma_{\alpha\beta}{}^\gamma = f\Gamma_{1\alpha\beta}{}^\gamma + (1-f)\Gamma_{2\alpha\beta}{}^\gamma, \quad (6.14)$$

где $f \in \mathcal{C}^1(\mathbb{M})$, задает компоненты некоторой аффинной связности.

Определение. *Вариацией аффинной связности $\delta\Gamma_{\alpha\beta}{}^\gamma$ называется разность двух связностей, заданных на многообразии $\mathbb{M}$.* $\square$

Вариация связности является тензорным полем типа $(1, 2)$, т.к. неоднородные слагаемые сокращаются.

6.2 Кручение и неметричность

Продолжим локальное изучение аффинной связности. В общем случае компоненты связности $\Gamma_{\alpha\beta}{}^\gamma$ никакой симметрии по индексам не имеют и никак не связаны с метрикой $g_{\alpha\beta}$, поскольку эти понятия определяют на многообразии $\mathbb{M}$ разные геометрические операции. А именно, метрика многообразия определяет в каждой точке

скалярное произведение векторов из касательного пространства, а аффинная связность определяет ковариантное дифференцирование и параллельный перенос тензоров. Геометрия на многообразии M определяется метрикой и аффинной связностью. Будем говорить, что на M задана *аффинная геометрия*, если заданы достаточно гладкие метрика и аффинная связность, т.е. задано три объекта (M, g, Γ) .

Замечание. В общем случае метрика и аффинная связность задаются произвольным образом и являются совершенно независимыми геометрическими объектами. Поэтому при построении физических моделей их можно рассматривать как независимые поля, имеющие разную физическую интерпретацию. В настоящее время принято считать, что метрика описывает гравитационное взаимодействие. Физический смысл аффинной связности пока неясен. Это связано с тем, что физическая интерпретация связности зависит от конкретной модели. Соответствующие модели сложны с математической точки зрения и в настоящее время изучены недостаточно хорошо. $\square$

По определению, кручение многообразия в локальной системе координат равно антисимметричной части аффинной связности (5.38)

$$T_{\alpha\beta}{}^{\gamma} := \Gamma_{\alpha\beta}{}^{\gamma} - \Gamma_{\beta\alpha}{}^{\gamma}. \quad (6.15)$$

Из закона преобразования связности (5.32) следует, что кручение является тензорным полем типа $(1, 2)$.

Определение. Если на многообразии задана аффинная геометрия, то можно построить *тензор неметричности* $Q_{\alpha\beta\gamma}$. Он равен ковариантной производной от метрики:

$$-Q_{\alpha\beta\gamma} := \nabla_{\alpha}g_{\beta\gamma} = \partial_{\alpha}g_{\beta\gamma} - \Gamma_{\alpha\beta}{}^{\delta}g_{\delta\gamma} - \Gamma_{\alpha\gamma}{}^{\delta}g_{\beta\delta}. \quad \square \quad (6.16)$$

Тензор неметричности, по построению, симметричен относительно перестановки двух последних индексов:

$$Q_{\alpha\beta\gamma} = Q_{\alpha\gamma\beta}.$$

Заметим, что для определения неметричности необходимы оба объекта: и метрика, и связность.

Таким образом по заданной метрике и аффинной связности построено два тензорных поля: кручение и тензор неметричности. Докажем, что по заданной метрике, кручению и тензору неметричности можно однозначно восстановить соответствующую аффинную связность. Уравнение (6.16) всегда можно решить относительно связности Γ . Действительно, линейная комбинация

$$\nabla_{\alpha}g_{\beta\gamma} + \nabla_{\beta}g_{\gamma\alpha} - \nabla_{\gamma}g_{\alpha\beta}$$

приводит к следующему общему решению для аффинной связности с опущенным верхним индексом

$$\begin{aligned} \Gamma_{\alpha\beta\gamma} := \Gamma_{\alpha\beta}{}^{\delta}g_{\delta\gamma} = & \frac{1}{2}(\partial_{\alpha}g_{\beta\gamma} + \partial_{\beta}g_{\gamma\alpha} - \partial_{\gamma}g_{\alpha\beta}) + \frac{1}{2}(T_{\alpha\beta\gamma} - T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta}) + \\ & + \frac{1}{2}(Q_{\alpha\beta\gamma} + Q_{\beta\gamma\alpha} - Q_{\gamma\alpha\beta}). \end{aligned} \quad (6.17)$$

Правая часть этого равенства симметрична по индексам α и β за исключением одного слагаемого, $\frac{1}{2}T_{\alpha\beta\gamma}$, что согласуется с определением тензора кручения (6.15). Заметим, что слагаемые с метрикой и тензором неметричности в (6.17) имеют одинаковый порядок индексов и знаков, а слагаемые с кручением при том же порядке индексов отличаются знаками. Таким образом, справедливо

Предложение 6.2.1. Для того, чтобы на многообразии M задать аффинную геометрию (g, Γ) , необходимо и достаточно задать три тензорных поля: метрику g , кручение T и неметричность Q .

Подчеркнем еще раз, что все три объекта можно задать совершенно независимым образом, и в моделях математической физики их можно рассматривать как независимые динамические переменные.

Определение. Второе слагаемое в правой части равенства (6.17)

$$K_{\alpha\beta\gamma} := \frac{1}{2}(T_{\alpha\beta\gamma} - T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta}) \quad (6.18)$$

называется тензором *кокручения*. $\square$

Заметим, что уравнение (6.16) симметрично по индексам β, γ , т.е. число алгебраических уравнений в точности равно числу компонент симметричной части связности $\Gamma_{\beta\gamma}^\alpha$. Поэтому весь произвол в решении этой системы уравнений определяется тензором кручения.

Вычислив ковариантную производную от тождества $g^{\alpha\beta}g_{\beta\gamma} = \delta_\gamma^\alpha$, находим, что ковариантная производная от обратной метрики,

$$\nabla_\alpha g^{\beta\gamma} = Q_\alpha^{\beta\gamma}, \quad Q_\alpha^{\beta\gamma} := g^{\beta\delta}g^{\gamma\epsilon}Q_{\alpha\delta\epsilon}, \quad (6.19)$$

отличается знаком от (6.16). Если тензор неметричности отличен от нуля, то подъем и опускание индексов не коммутирует с ковариантным дифференцированием.

Пример 6.2.1. $\nabla_\alpha X_\beta = (\nabla_\alpha X^\gamma)g_{\gamma\beta} - Q_{\alpha\gamma\beta}X^\gamma.$ $\square$

Для задания аффинной геометрии на многообразии необходимо задать метрику и аффинную связность. Рассмотрим частные случаи аффинной геометрии.

Определение. При попытке объединить гравитацию с электромагнетизмом Г. Вейль рассмотрел тензор неметричности специального вида [56]

$$Q_{\alpha\beta\gamma} := W_\alpha g_{\beta\gamma}, \quad (6.20)$$

где W_α – форма Вейля, отождествленная с электромагнитным потенциалом, (при этом предполагалось, что кручение тождественно равно нулю). Будем говорить, что на многообразии задана *геометрия Римана–Картана–Вейля*, если на нем задана метрика, кручение и неметричность специального вида (6.20).

Если тензор неметричности тождественно равен нулю, а метрика и кручение нетривиальны, то будем говорить, что на многообразии задана *геометрия Римана–Картана*. В этом случае из уравнения (6.17) следует, что аффинная связность однозначно определяется метрикой и кручением:

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}(\partial_\alpha g_{\beta\gamma} + \partial_\beta g_{\alpha\gamma} - \partial_\gamma g_{\alpha\beta}) + \frac{1}{2}(T_{\alpha\beta\gamma} - T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta}). \quad (6.21)$$

Такую связность называют *метрической*, поскольку ковариантная производная от метрики тождественно равна нулю,

$$\nabla_\alpha g_{\beta\gamma} = \partial_\alpha g_{\beta\gamma} - \Gamma_{\alpha\beta}^\delta g_{\delta\gamma} - \Gamma_{\alpha\gamma}^\delta g_{\beta\delta} = 0. \quad (6.22)$$

Это равенство называют *условием метричности*, и оно обеспечивает коммутативность ковариантного дифференцирования с опусканием и подъемом индексов. Условие метричности (6.22) эквивалентно условию $\nabla_\alpha g^{\beta\gamma} = 0$ в силу уравнения (6.19).

Если тензор кручения равен нулю, $T_{\alpha\beta}{}^\gamma = 0$, а неметричность имеет вид (6.20), то будем говорить, что задана геометрия *Римана–Вейля*. В геометрии Римана–Вейля выражение для аффинной связности можно записать в следующем виде:

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}[(\partial_\alpha + W_\alpha)g_{\beta\gamma} + (\partial_\beta + W_\beta)g_{\alpha\gamma} - (\partial_\gamma + W_\gamma)g_{\alpha\beta}]. \quad (6.23)$$

При этом она симметрична по двум первым индексам.

Если и тензор неметричности, и кручение тождественно равны нулю, а метрика нетривиальна и положительно определена, то будем говорить, что на многообразии задана *геометрия Римана*. В этом случае метрическая связность также симметрична по двум первым индексам и однозначно определяется метрикой:

$$\tilde{\Gamma}_{\alpha\beta}{}^\gamma = \frac{1}{2}g^{\gamma\delta}(\partial_\alpha g_{\beta\delta} + \partial_\beta g_{\alpha\delta} - \partial_\delta g_{\alpha\beta}). \quad (6.24)$$

Эта связность называется *связностью Леви–Чивиты* или *символами Кристоффеля*. Заметим, что в определении метрики мы требовали невырожденность матрицы $g_{\alpha\beta}$. Из выражения для символов Кристоффеля видно, что это условие необходимо, в частности, для того, чтобы метрика определяла связность Леви–Чивиты. Если метрика не является положительно определенной, то геометрия называется *псевдоримановой*. Для связности при этом сохраняются прежние названия. $\square$

Замечание. Если на многообразии задана аффинная геометрия общего вида (M, g, Γ) , то на нем определены две связности: аффинная связность Γ и связность Леви–Чивиты $\tilde{\Gamma}$, поскольку задана метрика. В такой ситуации над связностью Леви–Чивиты и построенных с ее помощью геометрических объектах мы будем писать знак тильды. $\square$

Пример 6.2.2. Связность Леви–Чивиты имеет наглядный геометрический смысл для двумерных поверхностей, изометрически вложенных в трехмерное евклидово пространство. Рассмотрим два вектора X_x и Y_y , касательных к поверхности в близких точках x и y . Леви–Чивита [57] предложил считать вектор Y_y параллельным вектору X_x , если его проекция в евклидовом пространстве на касательную плоскость в точке x параллельна вектору X_x . Можно проверить, что такая связность согласована с метрикой, индуцированной вложением, и имеет нулевое кручение. $\square$

В римановой геометрии ковариантная производная определяется компонентами связности с двумя нижними и одним верхним индексом, $\tilde{\Gamma}_{\alpha\beta}{}^\gamma := \tilde{\Gamma}_{\alpha\beta\delta}g^{\gamma\delta}$ и поэтому одинакова для двух метрик, отличающихся постоянным множителем.

Символы Кристоффеля (6.24) симметричны по первым двум индексам. Это свойство выполняется в голономном (координатном) базисе. В неголономном базисе эта симметрия нарушается (см. раздел 6.9).

Из выражения для символов Кристоффеля (6.24) или условия метричности (6.22) следует равенство

$$\partial_\alpha g_{\beta\gamma} = \tilde{\Gamma}_{\alpha\beta\gamma} + \tilde{\Gamma}_{\alpha\gamma\beta}. \quad (6.25)$$

Это доказывает следующее

Предложение 6.2.2. Для того, чтобы символы Кристоффеля в некоторой системе координат были равны нулю, необходимо и достаточно, чтобы в этой системе координат компоненты метрики были постоянны.

Поскольку символы Кристоффеля не являются компонентами тензора, то в другой системе координат они могут быть нетривиальны.

Пример 6.2.3. Символы Кристоффеля для евклидова пространства в декартовой системе координат равны нулю, но, скажем, в сферической или цилиндрической системе координат они отличны от нуля (см. раздел 7). $\square$

Определение. В случае, когда тензор неметричности и кручение тождественно равны нулю, и при этом в окрестности любой точки существует система координат, в которой метрика является единичной диагональной матрицей, и, следовательно, символы Кристоффеля равны нулю, геометрия называется *локально евклидовой* (или *псевдоевклидовой*, если метрика не является положительно определенной). Соответствующая система координат называется *декартовой*. $\square$

Пример 6.2.4. Рассмотрим двумерный тор, вложенный в трехмерное евклидово пространство. Если предположить, что кручение и неметричность на торе равны нулю, а метрика индуцирована вложением, то тензор кривизны тора равен нулю. В этом случае тор представляет собой нетривиальное локально евклидово многообразие. $\square$

Нетрудно посчитать число независимых компонент у связности, тензора кручения и неметричности:

$$[\Gamma_{\alpha\beta}^{\gamma}] = n^3, \quad [T_{\alpha\beta}^{\gamma}] = \frac{n^2(n-1)}{2}, \quad [Q_{\alpha\beta\gamma}] = \frac{n^2(n+1)}{2}.$$

Отсюда следует, что суммарное число независимых компонент кручения и неметричности равно числу компонент аффинной связности.

6.3 Ковариантная производная тензорных плотностей

Ковариантная производная тензорного поля естественным образом обобщается на тензорные плотности произвольной степени p . Для определенности рассмотрим тензорную плотность X_{α}^{β} степени p и типа $(1, 1)$. По определению, при преобразовании координат она преобразуется по правилу

$$X_{\alpha'}^{\beta'} = J^p \frac{dx^{\alpha}}{dx^{\alpha'}} X_{\alpha}^{\beta} \frac{dx^{\beta'}}{dx^{\beta}},$$

где $J := \det \partial_{\alpha} x^{\alpha'}$ – якобиан преобразования (1.69). Потребуем, чтобы ковариантная производная от тензорной плотности была тензорной плотностью той же степени,

$$\nabla_{\alpha'} X_{\beta'}^{\gamma'} = J^p \frac{dx^{\alpha}}{dx^{\alpha'}} \frac{dx^{\beta}}{dx^{\beta'}} \nabla_{\alpha} X_{\beta}^{\gamma} \frac{dx^{\gamma'}}{dx^{\gamma}}. \quad (6.26)$$

Отсюда следует, что ковариантная производная от тензорной плотности имеет вид

$$\nabla_{\alpha} X_{\beta}^{\gamma} = \partial_{\alpha} X_{\beta}^{\gamma} - \Gamma_{\alpha\beta}^{\delta} X_{\delta}^{\gamma} + X_{\beta}^{\delta} \Gamma_{\alpha\delta}^{\gamma} + p \Gamma_{\alpha} X_{\beta}^{\gamma}, \quad (6.27)$$

где $\Gamma_\alpha := \Gamma_{\alpha\beta}^\beta$ – след аффинной связности. Сравнивая эту ковариантную производную с ковариантной производной (6.6) от тензорного поля типа $(1, 1)$, находим, что различие состоит в появлении дополнительного слагаемого, пропорционального степени тензорной плотности и следу аффинной связности. Чтобы проверить выполнение закона преобразования (6.26) заметим, что в законе преобразования следа аффинной связности (6.47) содержится неоднородное слагаемое, которое компенсирует производную от якобиана

$$\partial_\alpha J^p = p J^p \frac{dx^\beta}{dx^{\beta'}} \partial_\alpha \left(\frac{dx^{\beta'}}{dx^\beta} \right).$$

Отсюда следует, что ковариантная производная от тензорной плотности произвольного типа отличается от ковариантной производной соответствующего тензора одним дополнительным слагаемым, пропорциональным степени тензорной плотности и следу аффинной связности.

Определение. Параллельный перенос для тензорных плотностей определяется так же, как и для тензоров. Мы говорим, что тензорная плотность *параллельно переносится* вдоль заданной кривой, если ее ковариантная производная вдоль этой кривой равна нулю. $\square$

Ковариантная производная от линейной комбинации тензорных плотностей X и Y одинакового ранга и степени равна линейной комбинации ковариантных производных:

$$\nabla_\alpha(aX + bY) = a\nabla_\alpha X + b\nabla_\alpha Y, \quad a, b \in \mathbb{R}.$$

Естественным образом определяется тензорное произведение $X \otimes Y$ тензорных плотностей произвольных рангов и степеней. При этом степени тензорных плотностей складываются. Нетрудно проверить, что для ковариантного дифференцирования справедливо правило Лейбница:

$$\nabla_\alpha(X \otimes Y) = (\nabla_\alpha X) \otimes Y + X \otimes (\nabla_\alpha Y).$$

Так же, как и для тензоров, свертка по индексам тензорных плотностей перестановочна с операцией ковариантного дифференцирования.

Пример 6.3.1. Рассмотрим частные случаи. Определители метрики g и элемент объема $\sqrt{|g|} = \det e_\alpha^a$ являются тензорными плотностями степеней -2 и -1 , соответственно. Используя выражение для следа аффинной связности (6.48), получаем следующие равенства:

$$\nabla_\alpha g = \partial_\alpha g - 2\Gamma_\alpha g = -Q_\alpha g, \quad (6.28)$$

$$\nabla_\alpha \sqrt{|g|} = \partial_\alpha \sqrt{|g|} - \Gamma_\alpha \sqrt{|g|} = -\frac{1}{2} Q_\alpha \sqrt{|g|}, \quad (6.29)$$

где

$$Q_\alpha := Q_{\alpha\beta}^\beta \quad (6.30)$$

– след тензора неметричности. Отсюда следует, что для метрической связности определители метрики и репера ковариантно постоянны. То есть в геометрии Римана–Картана и, в частности, в римановой геометрии ковариантные производные от определителя метрики и элемента объема равны нулю:

$$\tilde{\nabla}_\alpha g = 0, \quad \tilde{\nabla}_\alpha \sqrt{|g|} = 0.$$

При этом условие метричности для связности является достаточным, но не необходимым. Из выражения (6.29) следует, что для сохранения элемента объема при параллельном переносе вдоль произвольной кривой, необходимо и достаточно, чтобы след тензора неметричности был равен нулю. В геометрии Римана–Картана–Вейля справедлива формула

$$\nabla_\alpha \sqrt{|g|} = -\frac{n}{2} W_\alpha \sqrt{|g|},$$

где n – размерность многообразия. То есть форма Вейля в этом случае определяет изменение элемента объема при параллельном переносе. $\square$

6.4 Параллельный перенос

В разделе 5.2 был определен параллельный перенос векторов. С помощью аффинной связности на многообразии $\mathbb{M}$ можно определить *параллельный перенос* касательных векторов, а также тензоров произвольного ранга вдоль кривой.

Определение. Пусть дифференцируемая кривая $\gamma = x(t) = \{x^\alpha(t)\}$, $t \in [0, 1]$ соединяет две точки многообразия $p, q \in \mathbb{M}$: $x(0) = p$, $x(1) = q$. Касательный вектор к кривой (вектор скорости) имеет компоненты $u^\alpha := \dot{x}^\alpha$ и предполагается отличным от нуля. Инвариантное условие параллельного переноса (5.25),

$$\nabla_u X = 0, \quad (6.31)$$

в компонентах (5.26) принимает вид

$$\dot{X}^\alpha = -\dot{x}^\beta X^\gamma \Gamma_{\beta\gamma}^\alpha, \quad (6.32)$$

где все поля рассматриваются как функции от t . Это – система линейных обыкновенных дифференциальных уравнений на компоненты векторного поля $X^\alpha(t)$, которая решается с некоторым начальным условием $X^\alpha(0) := X_0^\alpha$. Если функции $\dot{x}^\beta \Gamma_{\beta\gamma}^\alpha$ непрерывны по t , то решение этой задачи существует и единственно для всех значений $t \in [0, 1]$ (см., например, [1], § 27). Решение системы уравнений (6.32) с начальным условием $X^\alpha(0) := X_0^\alpha$ называется *параллельным переносом* вектора X_0 в точку $x(t)$ вдоль кривой γ . $\square$

В общем случае, если правая часть системы линейных уравнений (6.32) непрерывна для всех $t \in \mathbb{R}$, то решение задачи Коши продолжается на всю вещественную ось (см., например, [1], § 27).

Замечание. В отличие от параллельного переноса, уравнения геодезических являются нелинейными, и поэтому общие теоремы для систем обыкновенных дифференциальных уравнений гарантируют существование и единственность решений только в некоторой достаточно малой окрестности начальной точки $t = 0$ при надлежащих условиях на правую часть. Это приводит к проблеме продолжения геодезических. $\square$

Условие параллельного переноса (6.32) можно переписать в интегральной форме

$$X^\alpha(t) = X_0^\alpha - \int_0^t ds \dot{x}^\beta X^\gamma \Gamma_{\beta\gamma}^\alpha, \quad (6.33)$$

где все функции в подынтегральном выражении рассматриваются, как функции от параметра s вдоль кривой и точка обозначает дифференцирование по s . При $t = 1$ получим компоненты вектора в конечной точке q .

Предложение 6.4.1. *Результат параллельного переноса вектора из точки p в точку q не зависит ни от параметризации кривой, ни от выбора системы координат.*

Доказательство. Независимость от параметризации кривой очевидна.

В новой штрихованной системе координат справедливо равенство

$$X^{\alpha'}(t) = X_0^{\alpha'} - \int_0^t ds \dot{x}^{\beta'} X^{\gamma'} \Gamma_{\beta' \gamma'}^{\alpha'}.$$

Это интегральное уравнение связано с уравнением (6.33) преобразованием координат. Чтобы доказать независимость параллельного переноса, представим неоднородное слагаемое в преобразовании компонент аффинной связности (6.5), умноженное на $\dot{x}^{\beta'} X^{\gamma'}$, в виде

$$\dot{x}^{\beta} X^{\gamma} \partial_{\beta\gamma}^2 x^{\alpha'} = X^{\gamma} \partial_t (\partial_{\gamma} x^{\alpha'})$$

и проинтегрируем по частям. $\square$

В то же время результат параллельного переноса вектора из точки p в точку q в общем случае зависит от кривой γ , соединяющей эти точки.

Параллельный перенос вектора вдоль кривой можно естественным образом обобщить на кусочно гладкие кривые, как последовательный параллельный перенос от одного излома к другому.

Аналогично определяется параллельный перенос 1-форм и тензоров произвольного ранга.

При параллельном переносе вектора из точки x в близкую точку $x + dx$, где инфинитезимальное приращение dx^{α} мы рассматриваем, как касательный вектор к кривой в точке x , компоненты вектора получают приращение

$$\delta X^{\alpha} = -dx^{\beta} X^{\gamma} \Gamma_{\beta\gamma}^{\alpha}. \quad (6.34)$$

Эти приращения линейны по дифференциалам dx^{α} и самому векторному полю. Формула (6.34) позволяет дать следующую интерпретацию ковариантной производной (6.2). Пусть на многообразии задано произвольное векторное поле $X(x)$. Чтобы получить ковариантную производную от него в точке x , необходимо взять значение векторного поля в точке $x + dx$, вычесть из него результат параллельного переноса вектора из точки x в точку $x + dx$ и разделить эту разность на dx .

Пример 6.4.1. При параллельном переносе скаляра f получаем постоянное скалярное поле вдоль кривой $\delta f = 0$. Результат параллельного переноса скаляра для любой аффинной связности не зависит от пути переноса $\square$

Пример 6.4.2. При параллельном переносе 1-форма получает приращение

$$\delta A_{\alpha} = dx^{\beta} \Gamma_{\beta\alpha}^{\gamma} A_{\gamma}, \quad (6.35)$$

которое отличается знаком от приращения вектора (6.34). $\square$

Пример 6.4.3. При параллельном переносе тензора типа $(1, 1)$ он получает приращение

$$\delta T^{\alpha}_{\beta} = dx^{\gamma} (-\Gamma_{\gamma\delta}^{\alpha} T^{\delta}_{\beta} + \Gamma_{\gamma\beta}^{\delta} T^{\alpha}_{\delta}).$$

Аналогично для тензора произвольного ранга в правой части будем иметь по одному слагаемому со знаком минус и плюс для каждого контра- и ковариантного индекса, соответственно. $\square$

Пример 6.4.4. В геометрии Римана–Картана, по определению, ковариантная производная от метрики равна нулю. Поэтому можно считать, что метрика получается в результате параллельного переноса симметричной невырожденной матрицы из произвольной точки многообразия. При этом результат параллельного переноса заданной матрицы не зависит от пути для любой метрической связности. Это будет ясно из дальнейшего рассмотрения. $\square$

Пусть на многообразии $\mathbb{M}$ помимо аффинной связности задана метрика. Рассмотрим, как меняется скалярное произведение (X, Y) двух векторов при параллельном переносе вдоль произвольной кривой γ .

Предложение 6.4.2. *Зависимость скалярного произведения (X, Y) двух векторов, которые параллельно переносятся вдоль γ , от точки кривой определяется только тензором неметричности:*

$$\partial_u(X, Y) = \nabla_u(X^\alpha Y^\beta g_{\alpha\beta}) = X^\alpha Y^\beta \nabla_u g_{\alpha\beta} = -u^\gamma X^\alpha Y^\beta Q_{\gamma\alpha\beta}. \quad (6.36)$$

Отсюда следует, что в геометрии Римана–Картана ($Q = 0$) скалярное произведение двух векторов при параллельном переносе вдоль произвольной кривой сохраняется. В частности, квадрат вектора скорости u^2 кривой γ в геометрии Римана–Картана постоянен вдоль нее.

Доказательство. Следует из определения тензора неметричности (6.16). $\square$

Следствие. В римановой геометрии и геометрии Римана–Картана длины векторов и углы между ними сохраняются при параллельном переносе вдоль произвольной кривой γ .

Задание аффинной связности позволяет сравнивать компоненты тензоров в бесконечно близких точках, причем делать это ковариантным образом. Трудность сравнения тензоров в различных точках связана с тем, что при преобразовании координат тензоры в разных точках преобразуются по-разному, и их сравнение (сложение компонент) теряет всякий смысл.

Ниже мы дадим геометрическую интерпретацию тензора кривизны, и с этой целью выполним следующее построение.

Выберем точку многообразия $p = \{p^\alpha\} \in \mathbb{M}$, $\dim \mathbb{M} = n$, и замкнутую кусочно гладкую кривую γ_p с началом и концом в точке p . Возьмем произвольный вектор X_p в точке p и перенесем его параллельно вдоль кривой γ_p . В результате получим новый вектор $X_p + \Delta X_p$, который может отличаться от исходного. Рассмотрим класс замкнутых кривых $\gamma_p = \{x^\alpha(t)\}$ малой “длины”, для которых выполнено неравенство

$$\oint_{\gamma_p} dt \sqrt{(\dot{x}^1)^2 + \dots + (\dot{x}^n)^2} < \epsilon \ll 1. \quad (6.37)$$

Это выражение нековариантно, т.е. мы работаем в фиксированной системе координат некоторой координатной окрестности. Тогда все рассматриваемые кривые заведомо лежат в малой окрестности точки $p \in \mathbb{U}_{p,\epsilon} \subset \mathbb{M}$, координаты точек которой удовлетворяют неравенствам:

$$\mathbb{U}_{p,\epsilon} := \{x \in \mathbb{M} : |x^\alpha - p^\alpha| < \epsilon, \alpha = 1, \dots, n\}.$$

Эта окрестность $\mathbb{U}_{p,\epsilon}$ диффеоморфна шару и поэтому односвязна. Рассмотрим двумерную поверхность $\mathbb{S}_\gamma \subset \mathbb{M}$ в многообразии $\mathbb{M}$, с границей γ_p и целиком лежащую в $\mathbb{U}_{p,\epsilon}$, которая задана параметрически: $\{x^\alpha(u, v)\}$, где u, v – координаты на поверхности. Тогда справедлива

Теорема 6.4.1 (Геометрический смысл кривизны). Пусть на многообразии $\mathbb{M}$ в некоторой координатной окрестности задана аффинная связность $\Gamma_{\alpha\beta}^{\gamma}$ с тензором кривизны $R_{\alpha\beta\gamma}^{\delta}$. Перенесем вектор X_p вдоль замкнутой кривой γ_p , для которой выполнено неравенство (6.37). Тогда компоненты вектора получают приращение порядка не выше ϵ^2 :

$$\Delta X_p^{\delta} = -X_p^{\gamma} R_{\alpha\beta\gamma}^{\delta}(p) \int_{\mathbb{S}_{\gamma}} du \wedge dv \frac{\partial x^{\alpha}}{\partial u} \frac{\partial x^{\beta}}{\partial v} + o(\epsilon^2), \quad (6.38)$$

где $\mathbb{S}_{\gamma} \subset \mathbb{U}_{p,\epsilon}$ – поверхность, ограниченная кривой γ_p .

Доказательство. При параллельном переносе вектора вдоль замкнутого пути приращение его компонент задается интегралом (6.33)

$$\Delta X_p^{\alpha} = - \oint_{\gamma_p} dt \dot{x}^{\beta} X_{\parallel}^{\gamma} \Gamma_{\beta\gamma}^{\alpha} = - \oint_{\gamma_p} dx^{\beta} X_{\parallel}^{\gamma} \Gamma_{\beta\gamma}^{\alpha}, \quad (6.39)$$

где $X_{\parallel}^{\gamma}(t)$ – результат параллельного переноса вектора X_p вдоль кривой γ до точки $x(t)$. Разложим подынтегральную функцию в ряд Тейлора:

$$X_{\parallel}^{\gamma} \Gamma_{\beta\gamma}^{\alpha} = (X_{\parallel}^{\gamma} \Gamma_{\beta\gamma}^{\alpha})_p + [\partial_{\delta}(X_{\parallel}^{\gamma} \Gamma_{\beta\gamma}^{\alpha})]_p (x^{\delta} - p^{\delta}) + \dots$$

Интеграл от первого слагаемого равен нулю, т.к. кривая замкнута, и, поэтому порядок приращения вектора при параллельном переносе не может превышать ϵ^2 . Это значит, что при вычислении интеграла в (6.39) с точностью ϵ^2 в подынтегральном выражении достаточно учитывать члены порядка ϵ .

Разнесем вектор X_p из точки p на всю окрестность $\mathbb{U}_{p,\epsilon}$ с помощью следующего соотношения

$$X^{\alpha}(x) := X_p^{\alpha} - (x^{\beta} - p^{\beta}) X_p^{\gamma} \Gamma_{\beta\gamma}^{\alpha}(p). \quad (6.40)$$

Это – параллельный перенос вектора X_p на окрестность $\mathbb{U}_{p,\epsilon}$ с точностью ϵ . С указанной точностью параллельный перенос вектора X_p не зависит от пути, т.к. приращение компонент вектора по замкнутому пути порядка ϵ^2 . В результате мы получаем гладкое векторное поле $X(x)$ на $\mathbb{U}_{p,\epsilon}$. Поэтому в подынтегральном выражении (6.39) можно заменить векторное поле вдоль кривой $X_{\parallel}^{\alpha}$, которое в общем случае не является непрерывным, на гладкое векторное поле X . Тогда можно воспользоваться формулой Грина (3.89):

$$\oint_{\gamma_p} dt \dot{x}^{\beta} X^{\gamma} \Gamma_{\beta\gamma}^{\alpha} = \int_{\mathbb{S}_{\gamma}} du \wedge dv \left[\frac{\partial}{\partial u} \left(\frac{\partial x^{\beta}}{\partial v} X^{\gamma} \Gamma_{\beta\gamma}^{\alpha} \right) - \frac{\partial}{\partial v} \left(\frac{\partial x^{\beta}}{\partial u} X^{\gamma} \Gamma_{\beta\gamma}^{\alpha} \right) \right].$$

Для векторного поля (6.40) выполнено равенство

$$\frac{\partial X^{\gamma}}{\partial u} = - \frac{\partial x^{\delta}}{\partial u} X_p^{\beta} \Gamma_{\delta\beta}^{\gamma}(p)$$

и аналогичную формулу для частной производной по v . В результате с точностью ϵ^2 получаем формулу (6.38). $\square$

Замечание. С точностью ϵ^2 приращение компонент вектора (6.38) ковариантно. По построению, это приращение не зависит от выбора поверхности $\mathbb{S}_{\gamma} \subset \mathbb{U}_{p,\epsilon}$, натянутой на контур γ_p . $\square$

Пример 6.4.5. Рассмотрим “параллелограмм” γ_p со сторонами $\epsilon_1^\alpha, \epsilon_2^\alpha < \epsilon$ и вершинами в точках $\{p^\alpha\}$, $\{p^\alpha + \epsilon_1^\alpha\}$, $\{p^\alpha + \epsilon_1^\alpha + \epsilon_2^\alpha\}$ и $\{p^\alpha + \epsilon_2^\alpha\}$. Пусть на него натянута поверхность $\mathbb{S}_\gamma = \{p^\alpha + \epsilon_1^\alpha u + \epsilon_2^\alpha v\}$; $u, v \in [0, 1]$. Тогда при обходе по “параллелограмму” вектор получит приращение

$$\Delta X_p^\delta = -\epsilon_1^\alpha \epsilon_2^\beta X_p^\gamma R_{\alpha\beta\gamma}^\delta(p),$$

что следует из (6.38). См. рис.6.1. $\square$

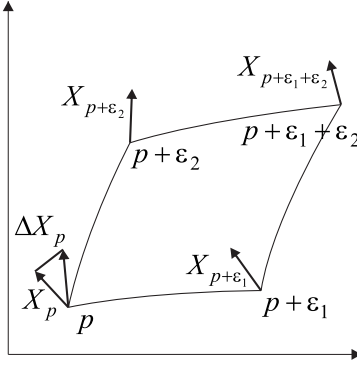


Рис. 6.1: Геометрический смысл кривизны: при параллельном переносе вектора X_p вдоль замкнутого контура он получает приращение ΔX_p .

Замечание. Теорема о геометрическом смысле кривизны не зависит от того задана на многообразии $\mathbb{M}$ метрика или нет, т.к. параллельный перенос определяется только связностью. $\square$

Пусть на многообразии $\mathbb{M}$ задана также метрика $g_{\alpha\beta}$. Тогда на поверхности $\mathbb{S}_\gamma$ индуцируется метрика $g_{ij} = \partial_i x^\alpha \partial_j x^\beta$ ($i, j = 1, 2$, где $\{u^i\} = (u, v)$). Следовательно, определена форма объема (площади поверхности) (3.79)

$$v = \frac{1}{2} du^i \wedge du^j \epsilon_{ij}.$$

Тогда интеграл в (6.38) можно переписать в виде

$$\frac{1}{2} \int_{\mathbb{S}_\gamma} du^i \wedge du^j \frac{\partial x^\alpha}{\partial u^i} \frac{\partial x^\beta}{\partial u^j} = \frac{1}{2} \int_{\mathbb{S}_\gamma} du dv \sqrt{|g|} \epsilon^{ij} \frac{\partial x^\alpha}{\partial u^i} \frac{\partial x^\beta}{\partial u^j} \text{sgn},$$

где $\text{sgn} := \text{sign}(\det g_{ij})$ и $\sqrt{|g|} := \sqrt{|\det g_{ij}|}$. С точностью ϵ^2 подынтегральное выражение можно вынести из под знака интегрирования и формула для приращения компонент вектора при параллельном переносе вдоль малого замкнутого пути (6.38) принимает вид

$$\Delta X_p^\delta = -\frac{1}{\sqrt{|g_p|}} \left(\frac{\partial x^\alpha}{\partial u} \frac{\partial x^\beta}{\partial v} \right)_p X_p^\gamma R_{\alpha\beta\gamma}^\delta(p) S, \quad (6.41)$$

где

$$S := \int_{\mathbb{S}_\gamma} du dv \sqrt{|g|}$$

— площадь поверхности $\mathbb{S}_\gamma$, которая имеет порядок ϵ^2 . Полученная формула ковариантна относительно преобразований координат на многообразии x^α и инвариантна относительно выбора координат u, v на поверхности.

Замечание. Геометрический смысл тензора кривизны лег в основу его физической интерпретации, как поверхностной плотности вектора Франка, характеризующего распределение дисклинаций в упругой среде со спиновой структурой (см. [58]). $\square$

Формулы переноса вектора вдоль малого замкнутого контура без труда обобщаются на произвольные тензорные поля.

Пример 6.4.6. Рассмотрим такой же “параллелограмм” со сторонами ϵ_1 и ϵ_2 как и в примере 6.4.5. При параллельном переносе вдоль него компоненты тензора T типа $(1, 1)$ получают приращение

$$\Delta T_p^\alpha{}_\beta = -\epsilon_1^\gamma \epsilon_2^\delta [R_{\gamma\delta\epsilon}{}^\alpha(p) T_p^\epsilon{}_\beta - R_{\gamma\delta\beta}{}^\epsilon(p) T_p^\alpha{}_\epsilon].$$

Для тензора произвольного ранга в правой части равенства будем иметь по одному слагаемому со знаками плюс и минус, соответственно, для каждого контравариантного и ковариантного индекса. $\square$

6.5 Геометрический смысл кручения

Рассмотрим многообразие $\mathbb{M}$ с заданной аффинной связностью. Пусть $\mathcal{U} \subset \mathbb{M}$ – координатная окрестность и $\Gamma_{\alpha\beta}^\gamma$ – компоненты связности. Пусть в произвольной точке $p \in \mathcal{U}$ заданы два касательных вектора $X_p = \{X_p^\alpha\}$ и $Y_p = \{Y_p^\alpha\}$, рис. 6.2. Чтобы дать

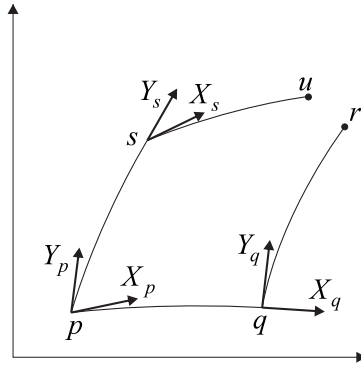


Рис. 6.2: Геометрический смысл кручения: “параллелограмм” с направляющими $X_p\epsilon$ и $Y_p\epsilon$ разомкнут.

геометрическую интерпретацию тензору кручения, произведем следующее построение. Перенесем вектор Y_p вдоль геодезической линии, которая касается вектора X_p , в близкую точку q , соответствующую каноническому параметру $t_q := t_p + \epsilon$, $\epsilon \ll 1$. Затем проведем геодезическую в точке q , которая касается вектора Y_q , и отметим на ней точку r , соответствующую параметру $t_q + \epsilon$. После этого проведем то же построение, но в обратном порядке: перенесем вектор X_p вдоль геодезической для Y_p в точку s , соответствующую параметру $t_p + \epsilon$, выпустим из нее геодезическую вдоль X_s и отметим точку u , соответствующую параметру $t_s + \epsilon$. Если кручение аффинной связности равно нулю, то точки r и u совпадут. При ненулевом кручении разность координат точек r и u имеет порядок ϵ^2 и, как будет показано, определяется тензором кручения.

Начнем построение. Рассмотрим геодезическую линию $\{x^\alpha(t)\}$ в $\mathbb{M}$, которая проходит через p и касается вектора X_p . Перенесем вектор X_p вдоль геодезической. В

результата получим векторное поле $X(t)$ на геодезической, компоненты которого, по построению, удовлетворяет уравнению

$$\nabla_{\dot{x}} X^\alpha = \frac{DX^\alpha}{dt} = \dot{x}^\beta (\partial_\beta X^\alpha + X^\gamma \Gamma_{\beta\gamma}^\alpha) = \dot{X}^\alpha + \dot{x}^\beta X^\gamma \Gamma_{\beta\gamma}^\alpha = 0, \quad (6.42)$$

где DX – ковариантный дифференциал векторного поля (6.8) и t – произвольный параметр вдоль геодезической. Тогда геодезическая будет интегральной кривой полученного векторного поля, проходящей через точку p .

Предположим, что выбран канонический параметр вдоль геодезической. Тогда геодезическая удовлетворяет уравнению

$$\ddot{x}^\alpha = -\Gamma_{\beta\gamma}^\alpha \dot{x}^\beta \dot{x}^\gamma$$

с начальными условиями:

$$\dot{x}^\alpha|_{t=t_p} = X^\alpha, \quad x^\alpha|_{t=t_p} = x_p^\alpha. \quad (6.43)$$

Решение этой задачи Коши единственно. Следовательно, канонический параметр фиксируется начальными условиями единственным образом. Для наших целей вычисления достаточно провести во втором порядке по ϵ . Ниже мы будем писать знак равенства, понимая его с точностью $o(\epsilon^2)$. В этом порядке координаты точки q равны

$$x_q^\alpha = x_p^\alpha + X_p^\alpha \epsilon + \frac{1}{2} \dot{X}_p^\alpha \epsilon^2 = x_p^\alpha + X_p^\alpha \epsilon - \frac{1}{2} X_p^\beta X_p^\gamma \Gamma_{\beta\gamma}^\alpha \epsilon^2,$$

где мы учли уравнения (6.42) и (6.43). Теперь параллельно перенесем вектор Y_p в точку q :

$$Y_q^\alpha = Y_p^\alpha - (x_q^\beta - x_p^\beta) Y_p^\gamma \Gamma_{p\beta\gamma}^\alpha = Y_p^\alpha - X_p^\beta Y_p^\gamma \Gamma_{p\beta\gamma}^\alpha \epsilon,$$

где достаточно учесть только линейный член разложения. Теперь найдем координаты точки r вдоль геодезической, выпущенной из точки q :

$$\begin{aligned} x_r^\alpha &= x_q^\alpha + Y_q^\alpha \epsilon - \frac{1}{2} Y_q^\beta Y_q^\gamma \Gamma_{q\beta\gamma}^\alpha \epsilon^2 = \\ &= x_p^\alpha + (X_p^\alpha + Y_p^\beta) \epsilon^2 - \left(\frac{1}{2} X_p^\beta X_p^\gamma + X_p^\beta Y_p^\gamma + \frac{1}{2} Y_p^\beta Y_p^\gamma \right) \Gamma_{p\beta\gamma}^\alpha \epsilon^2. \end{aligned} \quad (6.44)$$

Аналогично, координаты точки u во втором порядке по ϵ равны

$$x_u^\alpha = x_p^\alpha + (X_p^\alpha + Y_p^\beta) \epsilon^2 - \left(\frac{1}{2} X_p^\beta X_p^\gamma + Y_p^\beta X_p^\gamma + \frac{1}{2} Y_p^\beta Y_p^\gamma \right) \Gamma_{p\beta\gamma}^\alpha \epsilon^2.$$

Таким образом, получаем равенство

$$x_u^\alpha - x_r^\alpha = X_p^\beta Y_p^\gamma T_{p\beta\gamma}^\alpha \epsilon^2, \quad (6.45)$$

где $T_{p\beta\gamma}^\alpha$ – компоненты тензора кручения в точке p .

Формула (6.45) позволяет дать геометрическую интерпретацию тензору кручения: главная часть разомкнутости бесконечно малого “параллелограмма” с направляющими X_p и Y_p квадратична по ϵ и определяется тензором кручения.

Замечание. Геометрический смысл кручения лег в основу физической интерпретации кручения, как поверхностной плотности вектора Бюргерса, характеризующего дислокации в упругой среде (см. раздел 36.2). $\square$

6.6 Свойства аффинной связности

Во многих важных моделях математической физики пространство-время рассматривается, как многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, на котором задана аффинная геометрия, т.е. метрика и аффинная связность. В настоящем разделе приведены тождества, включающие аффинную связность, которые полезны для приложений при проведении вычислений. Кроме того, определены инвариантные дифференциальные операторы второго порядка и приведена формула интегрирования по частям. Все формулы настоящего раздела доказываются прямыми вычислениями.

Определим *след аффинной связности*, который получается после свертки последней пары индексов:

$$\Gamma_\alpha := \Gamma_{\alpha\beta}{}^\beta. \quad (6.46)$$

Он не является ковекторным полем, поскольку закон преобразования содержит неоднородное слагаемое:

$$\Gamma_{\alpha'} = \partial_{\alpha'} x^\alpha \Gamma_\alpha - \partial_{\alpha'} x^\alpha \partial_{\beta'} x^\beta \partial_{\alpha\beta}^2 x^{\beta'} = \partial_{\alpha'} x^\alpha \Gamma_\alpha + \partial_{\alpha'\beta'}^2 x^\alpha \partial_\alpha x^{\beta'}. \quad (6.47)$$

Из выражения (6.17) для компонент аффинной связности следует, что след аффинной связности (6.46) равен

$$\Gamma_\alpha = \tilde{\Gamma}_\alpha + \frac{1}{2} Q_\alpha, \quad (6.48)$$

где след символов Кристоффеля, $\tilde{\Gamma}_\alpha := \tilde{\Gamma}_{\alpha\beta}{}^\beta = \tilde{\Gamma}_{\beta\alpha}{}^\beta$, имеет вид

$$\tilde{\Gamma}_\alpha = \frac{1}{2} g^{\beta\gamma} \partial_\alpha g_{\beta\gamma} = -\frac{1}{2} g_{\beta\gamma} \partial_\alpha g^{\beta\gamma} = \frac{1}{2} \frac{\partial_\alpha g}{g} = \frac{\partial_\alpha \sqrt{|g|}}{\sqrt{|g|}} = e^\beta{}_b \partial_\alpha e_\beta{}^b, \quad (6.49)$$

и введен след тензора неметричности по последним индексам:

$$Q_\alpha := Q_{\alpha\beta}{}^\beta. \quad (6.50)$$

В тождествах (6.49)

$$g := \det g_{\alpha\beta}, \quad \sqrt{|g|} = \det e_\alpha{}^a.$$

Замечание. След аффинной связности не зависит от кручения. $\square$

Приведем также несколько полезных формул, справедливых в (псевдо)римановой геометрии:

$$\partial_\alpha g = g g^{\beta\gamma} \partial_\alpha g_{\beta\gamma}, \quad (6.51)$$

$$\partial_\alpha \sqrt{|g|} = \frac{1}{2} \sqrt{|g|} g^{\beta\gamma} \partial_\alpha g_{\beta\gamma} = \sqrt{|g|} \tilde{\Gamma}_{\beta\alpha}{}^\beta = \sqrt{|g|} e^\beta{}_a \partial_\alpha e_\beta{}^a, \quad (6.52)$$

$$\partial_\alpha g^{\beta\gamma} = -g^{\beta\delta} g^{\gamma\epsilon} \partial_\alpha g_{\delta\epsilon} = -g^{\beta\delta} e^\gamma{}_a \partial_\alpha e_\delta{}^a - g^{\gamma\delta} e^\beta{}_a \partial_\alpha e_\delta{}^a, \quad (6.53)$$

$$\partial_\alpha (\sqrt{|g|} g^{\beta\gamma}) = \sqrt{|g|} (g^{\beta\gamma} \tilde{\Gamma}_\alpha - g^{\beta\delta} \tilde{\Gamma}_{\alpha\delta}{}^\gamma - g^{\gamma\delta} \tilde{\Gamma}_{\alpha\delta}{}^\beta), \quad (6.54)$$

$$\partial_\beta (\sqrt{|g|} g^{\beta\alpha}) = -\sqrt{|g|} g^{\beta\gamma} \tilde{\Gamma}_{\beta\gamma}{}^\alpha, \quad (6.55)$$

$$e_\alpha{}^a \partial_\beta (\sqrt{|g|} e_\alpha{}^a) = (n-1) \sqrt{|g|} e_\alpha{}^a \partial_\beta e_\alpha{}^a. \quad (6.56)$$

Из равенства (6.55) следует полезная для вычислений формула

$$g^{\alpha\beta} \tilde{\Gamma}_{\alpha\beta}{}^\gamma = -\frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} g^{\alpha\gamma}).$$

В пространстве Римана–Картана–Вейля след неметричности пропорционален форме Вейля W_α из (6.20):

$$Q_\alpha = nW_\alpha.$$

Свертка аффинной связности (6.17) по первому и третьему индексам приводит к равенству

$$\Gamma_{\beta\alpha}{}^\beta = \tilde{\Gamma}_\alpha + T_\alpha + \frac{1}{2}Q_\alpha = \Gamma_\alpha + T_\alpha, \quad (6.57)$$

где введен след тензора кручения

$$T_\alpha := T_{\beta\alpha}{}^\beta. \quad (6.58)$$

Предложение 6.6.1. *Справедливо равенство*

$$\nabla_\alpha \varepsilon_{\alpha_1 \dots \alpha_n} = -\frac{1}{2}Q_\alpha \varepsilon_{\alpha_1 \dots \alpha_n}. \quad (6.59)$$

Доказательство. Из определения ковариантной производной и связи между полностью антисимметричным тензором и тензорной плотностью (2.63) следует равенство

$$\nabla_\alpha \varepsilon_{\alpha_1 \dots \alpha_n} = \hat{\varepsilon}_{\alpha_1 \dots \alpha_n} \partial_\alpha \sqrt{|g|} - \Gamma_{\alpha\alpha_1}{}^\beta \varepsilon_{\beta\alpha_2 \dots \alpha_n} - \dots - \Gamma_{\alpha\alpha_n}{}^\beta \varepsilon_{\alpha_1 \dots \alpha_{n-1}\beta}.$$

Это выражение антисимметрично по индексам $\alpha_1, \dots, \alpha_n$ и, следовательно, пропорционально полностью антисимметричному тензору:

$$\nabla_\alpha \varepsilon_{\alpha_1 \dots \alpha_n} = \varepsilon_{\alpha_1 \dots \alpha_n} X_\alpha,$$

где X_α – компоненты некоторого ковектора. Свертка полученного равенства с контравариантным тензором $\varepsilon^{\alpha_1 \dots \alpha_n}$ с учетом равенств (6.48), (6.49) и (37.57) определяет

$$X_\alpha = -\frac{1}{2}Q_\alpha. \quad \square$$

Следствие. Для метрической связности полностью антисимметричный тензор является ковариантно постоянным. $\square$

В дифференциальной геометрии естественным образом вводится понятие градиента и дивергенции.

Определение. Назовем *градиентом* скалярного поля (функции) f ковекторное поле (1-форму):

$$\text{grad } f := dx^\alpha \nabla_\alpha f = df = dx^\alpha \partial_\alpha f. \quad (6.60)$$

Назовем *дивергенцией* векторного поля X скалярное поле

$$\text{div } X := \nabla_\alpha X^\alpha = \partial_\alpha X^\alpha + \Gamma_{\beta\alpha}{}^\beta X^\alpha, \quad (6.61)$$

построенное с помощью ковариантной производной. $\square$

С учетом выражения аффинной связности через кручение и неметричность (6.17) дивергенция векторного поля (6.61) принимает вид

$$\nabla_\alpha X^\alpha = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} X^\alpha) + T_\alpha X^\alpha + \frac{1}{2}Q_\alpha X^\alpha. \quad (6.62)$$

В частности, в (псевдо)римановой геометрии

$$\tilde{\nabla}_\alpha X^\alpha = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} X^\alpha). \quad (6.63)$$

Эта полезная формула позволяет переписать инвариантный интеграл от дивергенции произвольного векторного поля в (псевдо)римановой геометрии в виде

$$\int dx \sqrt{|g|} \tilde{\nabla}_\alpha X^\alpha = \int dx \partial_\alpha (\sqrt{|g|} X^\alpha)$$

и воспользоваться формулой Стокса (см. раздел 3.7.2). В пространстве аффинной связности после применения формулы Стокса возникают дополнительные объемные интегралы с кручением и неметричностью (6.62).

Замечание. В разделе 3.5 мы определили дивергенцию (3.64) для внешних форм. Пусть на многообразии $\mathbb{M}$ задана метрика $g_{\alpha\beta}$ и 1-форма $A = dx^\alpha A_\alpha$, соответствующая вектору X , где $A_\alpha := X^\beta g_{\alpha\beta}$. Тогда из сравнения формул (3.57) и (6.61) следует, что $\operatorname{div} A = \operatorname{div} X$ тогда и только тогда, когда

$$T_\alpha X^\alpha + \frac{1}{2} Q_\alpha X^\alpha = 0.$$

В частности, равенство $\operatorname{div} A = \operatorname{div} X$ имеет место в (псевдо)римановой геометрии. $\square$

Дивергенция от антисимметричного тензорного поля, $F^{\alpha\beta} = -F^{\beta\alpha}$, равна

$$\nabla_\alpha F^{\alpha\beta} = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} F^{\alpha\beta}) + F^{\alpha\beta} \left(T_\alpha + \frac{1}{2} Q_\alpha \right) + \frac{1}{2} F^{\alpha\gamma} T_{\alpha\gamma}{}^\beta. \quad (6.64)$$

В (псевдо)римановой геометрии это тождество упрощается:

$$\tilde{\nabla}_\alpha F^{\alpha\beta} = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} F^{\alpha\beta}). \quad (6.65)$$

Приведем также формулу для антисимметричной ковариантной производной от 1-формы:

$$\nabla_\alpha X_\beta - \nabla_\beta X_\alpha = \partial_\alpha X_\beta - \partial_\beta X_\alpha - T_{\alpha\beta}{}^\gamma X_\gamma. \quad (6.66)$$

Из этого равенства следует, что антисимметризация обычной производной от 1-формы дает тензорное поле типа $(0, 2)$, т.к. последнее слагаемое является тензором. Антисимметризация индексов положена в основу определения внешней производной от произвольной формы (см. раздел 3.3).

Наличие ковариантной производной ∇_α и метрики $g_{\alpha\beta}$ позволяет строить ковариантные дифференциальные операторы, действующие на произвольные тензорные поля на многообразии. Для их построения достаточно взять произвольный дифференциальный оператор в (псевдо)евклидовом пространстве, заменить частные производные ∂_α на ковариантные ∇_α и (псевдо)евклидову метрику $\eta_{\alpha\beta}$ на метрику многообразия $g_{\alpha\beta}$. В общем случае эта процедура неоднозначна.

Инвариантный дифференциальный оператор второго порядка, действующий на произвольные тензорные поля на многообразии имеет вид

$$\Delta := g^{\alpha\beta} \nabla_\alpha \nabla_\beta = \nabla_\alpha g^{\alpha\beta} \nabla_\beta - Q_\alpha{}^{\alpha\beta} \nabla_\beta. \quad (6.67)$$

Он часто называется *оператором Лапласа–Бельтрами* независимо от сигнатуры метрики. В римановой геометрии оператор Лапласа–Бельтрами устроен проще:

$$\tilde{\Delta} := g^{\alpha\beta} \tilde{\nabla}_\alpha \tilde{\nabla}_\beta = \tilde{\nabla}_\alpha g^{\alpha\beta} \tilde{\nabla}_\beta. \quad (6.68)$$

Замечание. Если метрика имеет лоренцеву сигнатуру, то мы будем обозначать оператор Лапласа–Бельтрами квадратом $\square$, а не треугольником $\triangle$, по аналогии с оператором Даламбера. В этом случае уравнение $\square K_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r} = 0$ для каждой компоненты тензорного поля $K \in \mathcal{T}_s^r(\mathbb{M})$ будет гиперболического типа. $\square$

Действие оператора Лапласа–Бельтрами на скалярное поле можно записать в виде

$$\Delta f = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} g^{\alpha\beta} \partial_\beta f) + \left(T^\beta + \frac{1}{2} Q^\beta - Q_\alpha^{\alpha\beta} \right) \partial_\beta f. \quad (6.69)$$

В римановой геометрии правая часть этого равенства записывается через обычную производную от тензорной плотности:

$$\tilde{\Delta} f = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} g^{\alpha\beta} \partial_\beta f). \quad (6.70)$$

Приведем также формулу интегрирования по частям. Используя (6.49) и (6.57), проверяется, что

$$\partial_\alpha (\sqrt{|g|} X^{\alpha A} Y_A) = \sqrt{|g|} \nabla_\alpha X^{\alpha A} Y_A - \sqrt{|g|} \left(T_\alpha + \frac{1}{2} Q_\alpha \right) X^{\alpha A} Y_A + \sqrt{|g|} X^{\alpha A} \nabla_\alpha Y_A.$$

Здесь индекс A обозначает произвольную совокупность ковариантных и контравариантных индексов, по которым подразумевается суммирование. Интегрируя это соотношение по многообразию и пренебрегая граничными членами, получим равенство

$$\int dx \sqrt{|g|} X^{\alpha A} \nabla_\alpha Y_A = - \int dx \sqrt{|g|} \left[\nabla_\alpha X^{\alpha A} Y_A - \left(T_\alpha + \frac{1}{2} Q_\alpha \right) X^{\alpha A} Y_A \right]. \quad (6.71)$$

В римановой геометрии эта формула имеет тот же вид, что и обычное интегрирование по частям:

$$\int dx \sqrt{|g|} X^{\alpha A} \tilde{\nabla}_\alpha Y_A = - \int dx \sqrt{|g|} \left[\tilde{\nabla}_\alpha X^{\alpha A} Y_A \right]. \quad (6.72)$$

Отсюда следует, что с точностью до граничных членов имеет место равенство

$$\int dx \sqrt{|g|} X^A \tilde{\Delta} Y_A = \int dx \sqrt{|g|} \tilde{\Delta} X^A Y_A. \quad (6.73)$$

6.7 Локальное определение тензора кривизны

Помимо кручения аффинная связность $\Gamma_{\alpha\beta}^\gamma$ на многообразии $\mathbb{M}$ задает еще один важный геометрический объект – *тензор кривизны* аффинной связности (5.36) или *тензор Римана–Кристоффеля*. В локальной системе координат он имеет следующие компоненты:

$$R_{\alpha\beta\gamma}^\delta := \partial_\alpha \Gamma_{\beta\gamma}^\delta - \Gamma_{\alpha\gamma}^\epsilon \Gamma_{\beta\epsilon}^\delta - (\alpha \leftrightarrow \beta), \quad (6.74)$$

где скобки $(\alpha \leftrightarrow \beta)$ обозначают предыдущие слагаемые с переставленными индексами α и β .

Замечание. Тензор кривизны никакого отношения к метрике не имеет и определяется только связностью. $\square$

Тензор кривизны играет очень важную роль в дифференциальной геометрии и возникает в различных контекстах. Покажем, что тензор кривизны позволяет сформулировать критерий локальной тривиальности аффинной связности.

Пусть на многообразии $\mathbb{M}$ задана аффинная связность с нулевым кручением, $T_{\alpha\beta}^\gamma = 0$. При этом неметричность, если задана также метрика, может быть отлична от нуля. Рассмотрим соотношения (5.32) как уравнения на функции перехода $x^\alpha(x')$. Потребуем, чтобы в новой системе координат компоненты связности обращались в нуль в некоторой односвязной области. Тогда функции перехода должны удовлетворять уравнению

$$\frac{\partial^2 x^\alpha}{\partial x^{\beta'} \partial x^{\gamma'}} = - \frac{\partial x^\beta}{\partial x^{\beta'}} \frac{\partial x^\gamma}{\partial x^{\gamma'}} \Gamma_{\beta\gamma}^\alpha. \quad (6.75)$$

Дифференцируя это соотношение по $x^{\delta'}$ и исключая вторые производные от функций перехода с помощью исходного уравнения (6.75), получим равенство

$$\frac{\partial^3 x^\alpha}{\partial x^{\beta'} \partial x^{\gamma'} \partial x^{\delta'}} = \frac{\partial x^\beta}{\partial x^{\beta'}} \frac{\partial x^\gamma}{\partial x^{\gamma'}} \frac{\partial x^\delta}{\partial x^{\delta'}} (-\partial_{\delta'} \Gamma_{\beta\gamma}^\alpha + \Gamma_{\beta\delta'}^\epsilon \Gamma_{\epsilon\gamma}^\alpha + \Gamma_{\delta'\gamma}^\epsilon \Gamma_{\beta\epsilon}^\alpha).$$

Условия интегрируемости уравнений (6.75) получаются из этих уравнений антисимметризацией выражения в круглых скобках по индексам β', δ' или γ', δ' и приравниванием результата нулю. Обе антисимметризации приводят к единственному условию: равенству нулю тензора кривизны (6.74), $R_{\alpha\beta\gamma}^\delta = 0$.

На самом деле равенство нулю тензора кривизны является необходимым и достаточным условием локальной разрешимости системы уравнений (6.75) относительно матриц Якоби $\partial_{\alpha'} x^\alpha$ преобразования координат. После ее решения возникнет система равенств

$$\frac{\partial x^\alpha}{\partial x^{\alpha'}} = f_{\alpha'}^\alpha, \quad (6.76)$$

где в правой части стоят некоторые функции от x' . Эта система уравнений на функции перехода, в свою очередь, имеет свои условия интегрируемости:

$$\partial_{\alpha'} f_{\beta'}^\alpha - \partial_{\beta'} f_{\alpha'}^\alpha = 0. \quad (6.77)$$

Эти условия интегрируемости в общем случае выполнены не будут. Однако, можно доказать, что правая часть уравнений (6.76), как решений исходных уравнений (6.75) определена неоднозначно. Кроме того, этот произвол в выборе решений всегда можно зафиксировать таким образом, чтобы условия интегрируемости (6.77) для функций перехода были выполнены. Соответствующее доказательство в переменных Картана было дано в разделе 5.5. Таким образом, справедлива

Теорема 6.7.1. *Пусть на многообразии $\mathbb{M}$ задана геометрия Римана–Картана. Если тензоры кривизны и кручения равны нулю в некоторой односвязной области $\mathbb{U} \subset \mathbb{M}$, то, возможно, в меньшей окрестности существует такая система координат, в которой компоненты связности обратятся в нуль.*

Определение. Назовем аффинную связность на многообразии $\mathbb{M}$ *локально тривиальной*, если для любой точки $x \in \mathbb{M}$ найдется окрестность $\mathbb{U}_x \ni x$ и такая система координат, что компоненты связности $\Gamma_{\alpha\beta}^\gamma$ на $\mathbb{U}_x$ равны нулю. Такая аффинная связность называется также *интегрируемой*. $\square$

Таким образом, найден критерий локальной тривиальности аффинной связности.

Теорема 6.7.2. *Для локальной тривиальности аффинной связности на многообразии M необходимо и достаточно, чтобы ее кручение и тензор кривизны равнялись нулю на M .*

Замечание. Тензор кручения является тензором, и его компоненты нельзя обратить в нуль никаким преобразованием координат. Поэтому равенство нулю тензора кривизны необходимо для локальной тривиальности аффинной связности. $\square$

Пусть на многообразии M помимо связности $\Gamma_{\alpha\beta}^\gamma$ задана также метрика $g_{\alpha\beta}$. Тогда определен тензор неметричности (6.16). При ненулевом тензоре неметричности локальная тривиальность аффинной связности не означает, что многообразие является локально (псевдо)евклидовым. Действительно, при $\Gamma_{\alpha\beta}^\gamma = 0$ и $T_{\alpha\beta}^\gamma = 0$ уравнение (6.16) дает соотношение между метрикой и неметричностью

$$\partial_\alpha g_{\beta\gamma} = -Q_{\alpha\beta\gamma}. \quad (6.78)$$

В частности, метрику можно задать произвольно, и она будет определять неметричность. В римановой геометрии $T_{\alpha\beta}^\gamma = 0$ и $Q_{\alpha\beta\gamma} = 0$, и равенство нулю тензора кривизны, $\tilde{R}_{\alpha\beta\gamma}^\delta = 0$, означает, что найдется такая система координат, в которой все частные производные метрики обратятся в нуль, $\partial_\alpha g_{\beta\gamma} = 0$, т.е. метрика имеет постоянные компоненты в некоторой области $U \subset M$. С помощью последующего линейного преобразования координат ее всегда можно привести к диагональному виду, когда на диагонали будут стоять ± 1 в соответствии с исходной сигнатурой метрики. Это дает критерий локальной (псевдо)евклидовости (псевдо)риманова пространства. Сформулируем сразу глобальное утверждение. Оно использует геодезическую полноту многообразий, т.е. любая геодезическая для заданной связности может быть продолжена до бесконечных значений канонического параметра в обе стороны (см. раздел 16.1).

Теорема 6.7.3. *Пусть на многообразии M задана аффинная геометрия, и аффинная связность полна. Тогда, если тензоры кручения, неметричности и кривизны равны нулю на всем многообразии, то M изометрично либо (псевдо)евклидову пространству $\mathbb{R}^n$, либо факторпространству $\mathbb{R}^n/\mathbb{G}$, где $\mathbb{G}$ – группа преобразований, действующая на $\mathbb{R}^n$ свободно и собственнo разрывно.*

Доказательство. См., например, [54], следствие 1.9.5. $\square$

6.8 Свойства тензора кривизны

Помимо тождеств Бианки тензоры кручения и кривизны удовлетворяют ряду других дифференциальных соотношений. Выведем несколько полезных тождеств, исходя из определения тензора кривизны (6.74). Прямые вычисления с учетом выражения аффинной связности через метрику, кручение и неметричность показывают, что антисимметризация тензора кривизны по первым трем индексам определяется только тензором кручения и его ковариантными производными:

$$\begin{aligned} R_{\alpha\beta\gamma}^\delta + R_{\beta\gamma\alpha}^\delta + R_{\gamma\alpha\beta}^\delta = & \nabla_\alpha T_{\beta\gamma}^\delta + \nabla_\beta T_{\gamma\alpha}^\delta + \nabla_\gamma T_{\alpha\beta}^\delta + \\ & + T_{\alpha\beta}^\epsilon T_{\epsilon\gamma}^\delta + T_{\beta\gamma}^\epsilon T_{\epsilon\alpha}^\delta + T_{\gamma\alpha}^\epsilon T_{\epsilon\beta}^\delta. \end{aligned} \quad (6.79)$$

Это означает, что в римановой геометрии, а также в аффинной геометрии без кручения тензор кривизны удовлетворяет тождеству

$$R_{\alpha\beta\gamma}^\delta + R_{\beta\gamma\alpha}^\delta + R_{\gamma\alpha\beta}^\delta = 0. \quad (6.80)$$

Обратное утверждение в общем случае неверно. То есть из равенства (6.80) не следует, что кручение равно нулю.

Пусть на многообразии $\mathbb{M}$ задана аффинная геометрия, т.е. метрика и связность. Приведем явное выражение для тензора кривизны со всеми опущенными индексами. С учетом разложения связности (6.17) и тождества (6.25) получаем равенство

$$R_{\alpha\beta\gamma\delta} := R_{\alpha\beta\gamma}{}^{\epsilon} g_{\epsilon\delta} = \partial_{\alpha}\Gamma_{\beta\gamma\delta} + \Gamma_{\alpha\gamma}{}^{\epsilon}\Gamma_{\beta\delta\epsilon} - \Gamma_{\alpha\gamma}{}^{\epsilon}(T_{\beta\delta\epsilon} + Q_{\beta\delta\epsilon}) - (\alpha \leftrightarrow \beta). \quad (6.81)$$

Если у тензора кривизны со всеми опущенными индексами произвести симметризацию по последней паре индексов, то получим тождество

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\beta\delta\gamma} = \nabla_{\alpha}Q_{\beta\gamma\delta} - \nabla_{\beta}Q_{\alpha\gamma\delta} + T_{\alpha\beta}{}^{\epsilon}Q_{\epsilon\gamma\delta}. \quad (6.82)$$

Свертывая это тождество по последней паре индексов, получаем равенство

$$\begin{aligned} 2R_{\alpha\beta\gamma}{}^{\gamma} &= 2(\partial_{\alpha}\Gamma_{\beta} - \partial_{\beta}\Gamma_{\alpha}) = \nabla_{\alpha}Q_{\beta} - \nabla_{\beta}Q_{\alpha} + T_{\alpha\beta}{}^{\gamma}Q_{\gamma} = \\ &= \partial_{\alpha}Q_{\beta} - \partial_{\beta}Q_{\alpha}, \end{aligned} \quad (6.83)$$

где $Q_{\alpha} := Q_{\alpha\beta}{}^{\beta}$. Отсюда следует, что в геометрии Римана–Картана–Вейля свертка тензора кривизны по последним двум индексам дает напряженность для формы Вейля.

Рассмотрим симметрии тензора кривизны относительно перестановок индексов. В общем случае аффинной геометрии единственная симметрия тензора кривизны со всеми опущенными индексами – это антисимметрия по первой паре индексов:

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta}, \quad (6.84)$$

что сразу вытекает из определения (6.74). Отсюда следует, что в аффинной геометрии число линейно независимых компонент тензора кривизны равно

$$[R_{\alpha\beta\gamma\delta}] = \frac{n^3(n-1)}{2}.$$

В геометрии Римана–Картана тензор кривизны антисимметричен также и по второй паре индексов, что является следствием уравнения (6.82),

$$R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma}. \quad (6.85)$$

Поэтому число его линейно независимых компонент меньше:

$$[R_{\alpha\beta\gamma\delta}] = \frac{n^2(n-1)^2}{4}.$$

В (псевдо)римановой геометрии тензор кривизны обладает дополнительной симметрией: его антисимметризация по первым трем индексам тождественно обращается в нуль

$$\tilde{R}_{[\alpha\beta\gamma]\delta} = 0 \quad \Leftrightarrow \quad \tilde{R}_{\alpha\beta\gamma\delta} + \tilde{R}_{\beta\gamma\alpha\delta} + \tilde{R}_{\gamma\alpha\beta\delta} = 0. \quad (6.86)$$

что следует из уравнения (6.79). Число независимых антисимметризованных индексов $[\alpha\beta\gamma]$ в уравнении (6.86) равно

$$C_n^3 = \frac{n!}{3!(n-3)!} = \frac{n(n-1)(n-2)}{6}.$$

Следовательно, число линейно независимых компонент тензора кривизны равно

$$[\tilde{R}_{\alpha\beta\gamma\delta}] = \frac{n^2(n-1)^2}{4} - \frac{n^2(n-1)(n-2)}{6} = \frac{n^2(n^2-1)}{12}.$$

Из свойств (6.84), (6.85) и (6.86) следует, что в (псевдо)римановой геометрии тензор кривизны симметричен также относительно перестановки первой пары индексов со второй,

$$\tilde{R}_{\alpha\beta\gamma\delta} = \tilde{R}_{\gamma\delta\alpha\beta}. \quad (6.87)$$

Замечание. Обратные утверждения, связывающие геометрию с симметрией тензора кривизны неверны. Например, тензор кривизны может быть антисимметричен по второй паре индексов, и в то же время тензор неметричности может быть нетривиальным. $\square$

В (псевдо)римановой геометрии тензор кривизны со всеми опущенными индексами следующим образом выражается через метрику:

$$\tilde{R}_{\alpha\beta\gamma\delta} = \frac{1}{2}(\partial_{\alpha\gamma}^2 g_{\beta\delta} - \partial_{\alpha\delta}^2 g_{\beta\gamma} - \partial_{\beta\gamma}^2 g_{\alpha\delta} + \partial_{\beta\delta}^2 g_{\alpha\gamma}) + \tilde{\Gamma}_{\alpha\gamma}{}^{\epsilon} \tilde{\Gamma}_{\beta\delta\epsilon} - \tilde{\Gamma}_{\alpha\delta}{}^{\epsilon} \tilde{\Gamma}_{\beta\gamma\epsilon}. \quad (6.88)$$

Как видим, он линеен по вторым производным от метрики и квадратичен по первым производным.

Определение. По заданному тензору кривизны путем свертки пары индексов можно построить *тензор Риччи*

$$R_{\alpha\gamma} := R_{\alpha\beta\gamma}{}^{\beta}. \quad \square \quad (6.89)$$

Замечание. Тензор кривизны и тензор Риччи строятся только по аффинной связности, без использования метрики. $\square$

В общем случае тензор Риччи не обладает никакой симметрией по своим индексам. В (псевдо)римановой геометрии он симметричен относительно перестановки индексов:

$$\tilde{R}_{\alpha\gamma} = \tilde{R}_{\gamma\alpha}.$$

Определение. Свертывая тензор Риччи с обратной метрикой, получаем *скалярную кривизну* многообразия

$$R := R_{\alpha\beta} g^{\alpha\beta}. \quad (6.90)$$

Скалярная кривизна зависит и от связности, и от метрики. Функция

$$K := -\frac{1}{n(n-1)} R, \quad (6.91)$$

где R – скалярная кривизна, называется *гауссовой кривизной* многообразия. $\square$

В формуле (6.91) коэффициент подобран таким образом, что гауссова кривизна n -мерной сферы радиуса a равна $1/a^2$. При $n = 2$ гауссова кривизна поверхности равна произведению главных кривизн.

Пример 6.8.1. Пусть двумерная поверхность задана в трехмерном евклидовом пространстве $\mathbb{R}^3$ уравнением $z = f(x, y)$ с дважды непрерывно дифференцируемой правой частью, причем

$$f(0, 0) = 0, \quad f_x(0, 0) = 0, \quad f_y(0, 0) = 0,$$

где $f_x := \partial_x f$ и $f_y := \partial_y f$. Тогда скалярная кривизна $\tilde{R}$ в начале координат, соответствующая индуцированной метрике

$$ds^2 = dx^2 + dy^2 + dz^2 = (1 + f_x^2)dx^2 + 2f_x f_y dx dy + (1 + f_y^2)dy^2,$$

пропорциональна определителю гессиана:

$$\tilde{R} = -2 \det \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = -2K,$$

где K – гауссова кривизна поверхности. Поворотом евклидова пространства в плоскости x, y гессиан всегда можно привести к диагональному виду

$$\begin{pmatrix} f_{xx} & 0 \\ 0 & f_{yy} \end{pmatrix},$$

где, скажем, $f_{xx}(0, 0) \geq f_{yy}(0, 0)$. Если $f_{xx} > f_{yy}$, то оси x, y называются *главными направлениями* поверхности в данной точке. При $f_{xx} = f_{yy}$, главные направления не определены. Если гауссова кривизна положительна, $K > 0$, то поверхность в начале координат имеет локальный экстремум и лежит по одну сторону плоскости $z = 0$. Если кривизна меньше нуля, $K < 0$, то поверхность в начале координат имеет седловую точку. При $K = 0$ по крайней мере одна из вторых частных производных обращается в нуль, и мы имеем касание более высокого порядка. $\square$

Замечание. Многие авторы определяют тензор кривизны многообразия таким образом, чтобы скалярная кривизна совпала с гауссовой кривизной. Мы используем определения тензора кривизны, тензора Риччи и скалярной кривизны, которые широко используются в дифференциальной геометрии и приложениях. Это не приводит к недоразумениям. Однако надо помнить, что в наших обозначениях скалярная и гауссова кривизна двумерной сферы $\mathbb{S}_a^2$ радиуса a равны: $R = -2K = -2/a^2$ и $K = 1/a^2$ (см. вычисления далее в разделе 29.1). $\square$

В аффинной геометрии компоненты связности $\Gamma_{\alpha\beta}^\gamma$ инвариантны относительно изменения знака метрики:

$$g_{\alpha\beta} \mapsto -g_{\alpha\beta}, \quad T_{\alpha\beta}^\gamma \mapsto T_{\alpha\beta}^\gamma, \quad Q_{\alpha\beta}^\gamma \mapsto Q_{\alpha\beta}^\gamma. \quad (6.92)$$

Поэтому тензор кривизны и тензор Риччи также инвариантны. При этом скалярная кривизна (6.90) меняет знак.

Рассмотрим на многообразии $\mathbb{M}$ произвольное векторное поле $X = X^\alpha \partial_\alpha \in \mathcal{X}(\mathbb{M})$ и 1-форму $A = dx^\alpha A_\alpha \in \Lambda_1(\mathbb{M})$. Ковариантные производные от их компонент имеют вид (6.2), (6.3). Прямые вычисления показывают, что коммутатор двух ковариантных производных определяется тензором кривизны и кручения:

$$[\nabla_\alpha, \nabla_\beta]X^\gamma = R_{\alpha\beta}^\delta X^\delta - T_{\alpha\beta}^\delta \nabla_\delta X^\gamma, \quad (6.93)$$

$$[\nabla_\alpha, \nabla_\beta]A_\gamma = -R_{\alpha\beta\gamma}^\delta A_\delta - T_{\alpha\beta}^\delta \nabla_\delta A_\gamma. \quad (6.94)$$

и отличается только знаком перед слагаемым с кривизной. Коммутатор ковариантных производных от скалярного поля $\varphi(x)$ проще:

$$[\nabla_\alpha, \nabla_\beta]\varphi = -T_{\alpha\beta}{}^\gamma \nabla_\gamma \varphi \quad (6.95)$$

и определяется только тензором кручения.

Формулы (6.93), (6.94) обобщаются на тензоры произвольного ранга.

Пример 6.8.2. Коммутатор ковариантных производных от компонент тензора $K \in \mathcal{T}_1^1(\mathbb{M})$ типа $(1, 1)$ имеет вид

$$[\nabla_\alpha, \nabla_\beta]K^\gamma{}_\delta = R_{\alpha\beta\epsilon}{}^\gamma K^\epsilon{}_\delta - R_{\alpha\beta\delta}{}^\epsilon K^\gamma{}_\epsilon - T_{\alpha\beta}{}^\epsilon \nabla_\epsilon K^\gamma{}_\delta.$$

Для тензоров более высокого ранга будем иметь по одному слагаемому с кривизной со знаком плюс и минус, соответственно, для каждого контравариантного и ковариантного индекса и одно общее слагаемое с тензором кручения. $\square$

Формулы для коммутатора ковариантных производных можно использовать для доказательства некоторых тождеств. Докажем несколько формул, справедливых в (псевдо)римановой геометрии. Пусть $F^{\alpha\beta}$ – произвольный антисимметричный тензор второго ранга. Тогда в (псевдо)римановой геометрии справедливо равенство

$$\tilde{\nabla}_\alpha \tilde{\nabla}_\beta F^{\alpha\beta} = 0.$$

Действительно,

$$[\tilde{\nabla}_\alpha, \tilde{\nabla}_\beta]F^{\alpha\beta} = \tilde{R}_{\alpha\beta\gamma}{}^\alpha F^{\gamma\beta} + \tilde{R}_{\alpha\beta\gamma}{}^\beta F^{\alpha\gamma} = 2\tilde{R}_{\alpha\beta}F^{\alpha\beta} = 0,$$

поскольку тензор Риччи в этом случае симметричен.

Предложение 6.8.1. В (псевдо)римановой геометрии имеет место следующая формула

$$\tilde{\nabla}_\alpha \tilde{\nabla}_\beta \tilde{R}^{\alpha\beta\gamma\delta} = 0.$$

Доказательство. Справедливы следующие равенства:

$$\begin{aligned} 2\tilde{\nabla}_\alpha \tilde{\nabla}_\beta \tilde{R}^{\alpha\beta\gamma\delta} &= [\tilde{\nabla}_\alpha, \tilde{\nabla}_\beta]R^{\alpha\beta\gamma\delta} = \\ &= \tilde{R}_{\alpha\beta\epsilon}{}^\alpha \tilde{R}^{\epsilon\beta\gamma\delta} + \tilde{R}_{\alpha\beta\epsilon}{}^\beta \tilde{R}^{\alpha\epsilon\gamma\delta} + \tilde{R}_{\alpha\beta\epsilon}{}^\gamma \tilde{R}^{\alpha\beta\epsilon\delta} + \tilde{R}_{\alpha\beta\epsilon}{}^\delta \tilde{R}^{\alpha\beta\gamma\epsilon}. \end{aligned}$$

Первые два слагаемых обращаются в нуль в силу симметрии тензора Риччи. Последние два слагаемых сокращаются:

$$\tilde{R}_{\alpha\beta\epsilon}{}^\gamma \tilde{R}^{\alpha\beta\epsilon\delta} - \tilde{R}_{\alpha\beta\epsilon}{}^\delta \tilde{R}^{\alpha\beta\epsilon\gamma} = 0$$

в силу симметрии слагаемых по индексам γ, δ . $\square$

В заключение настоящего раздела приведем важную формулу, связывающую скалярную кривизну $R(g, \Gamma)$ в геометрии Римана–Картана со скалярной кривизной $\tilde{R}(g)$ в соответствующей (псевдо)римановой геометрии, вычисленной только по метрике при нулевом кручении:

$$R + \frac{1}{4}T_{\alpha\beta\gamma}T^{\alpha\beta\gamma} - \frac{1}{2}T_{\alpha\beta\gamma}T^{\gamma\alpha\beta} - T_\alpha T^\alpha - \frac{2}{\sqrt{|g|}}\partial_\alpha(\sqrt{|g|}T^\alpha) = \tilde{R}, \quad (6.96)$$

где $T_\alpha := T_{\beta\alpha}{}^\beta$ – след тензора кручения. Это тождество проверяется путем прямых вычислений и справедливо в пространстве произвольной размерности $n \geq 2$ и для метрики произвольной сигнатуры. Приведенное тождество можно переписать в явно ковариантном виде

$$R + \frac{1}{4}T_{\alpha\beta\gamma}T^{\alpha\beta\gamma} - \frac{1}{2}T_{\alpha\beta\gamma}T^{\gamma\alpha\beta} - T_\alpha T^\alpha - 2\tilde{\nabla}_\alpha T^\alpha = \tilde{R},$$

где мы воспользовались тождеством (6.63). Важность этой формулы заключается в том, что, поскольку обе скалярные кривизны приводят к уравнениям движения не выше второго порядка (в переменных Картана), то при выборе лагранжиана можно ограничиться одной скалярной кривизной и квадратами тензоров кручения в моделях гравитации с динамическим кручением.

6.9 Неголономный базис

Аффинная геометрия на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, задается метрикой $g_{\alpha\beta}$ и аффинной связностью $\Gamma_{\alpha\beta}{}^\gamma$ или, что эквивалентно, метрикой $g_{\alpha\beta}$, кручением $T_{\alpha\beta}{}^\gamma$ и неметричностью $Q_{\alpha\beta\gamma}$. При таком описании каждое преобразование координат сопровождается соответствующим преобразованием компонент тензорных полей относительно координатного базиса. Существует также другой способ описания геометрии, когда компоненты тензорных полей рассматриваются относительно репера, который не меняется при преобразовании координат. В этом случае на компоненты тензорных полей действует группа локальных преобразований $\mathrm{GL}(n, \mathbb{R})$, что соответствует вращению репера. В результате аффинная геометрия на многообразии $\mathbb{M}$ будет задана репером $e_\alpha{}^a$ и линейной или $\mathrm{GL}(n, \mathbb{R})$ связностью $\omega_{\alpha a}{}^b$ (см. раздел 5.4).

Определение. Переменные репер $e_\alpha{}^a(x)$ и $\mathrm{GL}(n, \mathbb{R})$ связность $\omega_{\alpha a}{}^b(x)$, задающие на многообразии $\mathbb{M}$ аффинную геометрию, называются *переменными Картана*. В четырехмерном пространстве-времени репер называется *тетрадой*. В двумерном и трехмерном пространстве репер называется, соответственно, *диадой* и *триадой*. $\square$

Замечание. В моделях математической физики переменные Картана, как правило, упрощают вычисления и необходимы при рассмотрении спинорных полей на многообразии $\mathbb{M}$. $\square$

Напомним, что координатный базис касательных пространств $T_x(\mathbb{M})$ во всех точках многообразия $x \in \mathbb{M}$ мы обозначаем ∂_α , и он называется *голономным*. Важным свойством координатных базисных векторов ∂_α является их коммутативность:

$$[\partial_\alpha, \partial_\beta] = 0.$$

Предположим, что в каждой точке многообразия $x \in \mathbb{M}$ задан произвольный базис касательного пространства $e_a(x)$ (репер) и дуальный к нему базис 1-форм $e^a(x)$ (корепер), $a = 1, \dots, n$. Дуальность означает, что значение 1-форм e^a на векторных полях e_b равно символу Кронекера: $e^a(e_b) = \delta_b^a$.

Замечание. Как уже отмечалось, репер может существовать глобально не для всех многообразий. Например, его не существует на неориентируемых многообразиях. В таких случаях все, сказанное ниже, имеет локальный характер. Тем не менее полученные формулы важны для вычислений, которые, как правило, проводятся в какой-либо системе координат. $\square$

Репер и корепер можно разложить по координатному базису:

$$e_a = e^\alpha_a \partial_\alpha, \quad e^a = dx^\alpha e_\alpha^a,$$

где e^α_a и e_α^a – взаимно обратные невырожденные матрицы, что является следствием дуальности базисов,

$$e^\alpha_a e_\alpha^b = \delta_a^b, \quad e^\alpha_a e_\beta^a = \delta_\beta^\alpha.$$

По предположению, матрицы e^α_a и e_α^a невырождены и достаточно гладко зависят от точки многообразия.

В общем случае репер представляет собой *неголономный базис* касательного пространства, т.е. не существует такой системы координат $x^a = x^a(x)$, что

$$e^\alpha_a = \partial_a x^\alpha. \quad (6.97)$$

Репер определен с точностью до локальных $\mathbb{GL}(n, \mathbb{R})$ преобразований, действующих на латинские индексы. Его важнейшей характеристикой являются *компоненты неголономности* c_{ab}^c , которые определяются коммутатором базисных векторных полей:

$$[e_a, e_b] := c_{ab}^c e_c \quad (6.98)$$

и антисимметричны по нижним индексам,

$$c_{ab}^c(x) = -c_{ba}^c(x).$$

Из тождеств Якоби для алгебры Ли векторных полей,

$$[e_a[e_b, e_c]] + [e_b[e_c, e_a]] + [e_c[e_a, e_b]] = 0,$$

следуют тождества для компонент неголономности:

$$\partial_a c_{bc}^d + \partial_b c_{ca}^d + \partial_c c_{ab}^d + c_{ab}^e c_{ce}^d + c_{bc}^e c_{ae}^d + c_{ca}^e c_{be}^d = 0, \quad (6.99)$$

где $\partial_a := e_a = e^\alpha_a \partial_\alpha$.

Нетрудно проверить, что равенство нулю компонент неголономности является необходимым и достаточным условием локальной разрешимости системы уравнений (6.97). Это означает, что, если компоненты неголономности равны нулю в некоторой области, то для любой точки из этой области существует окрестность, в которой можно выбрать такую систему координат, что базис станет голономным $e_a = \partial/\partial x^a$.

Из определения (6.98) следует явное выражение для компонент неголономности через компоненты репера и их производные

$$c_{ab}^c = (e^\alpha_a \partial_\alpha e_\beta^b - e^\alpha_b \partial_\alpha e_\beta^a) e_\beta^c. \quad (6.100)$$

Умножив это соотношение на обратные матрицы e_α^a , получим эквивалентную формулу

$$c_{\alpha\beta}^c := e_\alpha^a e_\beta^b c_{ab}^c = -\partial_\alpha e_\beta^c + \partial_\beta e_\alpha^c, \quad (6.101)$$

которую можно переписать в виде

$$de^c = -\frac{1}{2} e^a \wedge e^b c_{ba}^c, \quad (6.102)$$

где использовано определение внешнего умножения и дифференцирования форм (см. раздел 3).

Многие формулы содержат след компонент неголономности, который определяется следующим образом:

$$c_a := c_{ba}{}^b = \partial_a e^\alpha{}_a + \frac{\partial_a \sqrt{|g|}}{\sqrt{|g|}}. \quad (6.103)$$

Компоненты неголономности ковариантны относительно преобразования координат $x^\alpha \rightarrow x^{\alpha'}(x)$, но не являются компонентами какого-либо тензора относительно локальных $\mathbb{GL}(n, \mathbb{R})$ преобразований.

Использование неголономного базиса вместо координатного бывает значительно удобнее и часто используется в приложениях. Поэтому получим основные формулы дифференциальной геометрии в неголономном базисе.

Произвольное векторное поле можно разложить как по координатному, так и по некоординатному (неголономному) базису

$$X = X^\alpha \partial_\alpha = X^a e_a,$$

где $X^\alpha = e^\alpha{}_a X^a$ и $X^a = X^\alpha e_\alpha{}^a$. Предположим, что переход от греческих индексов к латинским и наоборот у компонент тензорных полей произвольного ранга всегда осуществляется с помощью компонент репера и корепера. При этом все симметрии относительно перестановок индексов, конечно, сохраняются.

Если на $\mathbb{M}$ задана метрика $g_{\alpha\beta}$, то ее компоненты в неголономном базисе имеют вид

$$g_{ab} = e^\alpha{}_a e^\beta{}_b g_{\alpha\beta}. \quad (6.104)$$

В общем случае компоненты метрики $g_{ab}(x)$ зависят от точки многообразия. Метрика g_{ab} всегда имеет ту же сигнатуру, что и метрика $g_{\alpha\beta}$, т.к. матрица $e^\alpha{}_a$ невырождена. Подъем и опускание греческих и латинских индексов осуществляется с помощью метрик $g_{\alpha\beta}$ и g_{ab} , соответственно.

Как правило, репер используют в тех случаях, когда матрица g_{ab} является диагональной и постоянной, а на диагонали расположены плюс и минус единицы:

$$g_{ab} = \eta_{ab} := \text{diag} \left(\underbrace{+\cdots+}_p \underbrace{-\cdots-}_q \right), \quad p + q = n.$$

Локально такой репер существует, поскольку уравнение (6.104) при одинаковых сигнатурах метрик g_{ab} и $g_{\alpha\beta}$ всегда разрешимо относительно репера. Такой репер называется *ортонормальным* и определен с точностью до $\mathbb{O}(p, q)$ вращений. Ортонормальный базис часто бывает более удобным, т.к. метрика в этом базисе постоянна.

Для римановой метрики множество реперов делится на два класса: с положительным и отрицательным определителем. Для многообразий с метрикой лоренцевой сигнатуры множество реперов можно разбить на четыре класса, по числу компонент связности группы Лоренца (см. раздел 1.9.1).

Компоненты тензоров второго и более высокого рангов могут содержать одновременно и греческие, и латинские индексы. По построению, ковариантная производная от компонент такого тензора содержит по одному слагаемому с аффинной связностью для каждого греческого индекса и одному слагаемому с линейной связностью для каждого латинского индекса. Если репер задан, то из определения (локальной формы) линейной связности (5.52) следует взаимно однозначная связь между линейной и аффинной связностью:

$$\omega_{\alpha a}{}^b = \Gamma_{\alpha\beta}{}^\gamma e^\beta{}_a e_\gamma{}^b - \partial_\alpha e^\beta{}_a e^\beta{}_b. \quad (6.105)$$

Эту формулу можно переписать в виде равенства нулю ковариантной производной от компонент корепера:

$$\nabla_\alpha e_\beta^a = \partial_\alpha e_\beta^a - \Gamma_{\alpha\beta}^\gamma e_\gamma^a + e_\beta^b \omega_{\alpha b}^a = 0. \quad (6.106)$$

Отсюда следует, что ковариантная производная от компонент репера также обращается в нуль

$$\nabla_\alpha e^\beta_a = \partial_\alpha e^\beta_a + \Gamma_{\alpha\gamma}^\beta e^\gamma_a - \omega_{\alpha a}^b e^\beta_b = 0.$$

Тогда, используя правило Лейбница, можно свободно переходить от греческих индексов к латинским и наоборот под знаком ковариантного дифференцирования:

$$\begin{aligned} \nabla_\alpha X^a &= \nabla_\alpha (X^\beta e_\beta^a) = (\nabla_\alpha X^\beta) e_\beta^a, \\ \nabla_\alpha X_a &= \nabla_\alpha (X_\beta e^\beta_a) = (\nabla_\alpha X_\beta) e^\beta_a, \end{aligned}$$

где

$$\begin{aligned} \nabla_\alpha X^a &= \partial_\alpha X^a + \omega_{\alpha b}^a X^b, \\ \nabla_\alpha X_a &= \partial_\alpha X_a - \omega_{\alpha a}^b X_b \end{aligned} \quad (6.107)$$

— ковариантные производные от компонент векторного поля относительно неголономного базиса.

Если аффинная связность не является метрической, то операция подъема и опускания индексов с помощью метрик $g_{\alpha\beta}$ и g_{ab} не коммутирует с ковариантной производной.

Замечание. Формулу (6.105) можно рассматривать, как калибровочное преобразование (вращение) $\Gamma \mapsto \omega$ в касательном пространстве, которое совпадает с преобразованием калибровочных полей Янга–Миллса (25.205). При этом репер $e_\alpha^a \in \mathbb{GL}(n, \mathbb{R})$ играет роль матрицы преобразования (локального вращения), а координаты многообразия не затрагиваются. $\square$

Формулы для кривизны (5.57) и кручения (5.56) содержат два греческих индекса. Эти индексы также можно преобразовать в неголономные. Простые вычисления приводят к следующим компонентам тензора кривизны и кручения в неголономном базисе:

$$R_{abc}^d = \partial_a \omega_{bc}^d - \partial_b \omega_{ac}^d - \omega_{ac}^e \omega_{be}^d + \omega_{bc}^e \omega_{ae}^d - c_{ab}^e \omega_{ec}^d, \quad (6.108)$$

$$T_{ab}^c = \omega_{ab}^c - \omega_{ba}^c - c_{ab}^c, \quad (6.109)$$

где $\omega_{ab}^c := e_a^\alpha \omega_{\alpha b}^c$ и $\partial_a := e_a^\alpha \partial_\alpha$. Эти формулы также часто используются в приложениях, особенно тогда, когда компоненты линейной связности ω_{abc} являются постоянными относительно некоторого неголономного базиса. В разделе 8 мы используем их для вычисления тензора кручения и кривизны групп Ли.

В координатном базисе преобразование координат сопровождается преобразованием компонент тензорных полей. Введение репера позволяет отделить преобразование координат от преобразований в касательном пространстве. Это достигается путем введения n^2 новых полей $e_\alpha^a(x)$. В результате появляется дополнительная возможность совершать локальные $\mathbb{GL}(n, \mathbb{R})$ преобразования, зависящие также от n^2 функций, в касательном пространстве, не затрагивая координат многообразия. Очевидно, что всегда можно совершить такое преобразование, что в некоторой области репер совпадет с координатным базисом $e_\alpha^a = \delta_\alpha^a$. В этом случае линейная связность

совпадет с аффинной $\omega_{\alpha a}^b = \Gamma_{\alpha a}^b$, а выражения для кривизны (6.108) и кручения (6.109) перейдут в уже знакомые формулы аффинной геометрии, т.к. компоненты неголономности обратятся в нуль: $c_{ab}^c = 0$.

Если на многообразии задана метрика g_{ab} общего вида, для которой

$$\nabla_a g_{bc} = -Q_{abc},$$

где Q_{abc} – тензор неметричности, то из уравнения (6.17) и определения (6.105) следует выражение для линейной связности со всеми неголономными индексами:

$$\begin{aligned} \omega_{abc} = & \frac{1}{2}(\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}) + \frac{1}{2}(c_{abc} - c_{bca} + c_{cab}) + \\ & + \frac{1}{2}(T_{abc} - T_{bca} + T_{cab}) + \frac{1}{2}(Q_{abc} + Q_{bca} - Q_{cab}), \end{aligned} \quad (6.110)$$

где $\omega_{abc} := \omega_{ab}^d g_{dc}$ и $c_{abc} := c_{ab}^d g_{dc}$. В этом случае линейную связность можно выразить через метрику, репер, кручение и неметричность.

Условие метричности связности в неголономном базисе принимает вид

$$\nabla_a g_{bc} = \partial_a g_{bc} - \omega_{abc} - \omega_{acb} = 0, \quad (6.111)$$

что эквивалентно равенству $Q_{abc} = 0$.

Предположим, что на $\mathbb{M}$ задана метрика и метрическая связность. Тогда выражение для тензора кривизны (6.108) со всеми опущенными индексами в неголономном базисе равно

$$R_{abcd} = \partial_a \omega_{bcd} - \partial_b \omega_{acd} - \omega_{bc}^e \omega_{ade} + \omega_{ac}^e \omega_{bde} - c_{ab}^e \omega_{ecd}, \quad (6.112)$$

где использовано равенство (6.111).

В геометрии Римана–Картана ($Q_{abc} = 0$) всегда можно выбрать ортонормальный репер, для которого $g_{ab} = \delta_{ab}$ или η_{ab} , если сигнатура метрики отличается от евклидовой. Такие реперы определены с точностью до локальных $\mathbb{SO}(n)$ вращений (или $\mathbb{SO}(p, q)$ вращений, $p+q = n$, для неевклидовой сигнатуры). Тогда $\mathbb{GL}(n, \mathbb{R})$ связность редуцируется к $\mathbb{SO}(n)$ связности (или $\mathbb{SO}(p, q)$ связности). В этом случае из условия метричности (6.111) вытекает, что компоненты линейной связности антисимметричны по последней паре индексов, $\omega_{abc} = -\omega_{acb}$, поскольку $\partial_a g_{bc} = 0$. Кроме этого, тензор кривизны со всеми ковариантными индексами (6.112) антисимметричен как по первой, так и по второй паре индексов.

В (псевдо)римановой геометрии кручение и неметричность равны нулю. Тогда из формулы (6.110) для ортонормального репера следует выражение для соответствующей $\mathbb{SO}(n)$ или $\mathbb{SO}(p, q)$ связности через компоненты неголономности:

$$\tilde{\omega}_{abc} = \frac{1}{2}(c_{abc} - c_{bca} + c_{cab}). \quad (6.113)$$

Если (псевдо)риманово многообразие допускает векторное поле Киллинга $K = K^a e_a$, то уравнение Киллинга (15.4) в неголономном базисе принимает вид

$$\tilde{\nabla}_a K_b + \tilde{\nabla}_b K_a = 0, \quad (6.114)$$

где

$$\tilde{\nabla}_a K_b = e^{\alpha}_a \partial_{\alpha} K_b - \tilde{\omega}_{ab}^c K_c, \quad K_b := K^a g_{ab}. \quad (6.115)$$

Как и ранее знак тильды означает, что связность построена по метрике при нулевом кручении и неметричности. Эта формула следует из того, что переход между индексами можно проводить под знаком ковариантного дифференцирования.

6.10 Тождества Бианки

Тождества Бианки играют большую роль в приложениях и поэтому в настоящем разделе мы рассмотрим их в компонентах. Громоздкие, но прямые вычисления позволяют записать тождества (5.45) и (5.46) после спуска на базу в переменных Картана:

$$\nabla_\alpha R_{\beta\gamma a}{}^b + \nabla_\beta R_{\gamma\alpha a}{}^b + \nabla_\gamma R_{\alpha\beta a}{}^b = T_{\alpha\beta}{}^\delta R_{\gamma\delta a}{}^b + T_{\beta\gamma}{}^\delta R_{\alpha\delta a}{}^b + T_{\gamma\alpha}{}^\delta R_{\beta\delta a}{}^b, \quad (6.116)$$

$$\begin{aligned} \nabla_\alpha T_{\beta\gamma}{}^a + \nabla_\beta T_{\gamma\alpha}{}^a + \nabla_\gamma T_{\alpha\beta}{}^a &= T_{\alpha\beta}{}^\delta T_{\gamma\delta}{}^a + T_{\beta\gamma}{}^\delta T_{\alpha\delta}{}^a + T_{\gamma\alpha}{}^\delta T_{\beta\delta}{}^a + \\ &+ R_{\alpha\beta\gamma}{}^a + R_{\beta\gamma\alpha}{}^a + R_{\gamma\alpha\beta}{}^a. \end{aligned} \quad (6.117)$$

В аффинной геометрии их можно свернуть, соответственно, с репером e^γ_b и e^γ_a :

$$\begin{aligned} \nabla_\alpha R_{\beta a} - \nabla_\beta R_{\alpha a} + \nabla_\gamma R_{\alpha\beta a}{}^\gamma &= -T_{\alpha\beta}{}^\delta R_{\delta a} - T_{\alpha\gamma}{}^\delta R_{\alpha\delta a}{}^\gamma + T_{\beta\gamma}{}^\delta R_{\alpha\delta a}{}^\gamma, \\ \nabla_\alpha T_\beta - \nabla_\beta T_\alpha - \nabla_\gamma T_{\alpha\beta}{}^\gamma &= T_{\alpha\gamma}{}^\delta T_{\beta\delta}{}^\gamma - T_{\beta\gamma}{}^\delta T_{\alpha\delta}{}^\gamma - T_{\alpha\beta}{}^\gamma T_\gamma + R_{\alpha\beta} - R_{\beta\alpha} - R_{\alpha\beta\gamma}{}^\gamma. \end{aligned}$$

В геометрии Римана–Картана, когда выполнено условие метричности, последнее слагаемое в последнем тождестве обращается в нуль. Тогда свернутые тождества Бианки можно переписать в виде

$$\nabla_\alpha R_{\beta\gamma} - \nabla_\beta R_{\alpha\gamma} + \nabla_\delta R_{\alpha\beta\gamma}{}^\delta = -T_{\alpha\beta}{}^\delta R_{\delta\gamma} - T_{\alpha\delta}{}^\epsilon R_{\beta\epsilon\gamma}{}^\delta + T_{\beta\delta}{}^\epsilon R_{\alpha\epsilon\gamma}{}^\delta, \quad (6.118)$$

$$\nabla_\alpha T_\beta - \nabla_\beta T_\alpha - \nabla_\gamma T_{\alpha\beta}{}^\gamma = T_{\alpha\gamma}{}^\delta T_{\beta\delta}{}^\gamma - T_{\beta\gamma}{}^\delta T_{\alpha\delta}{}^\gamma - T_{\alpha\beta}{}^\gamma T_\gamma + R_{\alpha\beta} - R_{\beta\alpha}, \quad (6.119)$$

где мы перешли ко всем греческим индексам.

Из последнего тождества (6.119) следует, что антисимметричная часть тензора Риччи в (псевдо)римановой геометрии (при нулевом кручении) всегда равна нулю: $\tilde{R}_{\alpha\beta} = \tilde{R}_{\beta\alpha}$.

Далее, свертка (6.118) с $g^{\alpha\gamma}$ (при нулевом тензоре неметричности) приводит к равенству

$$2\nabla_\beta R_\alpha{}^\beta - \nabla_\alpha R = 2T_{\alpha\beta\gamma}{}^\beta R^{\gamma\beta} - R_{\alpha\beta\gamma\delta} T^{\gamma\delta\beta}. \quad (6.120)$$

В (псевдо)римановой геометрии это тождество упрощается:

$$\tilde{\nabla}_\beta G^{\beta\alpha} = 0, \quad (6.121)$$

где

$$G_{\alpha\beta} := \tilde{R}_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\tilde{R}$$

– тензор Эйнштейна, построенный из тензора Риччи и скалярной кривизны при нулевом кручении и тензоре неметричности.

Тождество (6.119) антисимметрично по индексам $\alpha\beta$ и поэтому свертка с $g^{\alpha\beta}$ приводит к тождеству $0 = 0$.

Глава 7

Криволинейные координаты в $\mathbb{R}^3$

Формализм дифференциальной геометрии, развитый в предыдущих разделах, становится полезным и естественным даже в евклидовом пространстве, если вычисления проводятся в криволинейных системах координат. В настоящей главе мы продемонстрируем это на примере сферической и цилиндрической систем координат в трехмерном евклидовом пространстве $\mathbb{R}^3$. В частности, понятие метрики и символов Кристоффеля чрезвычайно полезно при нахождении явного вида ковариантных дифференциальных операторов, которые часто используются в приложениях.

7.1 Сферические координаты

Рассмотрим *сферические координаты* r, θ, φ в трехмерном евклидовом пространстве $\mathbb{R}^3$ с декартовыми координатами x, y, z , рис. 7.1. Функции перехода от сферических

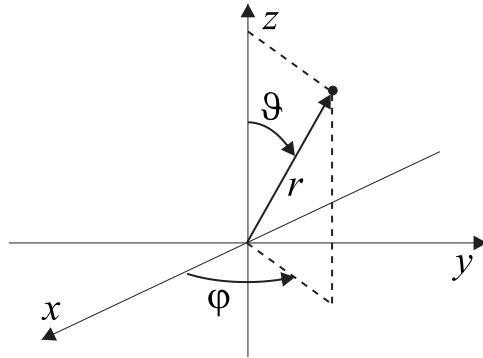


Рис. 7.1: Сферические координаты r, θ, φ в трехмерном евклидовом пространстве $\mathbb{R}^3$.

координат к декартовым имеют хорошо известный вид

$$\begin{aligned}x &= r \sin \theta \cos \varphi, \\y &= r \sin \theta \sin \varphi, \\z &= r \cos \theta,\end{aligned}\tag{7.1}$$

и определены при всех значениях r, θ, φ . Углы θ и φ называются, соответственно, *азимутальным* и *полярным*. Якобиан этого преобразования легко вычислить

$$J = r^2 \sin \theta.\tag{7.2}$$

Отсюда следует, что преобразование координат (7.1) вырождено при $r = 0$ или $\theta = 0, \pi$, т.е. на оси z . Действительно, обратные преобразования,

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}, \\ \theta &= \operatorname{arctg} \frac{\sqrt{x^2 + y^2}}{z}, \\ \varphi &= \operatorname{arctg} \frac{y}{x}, \end{aligned} \quad (7.3)$$

однозначно определены всюду, кроме оси z . При этом все евклидово пространство, из которого удалена полуплоскость $y = 0, x \geq 0$, включающая ось z , взаимно однозначно отображается на открытую область

$$0 < r < \infty, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi, \quad (7.4)$$

Если полуплоскость не удалять, то точки с координатами φ и $\varphi + 2\pi$ необходимо отождествить.

Переход к сферическим координатам сохраняет ориентацию, поскольку якобиан перехода (7.2) положителен в области (7.4).

Из формул (7.1) следуют правила преобразования дифференциалов:

$$\begin{aligned} dx &= \sin \theta \cos \varphi dr + r \cos \theta \cos \varphi d\theta - r \sin \theta \sin \varphi d\varphi, \\ dy &= \sin \theta \sin \varphi dr + r \cos \theta \sin \varphi d\theta + r \sin \theta \cos \varphi d\varphi, \\ dz &= \cos \theta dr - r \sin \theta d\theta. \end{aligned} \quad (7.5)$$

Подставляя эти выражения в евклидову метрику

$$ds^2 = dx^2 + dy^2 + dz^2, \quad (7.6)$$

получим следующее выражение для метрики в сферической системе координат

$$ds^2 = dr^2 + r^2 d\Omega, \quad (7.7)$$

где

$$d\Omega := d\theta^2 + \sin^2 \theta d\varphi^2 \quad (7.8)$$

– дифференциал телесного угла. Отсюда следует, что компоненты евклидовой метрики в сферической системе координат

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix} \quad (7.9)$$

являются функциями, а не константами. Отметим, что определитель метрики равен квадрату якобиана преобразования координат, $\det g_{\alpha\beta} = J^2$.

Помимо инвариантной квадратичной формы декартовых дифференциалов (7.6) существует еще одна линейно с ней независимая квадратичная форма

$$(xdx + ydy + zdz)^2,$$

инвариантная относительно $\mathbb{SO}(3)$ вращений. В сферических координатах также имеются две независимые инвариантные квадратичные формы, которые даются, например, выражениями dr^2 и $d\Omega$. Между этими формами существует связь:

$$\begin{aligned} dr^2 &= \frac{1}{r^2} (xdx + ydy + zdz)^2, \\ d\Omega &= \frac{1}{r^2} (dx^2 + dy^2 + dz^2) - \frac{1}{r^4} (xdx + ydy + zdz)^2. \end{aligned} \quad (7.10)$$

Обсудим, какие структуры существуют в евклидовом пространстве $\mathbb{R}^3$, инвариантные относительно $\mathbb{SO}(3)$ вращений. Общая теория инвариантных структур будет рассмотрена позже в разделе 9.3.

Из декартовых координат можно составить только одну функционально независимую сферически симметричную комбинацию – это радиус $r := \sqrt{x^2 + y^2 + z^2}$. Нетрудно показать, что любая сферически симметричная функция или плотность h может зависеть только от радиуса, $h = h(r)$.

Сферически симметричное векторное поле в сферической системе координат имеет только одну нетривиальную компоненту, направленную по радиусу,

$$X = \{X^r(r) := f(r), 0, 0\}.$$

В декартовой системе координат оно имеет вид

$$X = \left\{ \frac{x}{r}f, \frac{y}{r}f, \frac{z}{r}f \right\}$$

или

$$X^i = \frac{x^i}{r}f(r),$$

где $\{x^i\} = \{x, y, z\}$, $i = 1, 2, 3$, а $f(r)$ – некоторая функция только от радиуса.

Инвариантный симметричный тензор второго ранга удобно записывать в декартовой системе координат, выделив явно след и бесследовую часть (неприводимые компоненты):

$$X^{ij} = \delta^{ij}f_1(r) + \left(\delta^{ij} - 3\frac{x^i x^j}{r^2} \right) f_2(r),$$

где δ^{ij} – евклидова метрика и $f_{1,2}$ – некоторые функции. Первое слагаемое в этом разложении пропорционально следу тензора $X^{ij}\delta_{ij} = 3f_1$, а второе слагаемое имеет нулевой след. То есть сферически инвариантный тензор второго ранга взаимно однозначно определяется двумя функциями только от радиуса: $f_1(r)$ и $f_2(r)$.

Антисимметричный тензор второго ранга $Y^{ij} = -Y^{ji}$, инвариантный относительно полной группы вращений $\mathbb{O}(3)$, определяется одной псевдоскалярной функцией $f^*(r)$ от радиуса:

$$Y^{ij} = \varepsilon^{ijk}x_k f^*(r),$$

где ε^{ijk} – полностью антисимметричный псевдотензор третьего ранга. При отражении одной из декартовых осей координат ε^{ijk} и f^* меняют знаки, оставляя антисимметричный тензор без изменения.

Дифференциал телесного угла (7.8) задает метрику на сфере единичного радиуса с центром в начале координат. С топологической точки зрения евклидово пространство $\mathbb{R}^3$ с выколотым началом координат является прямым произведением сферы $\mathbb{S}^2$ и положительной полуоси $\mathbb{R}_+$:

$$\mathbb{R}^3 \setminus \{0\} = \mathbb{S}^2 \times \mathbb{R}_+.$$

Приведем явные формулы преобразования частных производных (координатного базиса векторных полей), которые часто используются в приложениях:

$$\begin{aligned} \partial_r &= \sin \theta \cos \varphi \partial_x + \sin \theta \sin \varphi \partial_y + \cos \theta \partial_z, \\ \partial_\theta &= r \cos \theta \cos \varphi \partial_x + r \cos \theta \sin \varphi \partial_y - r \sin \theta \partial_z, \\ \partial_\varphi &= -r \sin \theta \sin \varphi \partial_x + r \sin \theta \cos \varphi \partial_y, \end{aligned}$$

и их обратные:

$$\begin{aligned}\partial_x &= \sin \theta \cos \varphi \partial_r + \frac{\cos \theta \cos \varphi}{r} \partial_\theta - \frac{\sin \varphi}{r \sin \theta} \partial_\varphi, \\ \partial_y &= \sin \theta \sin \varphi \partial_r + \frac{\cos \theta \sin \varphi}{r} \partial_\theta + \frac{\cos \varphi}{r \sin \theta} \partial_\varphi, \\ \partial_z &= \cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta.\end{aligned}$$

В приложениях тензорные поля часто рассматриваются в ортонормированном базисе векторных полей и 1-форм, который мы пронумеруем индексами со шляпкой:

$$\begin{aligned}e_{\hat{r}} &:= \partial_r, & e_{\hat{\theta}} &:= \frac{1}{r} \partial_\theta, & e_{\hat{\varphi}} &:= \frac{1}{r \sin \theta} \partial_\varphi, \\ e^{\hat{r}} &= dr, & e^{\hat{\theta}} &= r d\theta, & e^{\hat{\varphi}} &= r \sin \theta d\varphi,\end{aligned}\tag{7.11}$$

что следует непосредственно из вида метрики (7.7). Этому базису $e_a = e^\alpha_a \partial_\alpha$, где $\alpha = (r, \theta, \varphi)$ и $a = (\hat{r}, \hat{\theta}, \hat{\varphi})$, соответствуют следующие компоненты ортонормированного репера и его обратного:

$$\begin{aligned}e_r^{\hat{r}} &= 1, & e_\theta^{\hat{\theta}} &= r, & e_\varphi^{\hat{\varphi}} &= r \sin \theta, \\ e^r_{\hat{r}} &= 1, & e^\theta_{\hat{\theta}} &= \frac{1}{r}, & e^\varphi_{\hat{\varphi}} &= \frac{1}{r \sin \theta}.\end{aligned}\tag{7.12}$$

При этом все недиагональные компоненты репера равны нулю. Здесь и далее шляпками обозначены индексы относительно этого репера. Векторное поле можно разложить как по координатному, так и по ортонормированному базису:

$$X = X^r \partial_r + X^\theta \partial_\theta + X^\varphi \partial_\varphi = X^{\hat{r}} e_{\hat{r}} + X^{\hat{\theta}} e_{\hat{\theta}} + X^{\hat{\varphi}} e_{\hat{\varphi}},$$

причем переход от компонент X^α к $X^{\hat{\alpha}}$ осуществляется с помощью репера:

$$X^r = X^{\hat{r}}, \quad X^\theta = \frac{1}{r} X^{\hat{\theta}}, \quad X^\varphi = \frac{1}{r \sin \theta} X^{\hat{\varphi}}.\tag{7.13}$$

Отметим, что, поскольку репер ортонормирован, то компоненты векторных полей и соответствующих 1-форм совпадают:

$$X_{\hat{r}} = X^{\hat{r}}, \quad X_{\hat{\theta}} = X^{\hat{\theta}}, \quad X_{\hat{\varphi}} = X^{\hat{\varphi}}.$$

Получим явные формулы для градиента, ротора, дивергенции и лапласиана в сферической системе координат, которые используются в задачах со сферической симметрией. Чтобы получить явное выражение для инвариантных дифференциальных операторов в сферической системе координат нам понадобятся символы Кристоффеля (6.24). Они имеют девять ненулевых компонент:

$$\begin{aligned}\Gamma_{r\theta}^\theta &= \Gamma_{\theta r}^\theta = \frac{1}{r}, & \Gamma_{\theta\varphi}^\varphi &= \Gamma_{\varphi\theta}^\varphi = \frac{\cos \theta}{\sin \theta}, \\ \Gamma_{r\varphi}^\varphi &= \Gamma_{\varphi r}^\varphi = \frac{1}{r}, & \Gamma_{\varphi\varphi}^r &= -r \sin^2 \theta, \\ \Gamma_{\theta\theta}^r &= -r, & \Gamma_{\varphi\varphi}^\theta &= -\sin \theta \cos \theta.\end{aligned}\tag{7.14}$$

Из условия равенства нулю ковариантной производной репера (7.12)

$$\nabla_\alpha e_\beta^a = \partial_\alpha e_\beta^a - \Gamma_{\alpha\beta}^\gamma e_\gamma^a + \omega_{\alpha b}^a e_\beta^b = 0, \quad (7.15)$$

где $\omega_{ab}^a - \mathbb{SO}(3)$ связность (без кручения), находим, что у $\mathbb{SO}(3)$ связности только 6 компонент отличны от нуля:

$$\begin{aligned} \omega_{\theta\hat{r}}^{\hat{\theta}} &= -\omega_{\theta\hat{\theta}}^{\hat{r}} = 1, \\ \omega_{\varphi\hat{r}}^{\hat{\varphi}} &= -\omega_{\varphi\hat{\varphi}}^{\hat{r}} = \sin \theta, \\ \omega_{\varphi\hat{\theta}}^{\hat{\varphi}} &= -\omega_{\varphi\hat{\varphi}}^{\hat{\theta}} = \cos \theta, \end{aligned} \quad (7.16)$$

где $\alpha, \beta = (r, \theta, \varphi)$ и $a, b = (\hat{r}, \hat{\theta}, \hat{\varphi})$.

Градиент скалярного поля f дает 1-форму:

$$df = e^{\hat{r}} \partial_r f + e^{\hat{\theta}} \frac{1}{r} \partial_\theta f + e^{\hat{\varphi}} \frac{1}{r \sin \theta} \partial_\varphi f. \quad (7.17)$$

Для получения этой формулы в дифференциале $df = dr \partial_r f + d\theta \partial_\theta f + d\varphi \partial_\varphi f$ дифференциалы dr , $d\theta$ и $d\varphi$ необходимо заменить на ортонормированный базис $e^{\hat{r}}$, $e^{\hat{\theta}}$ и $e^{\hat{\varphi}}$ с помощью соотношений (7.11).

Дивергенция вектора (6.63) дает функцию

$$\nabla_\alpha X^\alpha = \frac{1}{r^2} \partial_r (r^2 X^{\hat{r}}) + \frac{1}{r \sin \theta} \partial_\theta (\sin \theta X^{\hat{\theta}}) + \frac{1}{r \sin \theta} \partial_\varphi X^{\hat{\varphi}}. \quad (7.18)$$

Эта и последующие формулы получаются подстановкой символов Кристоффеля (7.14) в ковариантные производные и выражением компонент вектора в координатном базисе через его компоненты в ортонормированном базисе (7.13).

В инвариантном виде ротор 1-формы $A = dx^\alpha A_\alpha$ определяется оператором Ходжа (3.50) и внешним дифференцированием (3.34):

$$\text{rot } A = *dA := e^{\hat{r}} B_{\hat{r}} + e^{\hat{\theta}} B_{\hat{\theta}} + e^{\hat{\varphi}} B_{\hat{\varphi}}, \quad (7.19)$$

где компоненты разложения задаются формулами:

$$\begin{aligned} B_{\hat{r}} &= B^{\hat{r}} = \frac{1}{r \sin \theta} \partial_\theta (\sin \theta A_{\hat{\varphi}}) - \frac{1}{r \sin \theta} \partial_\varphi A_{\hat{\theta}}, \\ B_{\hat{\theta}} &= B^{\hat{\theta}} = \frac{1}{r \sin \theta} \partial_\varphi A_{\hat{r}} - \frac{1}{r} \partial_r (r A_{\hat{\varphi}}), \\ B_{\hat{\varphi}} &= B^{\hat{\varphi}} = \frac{1}{r} \partial_r (r A_{\hat{\theta}}) - \frac{1}{r} \partial_\theta A_{\hat{r}}. \end{aligned} \quad (7.20)$$

Лапласиан функции (6.70) дает также функцию:

$$\Delta f = \frac{1}{r^2} \partial_r (r^2 \partial_r f) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta f) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 f. \quad (7.21)$$

Лапласиан 1-формы снова дает 1-форму: $\Delta A_{\hat{\alpha}} = g^{\beta\gamma} \nabla_\beta \nabla_\gamma A_{\hat{\alpha}}$. После громоздких

вычислений можно получить явные выражения для компонент:

$$\begin{aligned}
\Delta A_{\hat{r}} &= \frac{1}{r^2} \partial_r (r^2 \partial_r A_{\hat{r}}) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta A_{\hat{r}}) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 A_{\hat{r}} - \frac{2}{r^2} A_{\hat{r}} - \\
&\quad - \frac{2}{r^2 \sin \theta} \partial_\theta (\sin \theta A_{\hat{\theta}}) - \frac{2}{r^2 \sin \theta} \partial_\varphi A_{\hat{\varphi}}, \\
\Delta A_{\hat{\theta}} &= \frac{1}{r} \partial_r^2 (r A_{\hat{\theta}}) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta A_{\hat{\theta}}) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 A_{\hat{\theta}} - \frac{1}{r^2 \sin^2 \theta} A_{\hat{\theta}} + \\
&\quad + \frac{2}{r^2} \partial_\theta A_{\hat{r}} - \frac{2 \cos \theta}{r^2 \sin^2 \theta} \partial_\varphi A_{\hat{\varphi}}, \\
\Delta A_{\hat{\varphi}} &= \frac{1}{r} \partial_r^2 (r A_{\hat{\varphi}}) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta A_{\hat{\varphi}}) + \frac{1}{r^2 \sin^2 \theta} \partial_\varphi^2 A_{\hat{\varphi}} - \frac{1}{r^2 \sin^2 \theta} A_{\hat{\varphi}} + \\
&\quad + \frac{2 \cos \theta}{r^2 \sin^2 \theta} \partial_\varphi A_{\hat{\theta}} + \frac{2}{r^2 \sin \theta} \partial_\varphi A_{\hat{r}}.
\end{aligned} \tag{7.22}$$

Отметим появление дополнительных слагаемых по сравнению с лапласианом функции (7.21).

Приведенные выше формулы для дифференциальных операторов можно также получить путем элементарных, но громоздких геометрических построений. Такой подход эквивалентен введенному ранее понятию ковариантной производной. Это показывает, что дифференциальная геометрия предоставляет естественный и конструктивный подход к получению явного вида дифференциальных операторов в различных криволинейных системах координат. Разумеется, геометрия пространства при этом остается евклидовой.

7.2 Цилиндрические координаты

Рассмотрим *цилиндрические координаты* r, φ, z в трехмерном евклидовом пространстве $\mathbb{R}^3$, рис. 7.2. При переходе к цилиндрическим координатам мы вводим полярные координаты на плоскости x, y , оставляя координату z без изменений. Функции перехода и их обратные имеют вид (за третьей координатой z можно не следить):

$$\begin{aligned}
x &= r \cos \varphi, & r &= \sqrt{x^2 + y^2}, \\
y &= r \sin \varphi, & \varphi &= \arctg(y/x).
\end{aligned}$$

Угол φ , как и в случае сферических координат, называется *полярным*. Якобиан преобразования координат равен $J = r$ и является вырожденным в начале полярных координат, что соответствует оси z . Область определения цилиндрических координат задается неравенствами:

$$0 < r < \infty, \quad 0 < \varphi < 2\pi, \quad -\infty < z < \infty,$$

что соответствует евклидову пространству с удаленной полуплоскостью $y = 0, x \geq 0$.

Дифференциалы декартовых и полярных координат (координатные базисы 1-форм) связаны преобразованием:

$$\begin{aligned}
dx &= \cos \varphi dr - r \sin \varphi d\varphi, & dr &= \cos \varphi dx + \sin \varphi dy, \\
dy &= \sin \varphi dr + r \cos \varphi d\varphi, & d\varphi &= -\frac{\sin \varphi}{r} dx + \frac{\cos \varphi}{r} dy.
\end{aligned} \tag{7.23}$$

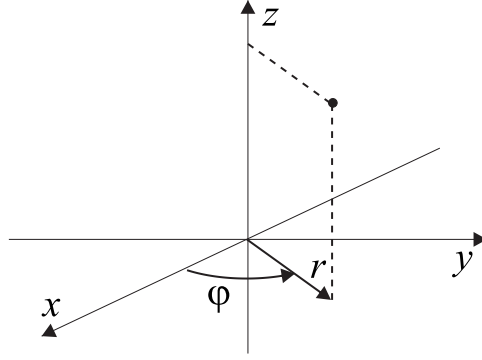


Рис. 7.2: Цилиндрические координаты r, φ, z в трехмерном евклидовом пространстве $\mathbb{R}^3$.

Частные производные в декартовых и полярных координатах (координатные базисы векторных полей) связаны простыми соотношениями:

$$\begin{aligned} \partial_x &= \cos \varphi \partial_r - \frac{\sin \varphi}{r} \partial_\varphi, & \partial_r &= \cos \varphi \partial_x + \sin \varphi \partial_y, \\ \partial_y &= \sin \varphi \partial_r + \frac{\cos \varphi}{r} \partial_\varphi, & \partial_\varphi &= -r \sin \varphi \partial_x + r \cos \varphi \partial_y. \end{aligned} \quad (7.24)$$

Прямая проверка показывает, что справедлива формула

$$(xdx + ydy)^2 = r^2 dr^2.$$

Соотношения (7.23) приводят к следующему выражению для евклидовой метрики (7.6) в цилиндрических координатах:

$$ds^2 = dr^2 + r^2 d\varphi^2 + dz^2.$$

В матричных обозначениях метрика и ее обратная имеют вид

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad g^{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Из символов Кристоффеля только три отличны от нуля:

$$\Gamma_{r\varphi}^\varphi = \Gamma_{\varphi r}^\varphi = \frac{1}{r}, \quad \Gamma_{\varphi\varphi}^r = -r.$$

Ортонормированный базис векторных полей и 1-форм имеет вид (для полярных координат)

$$\begin{aligned} e_{\hat{r}} &:= \partial_r, & e_{\hat{\varphi}} &:= \frac{1}{r} \partial_\varphi, \\ e^{\hat{r}} &= dr, & e^{\hat{\varphi}} &= r d\varphi. \end{aligned}$$

Этому базису соответствует ортонормированный репер $e_a = e^\alpha_a \partial_\alpha$ с компонентами

$$\begin{aligned} e_r^{\hat{r}} &= 1, & e_\varphi^{\hat{\varphi}} &= r, \\ e^r_{\hat{r}} &= 1, & e^\varphi_{\hat{\varphi}} &= \frac{1}{r}, \end{aligned}$$

где $\alpha = (r, \varphi)$ и $a = (\hat{r}, \hat{\varphi})$. При этом все недиагональные компоненты равны нулю.

Из условия (7.15) находим компоненты $\mathbb{SO}(2)$ связности. Несложные вычисления показывают, что только две компоненты отличны от нуля:

$$\omega_{\varphi\hat{r}}^{\hat{\varphi}} = -\omega_{\varphi\hat{\varphi}}^{\hat{r}} = 1.$$

Вычисления, аналогичные случаю сферических координат, приводят к следующим выражениям для простейших дифференциальных операторов в цилиндрических координатах:

$$df = e^{\hat{r}}\partial_r f + e^{\hat{\varphi}}\frac{1}{r}\partial_\varphi f + e^{\hat{z}}\partial_z f, \quad (7.25)$$

$$\nabla_\alpha X^\alpha = \frac{1}{r}\partial_r(rX^{\hat{r}}) + \frac{1}{r}\partial_\varphi X^{\hat{\varphi}} + \partial_z X^{\hat{z}}, \quad (7.26)$$

$$\text{rot } A = e^{\hat{r}}\left[\frac{1}{r}\partial_\varphi A_{\hat{z}} - \partial_z A_{\hat{\varphi}}\right] + e^{\hat{\varphi}}[\partial_z A_{\hat{r}} - \partial_r A_{\hat{z}}] + e^{\hat{z}}\frac{1}{r}[\partial_r(rA_{\hat{\varphi}}) - \partial_\varphi A_{\hat{r}}], \quad (7.27)$$

$$\Delta f = \frac{1}{r}\partial_r(r\partial_r f) + \frac{1}{r^2}\partial_{\varphi\varphi}^2 f + \partial_{zz}^2 f, \quad (7.28)$$

$$\Delta A_{\hat{r}} = \frac{1}{r}\partial_r(r\partial_r A_{\hat{r}}) + \frac{1}{r^2}\partial_{\varphi\varphi}^2 A_{\hat{r}} + \partial_{zz}^2 A_{\hat{r}} - \frac{1}{r^2}A_{\hat{r}} - \frac{2}{r^2}\partial_\varphi A_{\hat{\varphi}}, \quad (7.29)$$

$$\Delta A_{\hat{\varphi}} = \frac{1}{r}\partial_r(r\partial_r A_{\hat{\varphi}}) + \frac{1}{r^2}\partial_{\varphi\varphi}^2 A_{\hat{\varphi}} + \partial_{zz}^2 A_{\hat{\varphi}} - \frac{1}{r^2}A_{\hat{\varphi}} + \frac{2}{r^2}\partial_\varphi A_{\hat{r}}, \quad (7.30)$$

$$\Delta A_{\hat{z}} = \frac{1}{r}\partial_r(r\partial_r A_{\hat{z}}) + \frac{1}{r^2}\partial_{\varphi\varphi}^2 A_{\hat{z}} + \partial_{zz}^2 A_{\hat{z}}, \quad (7.31)$$

где $e^{\hat{z}} := dz$. Напомним, что шляпки над символами означают, что компоненты (ко)векторов рассматриваются относительно ортонормированного базиса.

Глава 8

Группы Ли

Группы Ли образуют один из наиболее важных классов многообразий и имеют широкие приложения в математической физике. В настоящем разделе мы дадим формальные определения, опишем основные свойства, а также рассмотрим группы Ли с дифференциально геометрической точки зрения. В частности, многообразия полупростых групп Ли будут рассмотрены, как римановы пространства и пространства абсолютного параллелизма, когда групповая операция отождествляется с параллельным переносом. В конце главы без доказательств приведена классификация простых групп Ли.

8.1 Группы Ли и локальные группы Ли

Определение. Множество элементов $a, b, c, \dots \in \mathbb{G}$ называется *группой Ли*, если оно является одновременно и группой, и гладким многообразием, при этом требуется, чтобы была определена гладкая групповая операция

$$\mathbb{G} \times \mathbb{G} \ni (a, b) \mapsto ab \in \mathbb{G}, \quad \forall a, b \in \mathbb{G}, \quad (8.1)$$

со следующими свойствами:

- 1) Ассоциативность: $(ab)c = a(bc), \quad \forall a, b, c \in \mathbb{G}$.
- 2) Существование единицы $e \in \mathbb{G}$: $ea = ae = a, \quad \forall a \in \mathbb{G}$.
- 3) Для каждого элемента $a \in \mathbb{G}$ существует обратный элемент $a^{-1} \in \mathbb{G}$ такой, что $a^{-1}a = aa^{-1} = e$, причем отображение $\mathbb{G} \ni a \mapsto a^{-1} \in \mathbb{G}$ гладкое. $\square$

Замечание. В определении оба условия гладкости отображений:

$$\mathbb{G} \times \mathbb{G} \ni (a, b) \mapsto ab \in \mathbb{G}, \quad (8.2)$$

$$\mathbb{G} \ni a \mapsto a^{-1} \in \mathbb{G}, \quad (8.3)$$

можно объединить в эквивалентное условие гладкости одного отображения:

$$\mathbb{G} \times \mathbb{G} \ni (a, b) \mapsto ab^{-1} \in \mathbb{G}. \quad (8.4)$$

То, что (8.4) следует из (8.2) и (8.3), очевидно. Обратно. Из (8.4) при $a = e$ следует гладкость отображения (8.3). Гладкость отображения (8.2) следует из (8.4) после гладкого отображения $b^{-1} \mapsto b$. $\square$

Каждому элементу группы $b \in \mathbb{G}$ поставим в соответствие отображение

$$\mathbb{G} \ni a \mapsto ab \in \mathbb{G}, \quad \forall b \in \mathbb{G}. \quad (8.5)$$

Тогда из определения группы Ли следует, что это отображение взаимно однозначно и гладко, т.е. является диффеоморфизмом $\mathbb{G} \xrightarrow{b} \mathbb{G}$. Следовательно, отображение (8.1) можно рассматривать как группу преобразований многообразия $\mathbb{G}$.

Перефразируем определение: группой Ли является группа, снабженная гладкой структурой многообразия. При этом не всякая группа допускает существование такой структуры.

Замечание. Если на произвольной группе задать дискретную топологию, то она будет группой Ли, рассматриваемой, как 0-мерное многообразие (если отбросить требование счетности базы топологии). В дальнейшем, чтобы исключить подобные ситуации, при рассмотрении групп Ли мы не будем рассматривать 0-мерные многообразия. Поэтому множество элементов группы Ли не может быть счетным, т.к. множество точек области в $\mathbb{R}^n$ несчетно. $\square$

Пример 8.1.1. Группа перестановок конечного числа элементов не является группой Ли, т.к. содержит конечное число элементов. $\square$

Можно сказать также, что группой Ли является многообразие, на котором задана групповая операция. При этом не на всяком многообразии можно определить гладкую бинарную операцию, удовлетворяющую групповым аксиомам.

Пример 8.1.2. Двумерная сфера $\mathbb{S}^2$ не может быть наделена групповой структурой. Действительно, зафиксируем произвольный отличный от нуля вектор в касательном пространстве к единице группы $T_e(\mathbb{G})$. С помощью дифференциала отображения (8.5) этот вектор можно разнести по всему групповому многообразию. В результате получим гладкое векторное поле на $\mathbb{G}$, которое всюду отлично от нуля, т.к. отображение (8.5) является диффеоморфизмом. Это противоречит теореме 10.2.1, утверждающей, что на двумерной сфере не существует непрерывного отличного от нуля векторного поля. $\square$

Поскольку группы Ли являются многообразиями, то на них без изменения переносится большинство характеристик многообразий.

Определение. *Размерностью* группы Ли называется ее размерность как многообразие. Группа Ли называется *компактной*, если групповое многообразие компактно. $\square$

Из непрерывности групповой операции следует, что, если подмножество $U \subset \mathbb{G}$ – открыто или замкнуто в $\mathbb{G}$, то таковым же будет подмножество U^{-1} , состоящее из обратных элементов, а также подмножества aU и Ua для всех $a \in \mathbb{G}$. Отсюда вытекает, что для открытого подмножества U подмножества $\forall U$ и UV также открыты в $\mathbb{G}$ каково бы ни было подмножество $V \in \mathbb{G}$.

Замечание. В определении группы Ли мы потребовали гладкость т.е. бесконечную дифференцируемость групповой операции и, соответственно, гладкость группы Ли, как многообразия. На самом деле ситуация следующая. Согласно теореме Глисона–Монтгомери–Циппина [59, 60] на всякой непрерывной группе (класса $\mathcal{C}^0$) можно ввести структуру аналитического многообразия (класса $\mathcal{C}^\omega$), совместимую с

групповой структурой. Эта теорема дает положительное решение пятой проблемы Гильберта: следует ли из существования в группе $\mathbb{G}$ каких-нибудь координат существование в ней дифференцируемых координат. Поэтому, не ограничивая общности, можно считать, что группы Ли являются вещественно аналитическими многообразиями, а групповая операция – вещественно аналитическое отображение. $\square$

Таким образом, группы Ли представляют собой специальный класс групп, для которых групповая операция может быть записана в координатной форме. Это важное свойство позволяет применить методы математического анализа для исследования свойств групп Ли.

Приведем несколько примеров групп Ли.

Пример 8.1.3. Прямая $\mathbb{R}$, на которой определены сдвиги (сложение) $a + b = c$, где $a, b, c \in \mathbb{R}$, является одномерной связной односвязной некомпактной абелевой группой Ли. $\square$

Пример 8.1.4. Проколтая прямая $\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$ является группой Ли по отношению к умножению, $xy = z$, где $x, y, z \in \mathbb{R}^\times$. Эта группа одномерна абелева и состоит из двух компонент связности. Каждая компонента связности некомпактна. Связная компонента единицы $\mathbb{R}_+$ (положительные числа) изоморфна группе вещественных чисел по сложению, $\mathbb{R}_+ \simeq \mathbb{R}$. Изоморфизм задается показательной функцией с произвольным положительным основанием, отличным от единицы. Например, $x = e^a$, где $a \in \mathbb{R}$. $\square$

Пример 8.1.5. Евклидово пространство $\mathbb{R}^n$ (как и произвольное векторное пространство) является n -мерной связной односвязной некомпактной абелевой группой Ли относительно сложения векторов. $\square$

Пример 8.1.6. Комплексная плоскость с выколотым началом координат $\mathbb{C}^\times := \mathbb{C} \setminus \{0\}$ является комплексной группой Ли по отношению к обычному умножению комплексных чисел. Ее комплексная размерность равна единице. Эта группа некомпактна связна, но не односвязна. Ее фундаментальная группа равна $\mathbb{Z}$ (число обходов вокруг начала координат). $\square$

Пример 8.1.7. Единичная окружность $\mathbb{S}^1$, точками которой являются комплексные числа $z = e^{i\varphi}$ является абелевой компактной группой Ли по отношению к умножению. Она обозначается через $U(1)$ (унитарные матрицы размера 1×1) и изоморфна группе собственных двумерных вращений $SO(2)$. Эта группа связна, но не односвязна. Ее фундаментальная группа изоморфна группе целых чисел $\mathbb{Z}$. $\square$

Пример 8.1.8. Трехмерная сфера $\mathbb{S}^3$ в пространстве кватернионов, точками которой являются кватернионы q , для которых $|q| = 1$, может быть снабжена групповой структурой (см. раздел 1.8). Это – трехмерная связная односвязная неабелева компактная группа Ли, которая изоморфна группе унитарных матриц размерности 2×2 с единичным определителем $SU(2)$. $\square$

Можно показать, что сферы только двух размерностей $\mathbb{S}^1$ и $\mathbb{S}^3$ исчерпывают все сферы на которых можно задать структуру группы Ли.

В рассмотренных примерах есть две одномерные абелевы группы Ли: $\mathbb{R}$ и $\mathbb{S}^1$. Любая связная одномерная группа Ли является абелевой, и изоморфна либо вещественной прямой $\mathbb{R}$, либо окружности $\mathbb{S}^1$. Этот результат имеет широкие применения, т.к. каждая группа Ли имеет одномерные подгруппы. Если одномерная группа Ли не

является связной, то она может быть неабелевой. Примеры дают группа вращений $\mathbb{O}(2)$ и группа Лоренца $\mathbb{O}(1, 1)$ (см. разделы 1.6 и 1.7).

В случае связных одномерных групп Ли можно описать все возможные гомоморфизмы между ними:

$$\begin{aligned}\mathbb{R} \ni x &\mapsto ax && \in \mathbb{R}, \quad a \in \mathbb{R}; \\ \mathbb{R} \ni x &\mapsto e^{iax} && \in \mathbb{S}^1, \quad a \in \mathbb{R}; \\ \mathbb{S}^1 \ni z &\mapsto 0 && \in \mathbb{R}; \\ \mathbb{S}^1 \ni z &\mapsto z^n && \in \mathbb{S}^1, \quad n \in \mathbb{Z};\end{aligned}$$

где x – точка вещественной прямой, z – комплексное число, равное по модулю единице, а числа a и n параметризуют гомоморфизмы.

Прямое произведение групп Ли $\mathbb{G}_1 \times \mathbb{G}_2$ является группой Ли с дифференцируемой структурой прямого произведения многообразий и групповой структурой прямого произведения групп.

Пример 8.1.9. Тор $\mathbb{T}^n := \underbrace{\mathbb{S}^1 \times \dots \times \mathbb{S}^1}_n \approx \underbrace{\mathbb{U}(1) \times \dots \times \mathbb{U}(1)}_n$ для всех n является n -мерной абелевой компактной группой Ли как прямое произведение окружностей. Эта группа связна, но не односвязна. $\square$

Предложение 8.1.1. Всякая компактная связная абелева группа Ли $\mathbb{G}$ размерности n изоморфна n -мерному тору $\mathbb{T}^n$. $\square$

Доказательство. См., например, [42], глава 3, упражнение 18. $\square$

С топологической точки зрения группы Ли $\mathbb{G}$ могут состоять из нескольких компонент связности¹. При этом каждая компонента имеет одинаковую размерность, равную размерности группы. Связная компонента единицы $\mathbb{G}_0$ является открытой подгруппой в $\mathbb{G}$ и представляет собой нормальный делитель. Остальные компоненты представляют собой смежные классы по этой подгруппе, и фактор группа $\mathbb{G}/\mathbb{G}_0$ состоит не более, чем из счетного числа элементов. Если группу Ли рассматривать, как группу преобразований, то каждая связная компонента представляет собой орбиту произвольного элемента из этой компоненты относительно действия элементов группы из связной компоненты единицы $\mathbb{G}_0$. Ясно, что все компоненты связности группы Ли диффеоморфны между собой.

Теперь перейдем к координатному описанию. Поскольку группа Ли является многообразием, то групповое умножение можно записать в координатах. Рассмотрим связную компоненту единицы $\mathbb{G}_0$ группы Ли размерности N . Пусть в окрестности единицы задана некоторая система координат $\{a^A\}$, $A = 1, \dots, N$. Не ограничивая общности, будем считать, что начало координат $a^A = 0$ совпадает с единицей группы. Тогда групповое умножение

$$c = ab = f(a, b), \quad a, b, c \in \mathbb{G}_0, \quad (8.6)$$

и групповые аксиомы (см. раздел 1.2) можно записать в координатном виде (в до-

¹Поскольку мы предполагаем, что многообразие обладает счетной базой, то число компонент связности у групп Ли может быть не более, чем счетно.

статочной малой окрестности единицы):

- 1) $c^\wedge = f^\wedge(a, b)$ – закон композиции,
 - 2) $f^\wedge(a, f(b, c)) = f^\wedge(f(a, b), c)$ – ассоциативность,
 - 3) $a^\wedge = f^\wedge(a, 0) = f^\wedge(0, a)$ – существование единицы,
 - 4) $f^\wedge(a, a^{-1}) = f^\wedge(a^{-1}, a) = 0$ – существование обратного элемента.
- (8.7)

Набор N функций $f^\wedge$ от $2N$ переменных, удовлетворяющих условиям (8.7), называется *функцией композиции* для группы Ли $\mathbb{G}$. Она определена в некоторой окрестности единицы группы.

Связная компонента группы Ли $\mathbb{G}_0$ может оказаться нетривиальным многообразием и не покрываться одной картой. Поэтому введем новое понятие.

Определение. *Локальной группой Ли* называется пара $(\mathbb{U}, f)$, где $\mathbb{U} \subset \mathbb{R}^N$ – область евклидова пространства, содержащая начало координат, и f – гладкое отображение $\mathbb{U} \times \mathbb{U} \rightarrow \mathbb{R}^N$ такое, что выполнены условия (8.7). При этом мы предполагаем, что все равенства выполнены в той области, где определены все выражения, входящие в данное равенство. $\square$

Замечание. В определении локальной группы Ли мы не предполагаем замкнутости области $\mathbb{U}$ относительно группового умножения, т.е. возможно существование таких точек в $\mathbb{U}$, что их произведение не лежит в $\mathbb{U}$. Конечно, для таких точек про ассоциативность в определении ничего не говорится. $\square$

Каждая группа Ли $\mathbb{G}$ порождает бесконечное множество локальных групп Ли следующим образом. Пусть $(\mathbb{V}, \varphi)$ – карта на $\mathbb{G}$, содержащая единицу, которая отображается в начало координат евклидова пространства $\mathbb{R}^N$. Возьмем такую окрестность $\mathbb{U} \subset \mathbb{V}$, что выполнены включения $\mathbb{U} \cdot \mathbb{U} \subset \mathbb{V}$ и $\mathbb{U}^{-1} \subset \mathbb{V}$, где под произведением $\mathbb{U} \cdot \mathbb{U}$ мы понимаем множество всех произведений ab , где $a, b \in \mathbb{U}$. Существование такой окрестности следует из непрерывности отображений $(a, b) \mapsto ab$ и $a \mapsto a^{-1}$. Тогда каждая пара $(\mathbb{U}, f)$, где мы отождествили область $\mathbb{U} \subset \mathbb{G}$ с ее образом $\varphi(\mathbb{U}) \in \mathbb{R}^N$, а f – функция композиции на $\mathbb{G}$, записанная в координатах, является локальной группой Ли.

Определение. Две локальные группы Ли $(\mathbb{U}_1, f_1)$ и $(\mathbb{U}_2, f_2)$ называются *изоморфными*, если существуют такие окрестности нуля $\mathbb{U}'_1 \subset \mathbb{U}_1$ и $\mathbb{U}'_2 \subset \mathbb{U}_2$ и такой диффеоморфизм $h : \mathbb{U}'_1 \rightarrow \mathbb{U}'_2$, что диаграмма

$$\begin{array}{ccc} \mathbb{U}'_1 \times \mathbb{U}'_1 & \xrightarrow{f_1} & \mathbb{R}^N \supset \mathbb{U}'_1 \\ \downarrow h \times h & & \downarrow h \\ \mathbb{U}'_2 \times \mathbb{U}'_2 & \xrightarrow{f_2} & \mathbb{R}^N \supset \mathbb{U}'_2 \end{array}$$

коммутативна, т.е. $h(f_1(a, b)) = f_2(h(a), h(b))$ при $a, b \in \mathbb{U}'_1$, там, где обе части равенства определены. $\square$

Ясно, что все локальные группы Ли, полученные описанным выше способом из одной группы Ли, изоморфны между собой. Поэтому введем

Определение. Две группы Ли $\mathbb{G}_1$ и $\mathbb{G}_2$ называются *локально изоморфными*, если порождаемые ими локальные группы Ли изоморфны. $\square$

Конечно, две изоморфные группы Ли являются также локально изоморфными. Нетривиальность данного выше определения заключается в том, что существуют локально изоморфные группы Ли, которые в то же время не изоморфны между собой.

Пример 8.1.10. Пусть $\mathbb{G}_1 = \mathbb{R}$ и $\mathbb{G}_2 = \mathbb{S}^1$. В качестве $U'_1 \subset \mathbb{R}$ возьмем интервал $(-\pi, \pi)$, а в качестве $U'_2 \subset \mathbb{S}^1$ – дополнение к точке $-1 \in \mathbb{S}^1$. Отображение

$$h: \mathbb{R} \supset U'_1 \ni x \mapsto e^{ix} \in U'_2 \subset \mathbb{S}^1$$

устанавливает диффеоморфизм между U'_1 и U'_2 , перестановочный с умножением. Таким образом, группы Ли $\mathbb{R}$ и $\mathbb{S}^1$ локально изоморфны. В то же время они не изоморфны между собой. $\square$

Предложение 8.1.2. *Всякая N -мерная абелева группа Ли локально изоморфна группе трансляций N -мерного векторного пространства, т.е. евклидову пространству $\mathbb{R}^N$, рассматриваемому как группа сдвигов.*

Доказательство. См., например, [61], глава 10, §56 D. $\square$

Структура группы Ли накладывает жесткие условия на функцию композиции f^Λ . Как уже было отмечено, функция композиции $f^\Lambda(a, b)$ вещественно аналитична (т.е. разлагается в сходящиеся степенные ряды) во всех точках группового многообразия по всем $2N$ переменным.

Пример 8.1.11. Рассмотрим группу невырожденных вещественных $n \times n$ матриц $\mathrm{GL}(n, \mathbb{R})$ с обычным правилом умножения. Произвольную матрицу $M \in \mathrm{GL}(n, \mathbb{R})$ можно представить в виде

$$M = \mathbb{1} + (X_a^b), \quad a, b = 1, \dots, n,$$

где $\mathbb{1}$ – единичная матрица и X_a^b – набор n^2 чисел (элементы $n \times n$ матрицы), которые мы примем за координаты матрицы M . Полученное таким образом отображение всей группы $\mathrm{GL}(n, \mathbb{R})$ на область евклидова пространства $\mathbb{R}^{n^2}$ переводит единицу группы в начало координат. Групповое многообразие при этом представляет собой открытую область в $\mathbb{R}^{n^2}$, которая является дополнением замкнутого множества точек, определяемого уравнением $\det M = 0$. Функция композиции для группы $\mathrm{GL}(n, \mathbb{R})$ принимает вид

$$f_a^b(M, N) = X_a^b + Y_a^b + X_a^c Y_c^b,$$

где $N = \mathbb{1} + Y$. Таким образом, функция композиции является аналитической функцией. Аналогично, множество невырожденных $n \times n$ комплексных матриц $\mathrm{GL}(n, \mathbb{C})$ образует аналитическую группу Ли вещественной размерности $2n^2$. $\square$

Проанализируем функцию композиции. Из условия существования единицы (8.7) следуют равенства:

$$\left. \frac{\partial f^\Lambda(a, b)}{\partial a^B} \right|_{b=0} = \delta_B^A, \quad \left. \frac{\partial f^\Lambda(a, b)}{\partial b^B} \right|_{a=0} = \delta_B^A.$$

Из этих уравнений и дифференцируемости функции композиции вытекает, что вблизи единицы группы функция композиции разрешима относительно координат a^A и b^A соответствующих элементов группы cb^{-1} и $a^{-1}c$, где $c = f(a, b)$. При этом класс дифференцируемости функций $a^A(c, b)$ и $b^A(a, c)$ такой же, как и у функции композиции (бесконечный).

Разложим функцию композиции в ряд Тейлора в окрестности единицы группы с точностью до членов третьего порядка

$$f^A(a, b) = a^A + b^A + a^C b^B \hat{f}_{BC}^A + a^D a^C b^B g_{BCD}^A + a^D b^C b^B h_{BCD}^A + \dots \quad (8.8)$$

В этом разложении $\hat{f}_{BC}^A$, $g_{BCD}^A = g_{BDC}^A$, $h_{BCD}^A = h_{CBD}^A$ — некоторые постоянные. Из существования единичного элемента следует отсутствие нулевого члена разложения и вид линейных слагаемых. Слагаемые более высоких порядков обязательно должны содержать смешанное произведение $a^A b^B$, т.к. в противном случае условие 3) в определении функции композиции не может быть выполнено. Подстановка разложения (8.8) в условие ассоциативности накладывает условие на коэффициенты разложения:

$$\hat{f}_{BC}^E \hat{f}_{ED}^A + 2h_{BCD}^A = \hat{f}_{CD}^E \hat{f}_{BE}^A + 2g_{BCD}^A. \quad (8.9)$$

При антисимметризации этого выражения по индексам B, C, D слагаемые h_{BCD}^A и g_{BCD}^A выпадают из-за симметрии по паре индексов. В результате получаем ограничение на *структурные константы* группы Ли,

$$f_{AB}^C := -\hat{f}_{AB}^C + \hat{f}_{BA}^C, \quad f_{AB}^C = -f_{BA}^C, \quad (8.10)$$

известные как *тождества Якоби*

$$f_{AB}^D f_{CD}^E + f_{BC}^D f_{AD}^E + f_{CA}^D f_{BD}^E = 0, \quad (8.11)$$

где слагаемые отличаются циклической перестановкой индексов A, B, C. При свертке тождеств Якоби по индексам C, E последние два слагаемых сокращаются, и мы получаем тождество

$$f_{AB}^D f_{DC}^C = 0.$$

Для абелевых групп функция композиции симметрична, $f^A(a, b) = f^A(b, a)$, и структурные константы тождественно обращаются в нуль.

Координаты обратного элемента $a^{-1A}(a)$ зависят от элемента a . Разложив это выражение по a^A и подставив в определение обратного элемента, получим следующее разложение для координат обратного элемента с точностью до членов третьего порядка:

$$a^{-1A} = -a^A + a^C a^B \hat{f}_{BC}^A - \frac{1}{2} a^D a^C a^B \hat{f}_{BC}^E (\hat{f}_{DE}^A + \hat{f}_{ED}^A) + \dots \quad (8.12)$$

При получении этого выражения было использовано соотношение (8.9).

8.2 Действие группы слева

Рассмотрим группу Ли как группу преобразований группового многообразия, действующую слева. При этом каждому элементу $a \in \mathbb{G}$ ставится в соответствие отображение

$$a : \quad \mathbb{G} \ni \quad b \mapsto l_a(b) := ab \quad \in \mathbb{G}.$$

Это отображение не является автоморфизмом группы Ли $\mathbb{G}$, т.к. нетривиально действует на единичный элемент. Из единственности единичного элемента в группе следует, что группа левых преобразований действует свободно и транзитивно (см. раздел 9).

Функция $F(a) \in C^\infty(\mathbb{G})$, заданная на групповом многообразии, будет инвариантна относительно действия группы, если выполнено условие $F(ab) = F(b)$, $\forall a \in \mathbb{G}$, и, следовательно, также для всех $b \in \mathbb{G}$. Очевидно, что инвариантные функции на группе тождественно равны константе. Другими словами, значение инвариантной функции, например, в единице $F(e)$ разносится по всему групповому многообразию с помощью действия группы слева. При этом группа Ли $\mathbb{G}$ может состоять из нескольких компонент.

Перейдем к рассмотрению левоинвариантных векторных полей на групповом многообразии. Поскольку на группе задано отображение l_a , которое является диффеоморфизмом, то в касательном пространстве определен соответствующий дифференциал отображения l_{a*} :

$$X(l_a(b)) = l_{a*}X(b), \quad (8.13)$$

который переводит вектор $X(b)$ в точке b в некоторый вектор $X(l_a(b))$ в точке ab .

Определение. Векторное поле $X \in \mathcal{X}(\mathbb{G})$ на группе Ли $\mathbb{G}$ называется *левоинвариантным*, если выполнено условие (8.13) для всех $a \in \mathbb{G}$ и, следовательно, для всех $b \in \mathbb{G}$. $\square$

Левоинвариантные векторные поля определены глобально, т.е. на всем групповом многообразии.

Предложение 8.2.1. Любое левоинвариантное векторное поле является гладким.

Доказательство. Следствие гладкости групповой операции. См., например, [42], предложение 3.7. $\square$

Проиллюстрируем действие дифференциала отображения для соответствующей локальной группы Ли, на которой задана функция композиции в явном виде. В компонентах дифференциал отображения записывается в виде

$$X^A(b) \mapsto X^A(c) = X^B(b)U_B^A(a, b), \quad c = ab,$$

где $X^A(b)$ – компоненты касательного вектора в точке b относительно координатного базиса ∂_A , и дифференциал отображения задается матрицей

$$U_B^A(a, b) := \frac{\partial f^A(a, b)}{\partial b^B},$$

зависящей от координат двух точек $a, b \in \mathbb{G}$. Поскольку групповая операция, $(a, b) \mapsto f(a, b)$, является диффеоморфизмом группового многообразия на себя при фиксированном a или b , то матрица U_A^B невырождена для всех $a, b \in \mathbb{G}$, $\det U_A^B \neq 0$ (там, где определена функция композиции).

Пусть $X_0 = \{X_0^A\}$ – произвольный вектор из касательного пространства к единице группы. Тогда этот вектор можно разнести по всему групповому многообразию с помощью дифференциала левого действия элементов группы. По построению, это будет левоинвариантное векторное поле $X(a) := l_{a*}X_0$. Оно определено глобально.

Компоненты левоинвариантного векторного поля вблизи единицы группы имеют вид

$$X^A(a) = X_0^B L_B^A(a),$$

где

$$L_B^A(a) := U_B^A(a, 0) = \left. \frac{\partial f^A(a, b)}{\partial b^B} \right|_{b=0}.$$

Тем самым, каждому вектору из касательного пространства к единице однозначно ставится в соответствие левоинвариантное векторное поле, и, наоборот, каждое левоинвариантное векторное поле однозначно определяет некоторый вектор в касательном пространстве к единице группы.

Матрица $U_B^A(a, b)$ удовлетворяет простому тождеству. Дифференцируя равенство $f^A(0, b) = b^A$ по b^B , получаем тождество

$$U_B^A(0, b) = \delta_B^A, \quad (8.14)$$

которое нам понадобится в дальнейшем.

Множество векторных полей на группе Ли $\mathbb{G}$, как и на любом другом многообразии, образует бесконечномерную алгебру Ли $\mathcal{X}(\mathbb{G})$ (см. раздел 2.6.7) и $C^\infty(\mathbb{G})$ -модуль. Поскольку дифференциал отображения сохраняет скобку Ли (коммутатор) (2.74), то коммутатор двух левоинвариантных векторных полей будет левоинвариантным векторным полем. Следовательно, множество всех левоинвариантных векторных полей образует подалгебру Ли в $\mathcal{X}(\mathbb{G})$. Поскольку, как векторное пространство, множество левоинвариантных векторных полей изоморфно касательному пространству в единице группы $\mathbb{T}_e(\mathbb{G})$, то множество левоинвариантных векторных полей представляет собой конечномерную алгебру Ли такой же размерности N , что и сама группа.

Замечание. Алгебра Ли левоинвариантных векторных полей не выдерживает умножения на гладкие функции $f(a) \in C^\infty(\mathbb{G})$, и поэтому представляет собой линейное пространство, а не $C^\infty(\mathbb{G})$ -модуль. $\square$

Структура алгебры Ли (коммутатор) левоинвариантных векторных полей естественным образом переносится на касательное пространство к единице группы:

$$[X_0, Y_0] := [l_{a*}X_0, l_{a*}Y_0]_e.$$

Словами. Берем два произвольных вектора X_0 и Y_0 из касательного пространства к единице группы, разносим их с помощью действия группы слева по всему многообразию, вычисляем коммутатор получившихся левоинвариантных векторных полей и выбираем в касательном пространстве к единице группы тот вектор, который соответствует коммутатору.

Множество левоинвариантных векторных полей образует N -мерное векторное пространство над полем вещественных чисел, которое изоморфно касательному пространству в начале координат (или в любой другой точке группового многообразия). В качестве базиса левоинвариантных векторных полей удобно выбрать N левоинвариантных векторных полей (*образующих алгебры Ли*)

$$L_A(a) := l_{a*}(\partial_A|_e) = L_A^B \partial_B|_a, \quad (8.15)$$

где мы разложили левоинвариантный базис по координатному базису $\partial_B|_a$ в точке $a \in \mathbb{G}$. В построенном базисе произвольное левоинвариантное векторное поле X имеет постоянные компоненты, равные его компонентам в начале координат относительно координатного базиса:

$$X = X_0^A L_A.$$

Поскольку левоинвариантные векторные поля образуют алгебру Ли, то коммутатор двух базисных левоинвариантных векторных полей L_A будет левоинвариантным векторным полем и, следовательно, его можно разложить по базису L_A с некоторыми постоянными коэффициентами:

$$[L_A, L_B] = f_{AB}^C L_C. \quad (8.16)$$

То, что в правой части этого равенства стоят структурные константы f_{AB}^C , которые были определены ранее через функцию композиции (8.10) не случайно.

Докажем это, рассмотрев окрестность единицы группы и соответствующую локальную группу Ли. Вблизи единицы группы левоинвариантный базис в компонентах имеет вид

$$L_A(a) = L_A^B(a) \partial_B,$$

где ∂_B – координатный базис касательных пространств, что оправдывает выбранные обозначения в определении (8.15). Из разложения (8.8) следует разложение для компонент левоинвариантного базиса:

$$L_A^B = \delta_A^B + a^C \hat{f}_{AC}^B + a^D a^C g_{ACD}^B + \dots \quad (8.17)$$

Подстановка этого разложения в формулу (8.16) доказывает, что в правой части стоят структурные константы (8.10), потому что коэффициенты разложения f_{AB}^C постоянны. При этом возникают также ограничения на коэффициенты разложения при высших степенях a^A , на которых мы останавливаться не будем.

Замечание. Для всех $a \in \mathbb{G}$ выполнено равенство $\det L_A^B(a) \neq 0$, потому что групповая операция задает диффеоморфизм группы $\mathbb{G}$ на себя. При малых a^A это следует также из (8.17). $\square$

Продолжим рассмотрение локальной группы Ли и дадим независимый вывод формулы (8.16), исходя только из свойств функции композиции.

Предложение 8.2.2. *Базис алгебры Ли (8.15) удовлетворяет коммутационным соотношениям (8.16) с некоторым набором структурных констант, удовлетворяющих тождеству Якоби.*

Доказательство. Рассмотрим тождество $b = a^{-1}ab$ в касательном пространстве, получим равенство

$$L_A^{-1B}(a) = U_A^B(a^{-1}, a).$$

Далее, распишем условие ассоциативности $(ab)c = a(bc)$ в касательном пространстве:

$$U_A^B(f(a, b), c) = U_A^C(b, c) U_C^B(a, f(b, c)).$$

Полагая $c = 0$, получим равенства

$$\frac{\partial f^C(a, b)}{\partial b^A} = U_A^C(a, b) = L_A^{-1D}(b) L_D^C(f), \quad (8.18)$$

где $f = f(a, b)$. Рассмотрим эту систему уравнений, как уравнения на функцию композиции. Система уравнений (8.18) разрешима тогда и только тогда, когда выполнены условия совместности. Чтобы получить эти условия, продифференцируем уравнение (8.18) по b^B и антисимметризуем по индексам А и В. После несложных алгебраических преобразований получим уравнение

$$[L_A(b), L_B(b)]^D L_D^{-1C}(b) = [L_A(f), L_B(f)]^D L_D^{-1C}(f),$$

где квадратные скобки обозначают коммутатор векторных полей. Поскольку левая и правая часть этого равенства рассматриваются в различных точках группового многообразия, и точки могут выбираться произвольно, то они должны быть равны константам. Эти константы проще всего определить, рассмотрев линейное приближение для левоинвариантных векторных полей (8.17). В результате получим уравнение для образующих алгебры Ли (8.16).

Тождества Якоби для структурных констант следуют из тождеств Якоби для векторных полей. $\square$

Таким образом, мы доказали, что множество левоинвариантных векторных полей образует N -мерную алгебру Ли над полем вещественных чисел, которая является подалгеброй бесконечномерной алгебры Ли всех векторных полей $\mathcal{X}(\mathbb{G})$.

Определение. Алгебра левоинвариантных векторных полей называется *алгеброй Ли* $\mathfrak{g}$ группы Ли $\mathbb{G}$. Базисные векторы (образующие) L_A алгебры Ли $\mathfrak{g}$ называются *генераторами группы Ли*. $\square$

Пример 8.2.1. Вещественная прямая $x \in \mathbb{R}$ является абелевой группой по сложению. Ее алгебра Ли состоит из постоянных векторных полей $X = a\partial_x \in \mathcal{X}(\mathbb{R})$, $a \in \mathbb{R}$. Это – единственная одномерная алгебра Ли. $\square$

Предложение 8.2.3. *Существует две неизоморфных двумерных алгебры Ли.*

Доказательство. Пусть L_x, L_y – базис двумерной алгебры Ли. Тогда существует две возможности.

1) Если $[L_x, L_y] = 0$, то коммутатор любых двух элементов алгебры Ли равен нулю, и мы имеем двумерную абелеву алгебру Ли.

2) Пусть $[L_x, L_y] = \alpha L_x + \beta L_y$ для некоторых $\alpha, \beta \in \mathbb{R}$, причем хотя бы одно из чисел α, β отлично от нуля. Тогда линейным преобразованием базиса алгебру Ли можно привести к виду

$$[L_1, L_2] = -L_2, \quad [L_1, L_1] = [L_2, L_2] = 0. \quad \square \quad (8.19)$$

Единственная неабелева двумерная алгебра Ли (8.19) соответствует группе Ли аффинных преобразований прямой, которая будет подробно рассмотрена в разделе 8.7.

Все трехмерные алгебры Ли также классифицированы (классификация Бианки [62]).

Алгебра Ли определяется с точностью до изоморфизма (выбора базиса левоинвариантных векторных полей) только окрестностью единицы группы. Поэтому, если две группы Ли локально изоморфны, то их алгебры Ли также изоморфны. В то же время сами группы Ли могут не быть изоморфными.

Пример 8.2.2. Группы $\mathrm{SU}(2)$, $\mathrm{SO}(3)$ и $\mathbb{O}(3)$ имеют изоморфные алгебры Ли, однако сами не изоморфны. При этом группа $\mathrm{SU}(2)$ связна и односвязна и является универсальной накрывающей группы $\mathrm{SO}(3)$. Группа вращений $\mathrm{SO}(3)$ связна, но не односвязна. Группа $\mathbb{O}(3)$ состоит из двух компонент связности. $\square$

Левоинвариантные векторные поля L_A (генераторы) образуют базис алгебры Ли. Этот базис определен с точностью до выбора базиса в касательном пространстве к единице группы, т.е. с точностью до действия группы $\mathrm{GL}(n, \mathbb{R})$. При этом вид структурных констант зависит от выбора базиса. Легко проверить, что при преобразовании базиса структурные константы преобразуются как компоненты тензора третьего ранга с двумя ковариантными индексами и одним контравариантным.

Выше мы показали, как групповая операция определяет алгебру Ли левоинвариантных векторных полей. Проведем обратное построение и покажем, что структурные константы (8.10) со свойством (8.11) позволяют восстановить функцию композиции $f^A(a, b)$ в окрестности единицы. Это построение дает

Доказательство теоремы 8.11.3. Перепишем соотношение (8.16) в компонентах:

$$L_A^D \partial_D L_B^C - L_B^D \partial_D L_A^C = f_{AB}^D L_D^C. \quad (8.20)$$

Свертка этого выражения с тремя обратными матрицами L^{-1} по индексам A, B, C приводит к соотношению

$$\partial_A L_B^{-1C} - \partial_B L_A^{-1C} = -L_A^{-1E} L_B^{-1D} f_{DE}^C. \quad (8.21)$$

Это равенство является дифференциальным уравнением на матрицу $L^{-1}(a)$. Левая часть уравнения антисимметрична по индексам A и B и, следовательно, представляет собой 2-форму на $\mathbb{G}_0$. Для того, чтобы найти необходимые и достаточные условия интегрируемости этой системы уравнений, ее необходимо продифференцировать, скажем, по a^D , и антисимметризовать по индексам A, B и D . После несложных вычислений, получим, что необходимым и достаточным условием интегрируемости уравнений (8.21) являются тождества Якоби для структурных констант. Таким образом, матрицы $L^{-1}(a)$ и, следовательно, $L(a)$ определяются структурными константами с точностью до некоторой константы, которая фиксируется условием $L_A^B(0) = \delta_A^B$. После этого функция композиции однозначно находится из уравнения (8.18) с фиксированными условиями в нуле. Условия интегрируемости этих уравнений, как было показано выше, опять-таки сводятся к тождествам Якоби. Таким образом, если задан произвольный набор структурных констант $f_{AB}^C = -f_{BA}^C$, удовлетворяющих тождествам Якоби (8.11), то функция композиции однозначно определяется в окрестности единицы группы и, следовательно, групповая структура задается в некоторой окрестности единицы группы $\mathbb{G}$. Таким образом, доказано, что произвольная алгебра Ли определяет локальную группу Ли с точностью до изоморфизма.

В обратную сторону: каждая локальная группа Ли определяет алгебру Ли с точностью до изоморфизма. Это сразу следует из определения алгебры Ли, как множества левоинвариантных векторных полей. $\square$

Замечание. Явный вид функции композиции известен только в самых простейших случаях, например, для абелевой группы сложения векторов евклидова пространства $f^A = a^A + b^A$. В примере 8.1.11 функция композиции была построена для группы $\mathrm{GL}(n, \mathbb{R})$. В разделе 8.7 будет явно построена функция композиции в простейшем случае двумерной неабелевой группы Ли (группа аффинных преобразований прямой). В более сложных случаях решить явно уравнения (8.21) и (8.18) не удастся.

Но это обычно и не нужно. В приложениях, как правило, работают с матричным представлением алгебры Ли и структурными константами. $\square$

Определение. Пусть $X = X_0^A L_A^B(a) \partial_B$ – произвольное левоинвариантное векторное поле. Форма $\omega = dx^A \omega_A(a)$ называется *левоинвариантной*, если ее значение на левоинвариантном векторном поле равно константе,

$$\omega(X) = X_0^A L_A^B \omega_B = \text{const.} \quad \square$$

Предложение 8.2.4. *Любая левоинвариантная 1-форма является гладкой.*

Доказательство. Следствие гладкости групповой операции. См., например, [42], предложение 3.12. $\square$

Из определения следует, что компоненты левоинвариантной 1-формы в координатном базисе имеют вид

$$\omega_A(a) = L_A^{-1B}(a) \omega_{0B}, \quad (8.22)$$

где ω_{0B} – компоненты левоинвариантной 1-формы в начале координат. При работе с левоинвариантными 1-формами удобно ввести левоинвариантный базис, который дуален к левоинвариантным векторным полям: $\omega^A(L_B) = \delta_B^A$. Он имеет следующий вид в координатном базисе:

$$\omega^A(a) = dx^B L_B^{-1A}(a),$$

Тогда произвольная левоинвариантная 1-форма будет иметь в левоинвариантном базисе постоянные компоненты $\omega := \omega^A \omega_{0A}$.

Предложение 8.2.5 (Формула Маурера–Картана). *Пусть ω – произвольная левоинвариантная 1-форма и X, Y – два произвольных левоинвариантных векторных поля. Тогда значение внешней производной $d\omega$ на векторных полях X, Y пропорционально значению 1-формы ω на коммутаторе этих полей:*

$$d\omega(X, Y) = -\frac{1}{2} \omega([X, Y]). \quad (8.23)$$

Это тождество известно, как формула Маурера–Картана. Ее можно переписать для левоинвариантного базиса 1-форм:

$$d\omega^A = -\frac{1}{2} \omega^B \wedge \omega^C f_{BC}^A. \quad (8.24)$$

Доказательство. Прямая проверка с использованием тождества (8.21). $\square$

При рассмотрении главных расслоений нам понадобится рассматривать формы, заданные на некотором многообразии $\mathbb{M}$, со значениями в алгебре Ли $\mathfrak{g}$. Введем для таких форм понятие коммутатора. Пусть L_A , $A = 1, \dots, N$, – базис алгебры Ли. Пусть A и B – соответственно, r и s формы на $\mathbb{M}$ со значениями в алгебре Ли $\mathfrak{g}$. Тогда они имеют вид

$$A = A^A L_A, \quad B = B^A L_A,$$

где $A^A \in \Lambda_r(\mathbb{M})$ и $B^A \in \Lambda_s(\mathbb{M})$ для всех значений индекса $A = 1, \dots, N$. Определим коммутатор заданных форм следующим равенством

$$[A, B] := A^A \wedge B^B [L_A, L_B] = A^A \wedge B^B f_{AB}^C L_C. \quad (8.25)$$

Выражение в правой части равенства представляет собой $r + s$ форму на $\mathbb{M}$ со значениями в алгебре Ли $\mathfrak{g}$.

Если на многообразии $\mathbb{M}$ заданы три формы A , B и C степеней r , s и t , соответственно, со значениями в алгебре Ли $\mathfrak{g}$, то из тождеств Якоби вытекает равенство

$$(-1)^{rt}[[A, B], C] + (-1)^{rs}[[B, C], A] + (-1)^{st}[[C, A], B] = 0. \quad (8.26)$$

Пример 8.2.3. Пусть A и B – 1-формы на $\mathbb{M}$ со значениями в алгебре Ли $\mathfrak{g}$ и $X, Y \in \mathcal{X}(\mathbb{M})$ – два векторных поля на $\mathbb{M}$. Тогда значение коммутатора форм на векторных полях равно

$$\begin{aligned} [A, B](X, Y) &= A^A \wedge B^A(X, Y)[L_A, L_B] = \\ &= \frac{1}{2}(A^A(X)B^B(Y) - A^A(Y)B^B(X)) = \frac{1}{2}[A(X), B(Y)] - \frac{1}{2}[A(Y), B(X)]. \end{aligned}$$

В частности,

$$[A, A](X, Y) = [A(X), A(Y)]. \quad \square$$

Предложение 8.2.6. Если A и B – соответственно, r и s формы со значениями в алгебре Ли $\mathfrak{g}$, то справедливы следующие формулы:

$$\begin{aligned} [A, B] &= (-1)^{rs+1}[B, A], \\ d[A, B] &= [dA, B] + (-1)^r[A, dB]. \end{aligned}$$

Доказательство. Прямая проверка. □

В дальнейшем особую роль будут играть формы, заданные на группе Ли $\hat{\mathbb{G}}$, со значениями в ее алгебре Ли $\mathfrak{g}$. То есть в качестве многообразия $\mathbb{M}$ мы рассматриваем саму группу $\hat{\mathbb{G}}$, которую, для отличия, мы отметили шляпкой. Другими словами, рассмотрим два экземпляра одной и той же группы Ли: $\hat{\mathbb{G}} = \mathbb{G}$. Пусть $\hat{\omega}^A$ – левоинвариантный базис кокасательного расслоения к $\hat{\mathbb{G}}$. Реализуем алгебру Ли $\mathfrak{g}$, как алгебру левоинвариантных векторных полей на $\mathbb{G}$ с базисом L_A .

Определение. Левоинвариантная 1-форма на группе Ли $\hat{\mathbb{G}}$ со значениями в алгебре Ли $\mathfrak{g}$,

$$\theta := \hat{\omega}^A L_A, \quad (8.27)$$

называется *канонической*. □

Если $X(\hat{a}) \in \mathcal{X}(\hat{\mathbb{G}})$ – векторное поле на $\hat{\mathbb{G}}$, то его можно разложить по левоинвариантному базису: $X = X^A(\hat{a})\hat{L}_A$. Тогда значение канонической формы на этом поле равно

$$\theta(X)(\hat{a}) = X^A(\hat{a})L_A.$$

То есть каждой точке $\hat{a} \in \hat{\mathbb{G}}$ ставится в соответствие левоинвариантное векторное поле на $\mathbb{G}$. Эта форма будет использована при изучении связностей на главных расслоениях в разделе 13.1. Формулу Маурера–Картана (8.24) можно переписать для канонической формы

$$d\theta = -\frac{1}{2}[\theta, \theta], \quad (8.28)$$

где справа стоит коммутатор в алгебре Ли $\mathfrak{g}$ и внешние произведения 1-форм $\hat{\omega}^A$.

Наличие левоинвариантных форм на группе Ли позволяет доказать следующее

Предложение 8.2.7. *Любая группа Ли $\mathbb{G}$ является ориентируемым многообразием.*

Доказательство. Левоинвариантные 1-формы ω^A определены на всем групповом многообразии и линейно независимы. Поэтому N -форма $\omega^1 \wedge \dots \wedge \omega^N$ нигде не обращается в нуль. Следовательно, согласно теореме 10.4.3, любая группа Ли ориентируема. $\square$

8.3 Действие группы справа

Все, сказанное относительно действия группы на себя слева, естественным образом переносится на действие группы справа. Ниже мы рассмотрим отличия, которые при этом возникают. Для простоты, рассмотрим локальную группу Ли, т.е. окрестность единицы группы.

Рассмотрим группу Ли $\mathbb{G}$, как группу преобразований, действующую справа:

$$\mathbb{G} \ni a : \quad \mathbb{G} \ni \quad b \mapsto r_a(b) := ba \quad \in \mathbb{G}.$$

Пусть

$$V_B^A(a, b) := \frac{\partial f^A(a, b)}{\partial a^B}.$$

Дифференцируя тождество $f^A(a, 0) = a^A$ по a^B , получаем равенство

$$V_B^A(a, 0) = \delta_B^A,$$

которое нам понадобится в дальнейшем.

Аналогично случаю левого действия группы, построим правоинвариантный базис векторных полей

$$R_A(a) = R_A^B(a) \partial_B, \quad (8.29)$$

где

$$R_A^B(a) := V_A^B(0, a) = \left. \frac{\partial f^B(b, a)}{\partial b^A} \right|_{b=0}. \quad (8.30)$$

В этом базисе произвольное правоинвариантное векторное поле имеет постоянные компоненты. Как и для левоинвариантных векторных полей, $\det R_A^B(a) \neq 0$ для всех $a \in \mathbb{G}$.

Как и в случае левого действия группы, можно доказать, что правоинвариантные векторные поля образуют N -мерную алгебру Ли над полем вещественных чисел. При этом коммутатор векторов правоинвариантного базиса,

$$[R_A, R_B] = -f_{AB}^C R_C, \quad (8.31)$$

отличается знаком структурных констант от коммутатора левоинвариантного базиса. Алгебры Ли лево- и правоинвариантных векторных полей изоморфны, поскольку переходят друг в друга после отображения $R_A \mapsto -R_A$.

Формула для коммутатора правоинвариантных векторных полей (8.31) справедлива, как легко видеть, глобально на всей группе Ли.

Из разложения функции композиции (8.8) следует разложение для компонент правоинвариантных векторных полей относительно координатного базиса

$$R_B^A(a) = \delta_B^A + a^C \hat{f}_{CB}^A + a^D a^C h_{CDB}^A + \dots \quad (8.32)$$

Из тождества $b = baa^{-1}$, записанного в касательном пространстве, следует равенство

$$R_{\mathbf{B}}^{-1\mathbf{A}}(a) = V_{\mathbf{B}}^{\mathbf{A}}(a, a^{-1}). \quad (8.33)$$

Рассмотрим функцию $F(a) \in \mathcal{C}^\infty(\mathbb{G})$ на группе Ли $\mathbb{G}$. Подействуем на точки многообразия бесконечно малым элементом $\epsilon \in \mathbb{G}$ справа. Тогда точка a перейдет в точку $a' = a + da$, где

$$da^{\mathbf{A}} := f^{\mathbf{A}}(a, \epsilon) - a^{\mathbf{A}} = \epsilon^{\mathbf{B}} L_{\mathbf{B}}^{\mathbf{A}}(a).$$

Поскольку для функции $F'(a') = F(a)$, то изменение формы функции равно:

$$\delta F(a) := F'(a) - F(a) = -\epsilon^{\mathbf{A}} L_{\mathbf{A}}^{\mathbf{B}} \partial_{\mathbf{B}} F(a) = -\epsilon^{\mathbf{A}} L_{\mathbf{A}} F(a).$$

Таким образом, левоинвариантные векторные поля, с точностью до знака, являются генераторами действия группы Ли справа. Аналогично доказывается, что генераторами групповых преобразований слева являются правоинвариантные векторные поля.

Поскольку преобразования, вызванные действием группы слева и справа коммутируют, то отсюда следует, что лево- и правоинвариантные векторные поля коммутируют между собой,

$$[L_{\mathbf{A}}, R_{\mathbf{B}}] = 0. \quad (8.34)$$

Это равенство справедливо, конечно, глобально.

Из формулы для производной Ли от векторного поля (2.123) и равенства (8.34) следует, что производная Ли от правоинвариантного векторного поля вдоль левоинвариантного векторного поля равна нулю и наоборот.

В дальнейшем под алгеброй Ли $\mathfrak{g}$ группы Ли $\mathbb{G}$ всегда понимается множество левоинвариантных векторных полей, которые генерируют действие группы справа.

8.4 Присоединенное представление

Рассмотрим автоморфизм группы Ли,

$$\mathrm{ad}(a) : \quad \mathbb{G} \ni \quad b \mapsto \mathrm{ad}(a)b := aba^{-1} \quad \in \mathbb{G}, \quad (8.35)$$

который сопоставляется каждому элементу группы $a \in \mathbb{G}$. Каждый автоморфизм группы Ли порождает автоморфизм ее алгебры Ли, потому что левоинвариантные векторные поля получены с помощью группового действия. Поэтому автоморфизм $\mathrm{ad} a$ порождает автоморфизм алгебры Ли $\mathfrak{g}$, который также обозначается $\mathrm{ad} a \in \mathrm{GL}(N, \mathbb{R})$.

Определение. Представление $a \mapsto \mathrm{ad} a$, где

$$\mathrm{ad} a : \quad \mathfrak{g} \ni \quad X = X_0^{\mathbf{A}} L_{\mathbf{A}} \mapsto X' = X_0^{\mathbf{A}} S_{\mathbf{A}}^{\mathbf{B}}(a) L_{\mathbf{B}} \quad \in \mathfrak{g}, \quad (8.36)$$

индуцированное групповым действием (8.35), называется *присоединенным представлением* группы Ли $\mathbb{G}$ в ее алгебре Ли $\mathfrak{g}$. $\square$

Присоединенное представление каждому элементу группы $a \in \mathbb{G}$ ставит в соответствие $N \times N$ матрицу $S_{\mathbf{A}}^{\mathbf{B}}(a)$. Для краткости мы будем писать

$$\mathrm{ad}(a) = S_{\mathbf{A}}^{\mathbf{B}}(a).$$

Ясно, что единичному элементу группы соответствует единичная матрица

$$\text{ad}(e) = S_A^B(e) = \delta_A^B,$$

а обратному элементу группы – обратная матрица

$$\text{ad}(a^{-1}) = S_A^B(a^{-1}) = S_A^{-1B}(a).$$

В наших обозначениях матрица для произведения двух элементов группы равна произведению матриц для каждого элемента,

$$S_A^B(ab) = S_A^C(b)S_C^B(a),$$

взятых в обратном порядке.

Для абелевых групп присоединенное представление всегда тривиально.

Ядром отображения (8.35) является центр группы, т.е. множество всех тех элементов группы, которые коммутируют со всеми элементами группы $\mathbb{G}$. Каждому элементу центра группы соответствует единичная матрица присоединенного представления.

Для локальных неабелевых групп Ли можно построить явный вид матриц присоединенного представления в терминах функции композиции. С этой целью рассмотрим действие преобразования (8.35) в касательном пространстве в последовательности $b \mapsto (ab)a^{-1}$:

$$X^A(b) \mapsto X^C(b)U_C^B(a, b)V_B^A(f(a, b), a^{-1}).$$

Полагая $b = 0$ и учитывая равенство (8.33), получаем отображение

$$X_0^A \mapsto X_0^B L_B^C(a) R_C^{-1A}(a) := X_0^B S_B^A(a),$$

где введена матрица присоединенного представления $S_B^A(a) \in \text{GL}(n, \mathbb{R})$, сопоставляемая каждому элементу $a \in \mathbb{G}$. Таким образом, получаем явное выражение для матрицы присоединенного представления через производные от функции композиции:

$$S_B^A(a) = L_B^C(a) R_C^{-1A}(a) = R_B^C(a^{-1}) L_C^{-1A}(a^{-1}), \quad (8.37)$$

где последнее выражение получено в результате рассмотрения действия $\text{ad } a$ в последовательности $a(ba^{-1})$. Из этого представления следует, что

$$S_B^A(a^{-1}) = S_B^{-1A}(a).$$

Разложения (8.17) и (8.32) приводят к следующему разложению матрицы присоединенного представления вблизи начала координат

$$S_B^A(a) = \delta_B^A + a^C f_{CB}^A + \dots \quad (8.38)$$

Отсюда следует, что структурные константы, взятые с обратным знаком, можно рассматривать, как генераторы присоединенного представления: $L_{(A)B}^C = -f_{AB}^C$. При этом первый индекс A нумерует генераторы и, для отличия, взят в скобки. Вторым и третьим индексами структурных констант B, C рассматриваются, как матричные. Тождества Якоби (8.11) переписываются в виде

$$[L_{(A)}, L_{(B)}]_C^D = f_{AB}^E L_{(E)C}^D, \quad (8.39)$$

где квадратные скобки обозначают коммутатор матриц.

Рассмотрим два последовательных автоморфизма группы Ли $\mathbb{G}$, соответствующих элементам b и a :

$$c \mapsto bcb^{-1} \mapsto abcb^{-1}a^{-1} = (ab)c(ab)^{-1}.$$

Соответствующее преобразование в алгебре Ли $\mathfrak{g}$ задается матрицей присоединенного представления для произведения $f(a, b)$

$$S_A^B(f(a, b)) = S_A^C(b)S_C^B(a), \quad (8.40)$$

которая равна произведению матриц для каждого из элементов, но взятых в обратном порядке.

Пусть $b \ll 1$. Разложив равенство (8.40) в ряд по b с использованием формулы (8.38) и разложения для функции композиции,

$$f^A(a, b) = a^A + b^B L_B^A(a) + \dots,$$

в первом порядке по b получим правило дифференцирования матрицы присоединенного представления

$$L_C S_A^B = L_C^D \partial_D S_A^B = f_{CA}^D S_D^B. \quad (8.41)$$

Из этого равенства с учетом тождества $L_C(S_A^D S_D^{-1B}) = 0$ получаем правило действия левоинвариантных векторных полей на обратные матрицы присоединенного представления:

$$L_C S_A^{-1B} = -S_A^{-1D} f_{CD}^B. \quad (8.42)$$

Аналогично, раскладывая соотношение (8.40) в ряд по $a \ll 1$, получим равенства

$$R_C S_A^B = R_C^D \partial_D S_A^B = S_A^D f_{CD}^B. \quad (8.43)$$

С учетом формулы $R_C(S_A^D S_D^{-1B}) = 0$ это дает правило дифференцирования обратной матрицы

$$R_C S_A^{-1B} = -f_{CA}^D S_D^{-1B}.$$

Замечание. Полученные формулы дифференцирования матриц присоединенного представления (8.41) и (8.43) нековариантны, поскольку частная производная от тензора не является тензором. Однако при рассмотрении алгебр Ли левоинвариантные векторные поля имеют постоянные компоненты относительно левоинвариантного базиса и поэтому достаточно ограничиться аффинными преобразованиями координат с постоянными коэффициентами. В этом случае частные производные преобразуются по тензорным правилам. $\square$

Предложение 8.4.1. Структурные константы группы Ли $\mathbb{G}$ инвариантны относительно присоединенного действия группы:

$$f_{AB}^C = S_A^D S_B^E f_{DE}^F S_F^{-1C}. \quad (8.44)$$

Доказательство. Подействуем присоединенным представлением (8.35) на групповое многообразие $\mathbb{G}$. Тогда базис левоинвариантных векторных полей преобразуется в соответствии с присоединенным представлением

$$L_A \mapsto S_A^B L_B.$$

При этом коммутационные соотношения (8.16) сохраняют свой вид, поскольку действие дифференциала отображения сохраняет коммутатор векторных полей. Следовательно, для структурных констант справедливо соотношение

$$S_A^D S_B^E f_{DE}^C = f_{AB}^D S_D^C.$$

Умножив это равенство справа на S^{-1} получим (8.44). $\square$

8.5 Группы Ли как (псевдо)римановы пространства

На групповом многообразии $\mathbb{G}$ можно определить метрику, превратив тем самым группу Ли в (псевдо)риманово пространство. Среди всех возможных метрик важную роль играют метрики, инвариантные относительно групповых преобразований. Эти метрики описываются следующим образом. Зададим в начале координат произвольную симметричную невырожденную матрицу θ_{AB} и разнесем ее по групповому многообразию с помощью действия группы слева или справа. В результате получим, соответственно, лево- и правоинвариантные метрики. Компоненты этих метрик в лево- и правоинвариантном базисе имеют тот же вид θ_{AB} , что и в начале координат. В координатном базисе $g = dx^A \otimes dx^B g_{AB}$ и компоненты метрик нетривиально зависят от точки $a \in \mathbb{G}$:

$$\begin{aligned} g_{AB}^L(a) &= L_A^{-1C}(a) L_B^{-1D}(a) \theta_{CD}, \\ g_{AB}^R(a) &= R_A^{-1C}(a) R_B^{-1D}(a) \theta_{CD}, \end{aligned} \quad (8.45)$$

где индексами L и R отмечены, соответственно, лево- и правоинвариантная метрики. В общем случае лево- и правоинвариантные метрики различны и их компоненты в координатном базисе нетривиально зависят от точки группового многообразия.

Для абелевых групп лево- и правоинвариантные метрики равны, т.к. совпадают действия группы слева и справа.

Скалярное произведение левоинвариантных векторных полей: $X = X_0^A L_A$ и $Y = Y_0^A L_A$, определенное левоинвариантной метрикой, равно

$$g^L(X, Y) = X_0^A Y_0^B g^L(L_A, L_B) = X_0^A Y_0^B \theta_{AB}$$

и не зависит от точки группы Ли $\mathbb{G}$. Аналогично, скалярное произведение правоинвариантных векторных полей, определенное правоинвариантной метрикой равно константе.

Условия левой и правой инвариантности метрик в инвариантном виде имеют следующий вид. Пусть $X, Y \in \mathcal{X}(\mathbb{G})$ – два произвольных векторных поля на группе Ли $\mathbb{G}$. Тогда лево- и правоинвариантные метрики определены соотношениями

$$\begin{aligned} l_a^* g^L(X, Y) &= g^L(l_{a*} X, l_{a*} Y), \\ r_a^* g^R(X, Y) &= g^R(r_{a*} X, r_{a*} Y), \end{aligned}$$

где l_a^* и r_a^* – возвраты левого и правого действия группы.

Особую роль играют двусторонне инвариантные метрики $g^L = g^R$ на неабелевых группах Ли. Для таких метрик скалярное произведение левоинвариантных векторных полей инвариантно относительно действия группы справа. Двусторонне инвариантные метрики существуют далеко не на всех группах Ли. Условие двусторонней инвариантности немедленно следует, например, из (8.45),

$$\theta_{AB} = S_A^C(a) S_B^D(a) \theta_{CD}, \quad \forall a \in \mathbb{G},$$

где S_A^B – матрица присоединенного представления. Для малых a , т.е. вблизи единицы группы это уравнение сводится к условию

$$f_{AB}^D \theta_{DC} + f_{AC}^D \theta_{DB} = 0. \quad (8.46)$$

То есть, если двусторонне инвариантная метрика на группе Ли существует, то структурные константы со всеми опущенными индексами должны быть антисимметричны по всем трем индексам.

Замечание. Двусторонне инвариантные метрики на группе Ли играют ту же роль, что и метрика Лоренца в пространстве Минковского $\mathbb{R}^{1,n-1}$ или евклидова метрика в евклидовом пространстве $\mathbb{R}^n$. $\square$

Построим двусторонне инвариантную метрику на неабелевой группе Ли в том случае, когда это возможно.

Определение. Свертка структурных констант алгебры Ли

$$\eta_{AB} := -f_{AC}{}^D f_{BD}{}^C = -\text{tr}(f_A f_B) \quad (8.47)$$

называется *формой Киллинга–Картана*. $\square$

Здесь мы рассматриваем структурные константы $f_{AC}{}^D$ в виде набора матриц. При этом второй и третий индексы рассматриваются, как матричные, а первый – нумерует матрицы. Форма (8.47) двусторонне инвариантна, что следует из инвариантности структурных констант (8.44).

Предложение 8.5.1. *Для произвольной группы Ли $\mathbb{G}$ структурные константы со всеми опущенными индексами,*

$$f_{ABC} := f_{AB}{}^D \eta_{DC}$$

антисимметричны по всем трем индексам,

$$f_{[abc]} = 0.$$

Доказательство. Прямое следствие тождеств Якоби (8.11), которое сводится к простой проверке. $\square$

Таким образом, если форма Киллинга–Картана является невырожденной, то она задает двусторонне инвариантную метрику на группе Ли. Форма Киллинга–Картана является невырожденной не для всякой группы Ли. Напомним, что связная группа Ли называется полупростой, если она не имеет нетривиальных инвариантных связных абелевых подгрупп.

Теорема 8.5.1 (Э. Картан). *Форма Киллинга–Картана для связной группы Ли является невырожденной тогда и только тогда, когда группа Ли является полупростой.*

Доказательство. Группа Ли полупроста если и только если она имеет полупростую алгебру Ли. Доказательство для полупростых алгебр Ли приведено, например, в [32], глава 1, §2, теорема 7. $\square$

Положительная определенность формы Киллинга–Картана связана с компактностью группового многообразия. А именно, справедлива

Теорема 8.5.2. *Связная полупростая группа Ли компактна тогда и только тогда, когда ее форма Киллинга–Картана положительно определена.*

Доказательство. Доказательство для полупростых алгебр Ли приведено, например, в [63], теорема 4.3.2. $\square$

Отсюда следует, что форма Киллинга–Картана превращает компактную полупростую группу Ли в риманово многообразие с инвариантной метрикой. Для некомпактных полупростых групп Ли форма Киллинга–Картана не является знакоопределенной. Следовательно, в этом случае групповое пространство становится псевдоримановым многообразием.

Полупростые группы не исчерпывают весь класс групп Ли, на которых можно задать двусторонне инвариантную метрику. Например, на абелевых группах любая невырожденная матрица в единице, разнесенная по групповому многообразию, задает двусторонне инвариантную метрику.

Пример 8.5.1. Двумерная связная абелева группа Ли изоморфна либо евклидовой плоскости $\mathbb{R}^2$ со сложением в качестве групповой операции, либо цилиндру, либо тору. Пусть x, y – декартовы координаты на плоскости. Тогда лево- и правоинвариантные векторные поля совпадают и равны

$$L_1 = R_1 = \partial_x, \quad L_2 = R_2 = \partial_y.$$

Двусторонне инвариантная метрика на $\mathbb{R}^2$ получается путем разнесения произвольной симметричной невырожденной 2×2 матрицы θ_{AB} , $A, B = 1, 2$, заданной в начале координат, с помощью действия группы. В результате получаем двусторонне инвариантную метрику, которая в декартовых координатах имеет вид θ_{AB} для всех точек $\mathbb{R}^2$. Евклидова плоскость с такой метрикой является в общем случае однородным, но не изотропным пространством. При $\theta_{AB} = \delta_{AB}$ евклидова плоскость становится также изотропной, т.е. инвариантной также относительно действия группы вращений $\mathbb{O}(2)$. Цилиндр и тор в этом случае не инвариантны относительно вращений. $\square$

Форма Киллинга–Картана (8.47) задает метрику на групповом многообразии полупростой группы Ли $\mathbb{G}$ либо в лево-, либо в правоинвариантном базисе. Для определенности будем считать, что метрика задана в левоинвариантном базисе. Тогда в координатном базисе она будет нетривиально зависеть от точки группового многообразия:

$$g_{AB}(a) = L_A^{-1C}(a) L_B^{-1D}(a) \eta_{CD}.$$

Двусторонняя инвариантность формы Киллинга–Картана приводит к существованию, по крайней мере, $2N$ векторов Киллинга L_A и R_A . Уравнения Киллинга в неголономном базисе (6.114) выполнены ввиду антисимметрии символов Кристоффеля (8.48) по первым двум индексам.

Вычислим геометрические характеристики группового многообразия неабелевой полупростой группы Ли, рассматривая его, как (псевдо)риманово пространство с нулевым кручением и неметричностью. Вычисления можно проводить в любом базисе, но в данном случае удобнее рассматривать все характеристики многообразия в неголономном базисе, определяемом левоинвариантными векторными полями L_A , поскольку компоненты метрики в этом базисе постоянны. Соответствующие формулы для вычислений приведены в разделе 6.9.

Компоненты неголономности базиса определяются коммутатором левоинвариантных векторных полей (8.16) и совпадают со структурными константами:

$$c_{AB}^C = f_{AB}^C.$$

Символы Кристоффеля в неголономном базисе, вычисленные по формуле (6.110), равны

$$\tilde{\omega}_{ABC} = \frac{1}{2} f_{ABC}, \quad (8.48)$$

где знак тильды означает, что кручение и неметричность положены равными нулю. В рассматриваемом случае символы Кристоффеля антисимметричны по первым двум индексам ввиду неголономности используемого базиса. Соответствующий тензор кривизны также задается структурными константами

$$\tilde{R}_{ABC}{}^D = \frac{1}{4} f_{AB}{}^E f_{CE}{}^D.$$

Отсюда вытекают выражения для тензора Риччи и скалярной кривизны:

$$\tilde{R}_{AB} = -\frac{1}{4} \eta_{AB}, \quad \tilde{R} = -\frac{N}{4},$$

где η_{AB} – форма Киллинга–Картана и N – размерность полупростой группы Ли.

Покажем, что полупростая группа Ли с формой Киллинга–Картана в качестве метрики, представляет собой пространство постоянной кривизны. В этом случае структурные константы f_{ABC} задают полностью антисимметричный ковариантный тензор третьего ранга в левоинвариантном базисе. Тогда ковариантная производная этого тензора со связностью (8.48),

$$\tilde{\nabla}_A f_{BCD} = L_A f_{BCD} - \frac{1}{2} f_{AB}{}^E f_{ECD} - \frac{1}{2} f_{AC}{}^E f_{BED} - \frac{1}{2} f_{AD}{}^E f_{BCE} = 0,$$

равна нулю в силу тождеств Бианки. Поскольку связность является метрической, а подъем и опускание индексов можно переставлять с оператором ковариантного дифференцирования, то ковариантная производная от тензора кривизны равна нулю,

$$\tilde{\nabla}_A \tilde{R}_{BCD}{}^E = 0.$$

Это и означает, что группа Ли как (псевдо)риманово пространство является пространством постоянной кривизны.

Тензор кривизны со всеми опущенными индексами в неголономном базисе имеет вид

$$\tilde{R}_{ABCD} = -\frac{1}{4} f_{AB}{}^E f_{CDE}, \quad (8.49)$$

в котором явно прослеживаются все свойства симметрии относительно перестановок индексов.

Замечание. Широкий класс (псевдо)римановых пространств $(\mathbb{M}, g)$ постоянной кривизны описывается метрикой, удовлетворяющей следующему соотношению

$$\tilde{R}_{\alpha\beta\gamma\delta} = C(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}), \quad (8.50)$$

где C – некоторая постоянная. Группы Ли с формой Киллинга–Картана в качестве метрики дают пример пространств постоянной кривизны другого типа, для которых равенство (8.50) не выполнено. Действительно, равенство

$$f_{AB}{}^E f_{CDE} = C(\eta_{AC}\eta_{BD} - \eta_{AD}\eta_{BC}), \quad C = \text{const},$$

не может быть выполнено ни при каких значениях константы C . Для доказательства достаточно свернуть это равенство сначала с η^{BD} , а затем с $f^{CD}{}_F$, что приведет к противоречию. $\square$

8.6 Группы Ли как пространства Римана–Картана

В предыдущем разделе связность на групповом многообразии определялась по метрике при нулевом кручении и неметричности. Ниже мы рассмотрим другой способ определения связности, который во многих отношениях является более естественным для групп Ли. А именно, отождествим групповое действие справа с параллельным переносом. Это означает, что результат параллельного переноса вектора из точки a в точку c совпадает с дифференциалом группового преобразования $c = f(a, b)$ под действием элемента $b := a^{-1}c$.

Формализуем данное определение связности. В рассматриваемом случае правоинвариантные векторные поля можно рассматривать, как результат параллельного переноса векторов из начала координат по всему групповому многообразию. Это означает, что в правоинвариантном базисе компоненты векторов постоянны, и, следовательно, компоненты связности равны нулю. Отсюда сразу вытекает равенство нулю тензора кривизны. То есть группа Ли с аффинной связностью, заданной групповой операцией, является пространством абсолютного параллелизма. Так и должно быть, потому что результат параллельного переноса вектора из точки a в точку $c = ab$ однозначно определяется начальной и конечной точкой и не зависит от пути, вдоль которого осуществляется параллельный перенос.

Кручение для рассматриваемой связности отлично от нуля. В правоинвариантном базисе, который удовлетворяет коммутационным соотношениям (8.31), компоненты тензора кручения, согласно формулам (6.109) имеют вид

$$T_{AB}{}^C = f_{AB}{}^C.$$

Замечание. В данном определении связности наличие или отсутствие метрики на групповом многообразии не играет никакой роли. $\square$

Посмотрим, как выглядят основные геометрические объекты в левоинвариантном базисе. В левоинвариантном базисе правоинвариантное векторное поле имеет непостоянные компоненты:

$$X = X_0^A R_A{}^B(a) \partial_B = X_0^A S_A{}^B(a^{-1}) L_B,$$

где мы воспользовались представлением (8.37) для матрицы присоединенного представления. Несложные вычисления показывают, что разница компонент векторов в соседних точках равна

$$\delta X^A(a) := X^A(a + da) - X^A(a) = -X^C da^B f_{BC}{}^A.$$

С другой стороны, при параллельном переносе компоненты вектора получают приращение, определяемое аффинной связностью:

$$\delta X^A = -X^C da^B \omega_{BC}{}^A.$$

Отсюда следует явное выражение для компонент аффинной связности в левоинвариантном базисе

$$\omega_{AB}{}^C = f_{AB}{}^C. \quad (8.51)$$

Связность (8.51) отличается от связности Леви-Чивиты (8.48) множителем $1/2$. Соответствующие тензоры кривизны и кручения, вычисленные по формулам для неголономного базиса (6.108) и (6.109), равны

$$R_{ABC}{}^D = 0, \quad (8.52)$$

$$T_{AB}{}^C = f_{AB}{}^C. \quad (8.53)$$

Равенство компонент тензора кривизны нулю соответствует тому, что групповое пространство является пространством абсолютного параллелизма. Компоненты тензора кручения в левоинвариантном базисе имеют тот же вид, что и в правоинвариантном. Это соответствует тому обстоятельству, что структурные константы инвариантны относительно действия присоединенного представления (8.44).

Компоненты связности (8.51) в левоинвариантном базисе определены независимо от метрики и для произвольных, не обязательно полупростых, групп. На полупростых группах Ли эта связность является метрической:

$$\nabla_A \eta_{BC} = -\omega_{AB}{}^D \eta_{DC} - \omega_{AC}{}^D \eta_{BD},$$

где η_{BC} – двусторонне инвариантная форма Киллинга–Картана. Это значит, что групповое многообразие полупростых групп Ли со связностью, определенной компонентами (8.51), представляет собой пространство Римана–Картана.

Группа Ли, рассматриваемая как пространство Римана–Картана является пространством постоянной (нулевой) кривизны и постоянного кручения. Действительно, ковариантная производная от тензора кривизны, очевидно, равна нулю. Ковариантная производная от тензора кручения со связностью (8.51),

$$\nabla_A T_{BC}{}^D = L_A T_{BC}{}^D - f_{AB}{}^E T_{EC}{}^D - f_{AC}{}^E T_{BE}{}^D + f_{AE}{}^D T_{BC}{}^E = 0,$$

также равна нулю в силу тождеств Якоби.

Данное определение связности не зависит от задания метрики на групповом многообразии, т.к. связность была определена исключительно групповой операцией. По построению, интегральные кривые правоинвариантных векторных полей на группе Ли являются геодезическими линиями для этой связности. Действительно, правоинвариантное векторное поле является касательным к интегральной кривой и в то же время является результатом параллельного переноса со связностью (8.51).

Замечание. Для задания на группе Ли инвариантной аффинной связности общего вида можно задать в некоторой точке метрику, кручение и тензор неметричности, а затем все эти объекты разнести по групповому многообразию. $\square$

8.7 Группа аффинных преобразований прямой

Проиллюстрируем общие свойства групп Ли на примере простейшей связной абелевой группы Ли $\mathbb{G}$. Размерность этой группы минимальна и равна двум. Этот пример интересен, поскольку нетривиален и в то же время достаточно прост для явного построения всех геометрических конструкций. Это – один из немногих случаев, когда функцию композиции можно записать в явном виде на всем групповом многообразии.

Простейшая двумерная алгебра Ли $\mathfrak{g}$ с образующими $\{L_A\} = (L_x, L_y)$, $A = x, y$, задается коммутационными соотношениями:

$$\begin{aligned} [L_x, L_x] &= [L_y, L_y] = 0, \\ [L_x, L_y] &= \alpha L_y, \quad \alpha \in \mathbb{R}. \end{aligned} \tag{8.54}$$

При $\alpha = 0$ эта группа является абелевой. При $\alpha \neq 0$ растяжкой координат, которая соответствует преобразованию $L_x \rightarrow \alpha L_x$, можно обратить структурную константу в единицу, $\alpha = 1$. Мы этого делать не будем, чтобы следить за пределом $\alpha \rightarrow 0$. Группа, соответствующая этой алгебре является подгруппой группы Лоренца $\mathbb{SO}(1, 2)$.

Алгебра Ли (8.54) содержит инвариантную абелеву подгруппу, которая порождается образующей L_y . Поэтому она не является полупростой.

Предложение 8.7.1. *Алгебра Ли (8.54) является единственной двумерной неабелевой алгеброй Ли с точностью до изоморфизма.*

Доказательство. В общем случае нетривиальное коммутационное соотношение для неабелевой алгебры Ли определяется двумя отличными от нуля постоянными α и β :

$$[L_x, L_y] = \beta L_x + \alpha L_y,$$

где хотя бы одна из постоянных отлична от нуля. Пусть $\beta \neq 0$. Тогда сдвигом

$$L_y \mapsto L_x + \frac{\alpha}{\beta} L_y$$

это коммутационное соотношение приводится к виду (8.54). $\square$

Покажем, что алгебра Ли (8.54) является алгеброй Ли группы аффинных преобразований прямой. Аффинные преобразования прямой $x \in \mathbb{R}$ параметризуются двумя числами $a, b \in \mathbb{R}$:

$$\mathbb{R} \ni x \mapsto ax + b \in \mathbb{R}, \quad a \neq 0.$$

Параметр a параметризует дилатации, а b – сдвиги. Генераторы дилатаций и сдвигов имеют, соответственно, следующий вид

$$L_1 := x\partial_x, \quad L_2 := \partial_x.$$

Нетрудно вычислить соответствующую алгебру Ли:

$$[L_1, L_2] = -L_2, \quad [L_1, L_1] = [L_2, L_2] = 0, \quad (8.55)$$

которая совпадает с алгеброй Ли (8.54) при $\alpha = -1$. Таким образом, алгебра Ли (8.54) изоморфна алгебре Ли аффинных преобразований прямой.

Группа аффинных преобразований прямой двумерна и неабелева. Она состоит из двух компонент связности: $a > 0$ и $a < 0$. Каждая компонента связности диффеоморфна $\mathbb{R}^2$.

Опишем явно связную компоненту единицы группы аффинных преобразований прямой. Зададим алгебру Ли (8.54) двумя левоинвариантными векторными полями (генераторами действия группы справа) на евклидовой плоскости $\mathbb{R}^2$ в декартовых координатах x, y :

$$L_x := \partial_x, \quad L_y := e^{\alpha x} \partial_y. \quad (8.56)$$

Нетрудно проверить, что эти векторные поля действительно удовлетворяют алгебре (8.54). При $\alpha = 0$ группа Ли является абелевой и совпадает с группой трансляций плоскости. Поэтому ограничимся неабелевым случаем $\alpha > 0$. Структурные константы группы Ли равны

$$\begin{aligned} f_{xy}^x &= -f_{yx}^x = 0, \\ f_{xy}^y &= -f_{yx}^y = \alpha. \end{aligned}$$

Простые вычисления дают форму Киллинга–Картана (8.47)

$$\eta_{AB} = \begin{pmatrix} \alpha^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Эта форма вырождена, поскольку группа Ли $\mathbb{G}$ не является полупростой.

Постоим функцию композиции $f^\Lambda(a, b)$ в явном виде. Обозначим координаты элементов группы $a, b \in \mathbb{G}$ через $a = (x_a, y_a)$ и $b = (x_b, y_b)$. Тогда система дифференциальных уравнений на функцию композиции (8.18) примет вид

$$\begin{aligned} \frac{\partial f^x(a, b)}{\partial x_b} &= 1, & \frac{\partial f^y(a, b)}{\partial x_b} &= 0, \\ \frac{\partial f^x(a, b)}{\partial y_b} &= 0, & \frac{\partial f^y(a, b)}{\partial y_b} &= e^{\alpha(f^x - x_b)}. \end{aligned}$$

Два уравнения на f^x с учетом начального условия $f^x(a, 0) = x_a$ имеют единственное решение $f^x(a, b) = x_a + x_b$. Аналогично, оставшиеся уравнения определяют вторую компоненту функции композиции $f^y(a, b) = y_a + e^{\alpha x_a} y_b$. Таким образом, дифференциальные уравнения (8.18) с начальными условиями $a^\Lambda = f^\Lambda(a, 0)$ однозначно определили функцию композиции по левоинвариантным векторным полям L_Λ :

$$\{f^\Lambda(a, b)\} = (x_a + x_b, y_a + e^{\alpha x_a} y_b). \quad (8.57)$$

Поскольку функция композиции определена для всех $a, b \in \mathbb{R}^2$ и отображает $\mathbb{R}^2 \times \mathbb{R}^2$ на всю плоскость $\mathbb{R}^2$, то группа Ли $\mathbb{G}$, как многообразие, диффеоморфна евклидовой плоскости $\mathbb{R}^2$ и, следовательно, некомпактна.

Несложные вычисления показывают, что построенная функция композиции (8.57) действительно удовлетворяет всем групповым аксиомам (8.7) с обратным элементом:

$$a^{-1} = (-x_a, -e^{-\alpha x_a} y_a).$$

Таким образом, мы построили неабелеву групповую структуру на евклидовой плоскости $\mathbb{R}^2$. Вспомним также, что плоскость $\mathbb{R}^2$ можно рассматривать, как абелеву группу Ли двумерных трансляций. Тем самым мы показали, что на одном и том же многообразии возможно задание различных групповых структур.

Знание функции композиции позволяет вычислить все объекты на групповом многообразии, которые задаются групповым действием. В частности, правоинвариантные векторные поля (8.29) имеют вид:

$$R_x = \partial_x + \alpha y \partial_y, \quad R_y = \partial_y, \quad (8.58)$$

и удовлетворяют коммутационным соотношениям:

$$\begin{aligned} [R_x, R_x] &= [R_y, R_y] = 0 \\ [R_x, R_y] &= -\alpha R_y, \end{aligned}$$

которые отличаются знаком от коммутационных соотношений алгебры Ли левоинвариантных векторных полей (8.54). Нетрудно также проверить, что лево- и правоинвариантные векторные поля коммутируют, $[L_\Lambda, R_\Lambda] = 0$.

Построим интегральные кривые левоинвариантных векторных полей. Произвольное левоинвариантное векторное поле задается двумя постоянными $X_0, Y_0 \in \mathbb{R}$

$$L = X_0 \partial_x + Y_0 e^{\alpha x} \partial_y.$$

Соответствующие интегральные кривые $(x(t), y(t))$ определяются системой уравнений

$$\dot{x} = X_0, \quad \dot{y} = Y_0 e^{\alpha x}.$$

При $X_0 \neq 0$, эти уравнения имеют общее решение, задающее интегральную кривую в параметрическом виде

$$\begin{aligned} x &= X_0 t + x_0, \\ y &= \frac{Y_0}{\alpha X_0} e^{\alpha x_0} (e^{\alpha X_0 t} - 1) + y_0, \end{aligned}$$

где постоянные интегрирования x_0, y_0 являются координатами точки, через которую проходит интегральная кривая при $t = 0$. Отсюда следует соотношение, определяющее форму интегральной кривой

$$y = \frac{Y_0}{\alpha X_0} (e^{\alpha x} - e^{\alpha x_0}) + y_0.$$

При $X_0 = 0$ интегральные кривые параллельны оси y . Вид интегральных кривых, проходящих через начало координат под разными углами, которые определяются постоянными X_0, Y_0 , показан на рис. 8.1 Интегральные кривые определены для всех

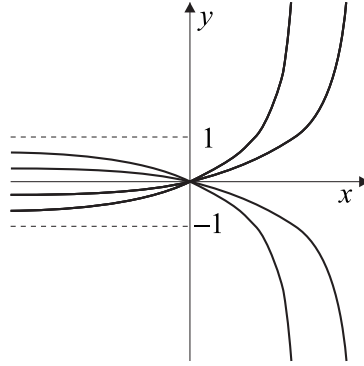


Рис. 8.1: Интегральные кривые для левоинвариантных векторных полей, проходящие через начало координат, $x_0 = 0, y_0 = 0$, под углом $\arctg(Y_0/X_0)$.

значений параметра $t \in \mathbb{R}$ и поэтому левоинвариантные векторные поля полны.

Зададим теперь геометрию, т.е. метрику и аффинную связность, на группе Ли аффинных преобразований прямой $\mathbb{G} \approx \mathbb{R}^2$. Конечно, интерес представляют те геометрические объекты, которые согласованы с групповой операцией. Начнем с метрики. Поскольку группа Ли $\mathbb{G}$ неабелева и не является полупростой, то на ней не существует двусторонне инвариантной метрики. Лучшее, что можно сделать, это построить право- или левоинвариантную метрику. С этой целью зададим матрицу θ_{AB} в начале координат и разнесем ее, например, с помощью действия группы справа. Правоинвариантная метрика в координатном базисе (8.45) имеет вид

$$g_{AB} := g_{AB}^R(a) = R_A^{-1C}(a) R_B^{-1D}(a) \theta_{CD}, \quad (8.59)$$

где матрица θ_{AB} симметрична и невырождена, а в остальном совершенно произвольна. В частности, она может иметь лоренцеву сигнатуру, т.е. на группе Ли $\mathbb{G}$ можно построить как риманову, так и псевдориманову геометрию.

Рассмотрим случай, когда в начале координат задана единичная матрица, $\theta_{AB} = \delta_{AB}$. Тогда формула (8.59), в которой матрицы R_A^B определены равенствами (8.58), задает следующую правоинвариантную метрику в декартовых координатах на плоскости $\mathbb{R}^2$

$$g_{AB} = \begin{pmatrix} 1 + \alpha^2 y^2 & -\alpha y \\ -\alpha y & 1 \end{pmatrix}. \quad (8.60)$$

Эта метрика имеет единичный определитель, $\det g_{AB} = 1$, и ее обратная имеет вид

$$g^{AB} = \begin{pmatrix} 1 & \alpha y \\ \alpha y & 1 + \alpha^2 y^2 \end{pmatrix}.$$

Прямые вычисления дают следующие выражения для символов Кристоффеля:

$$\begin{aligned} \tilde{\Gamma}_{xx}^x &= -\alpha^3 y^2, & \tilde{\Gamma}_{xx}^y &= -\alpha^2 y(1 + \alpha^2 y^2), \\ \tilde{\Gamma}_{xy}^x &= \tilde{\Gamma}_{yx}^x = \alpha^2 y, & \tilde{\Gamma}_{xy}^y &= \tilde{\Gamma}_{yx}^y = \alpha^3 y^2, \\ \tilde{\Gamma}_{yy}^x &= -\alpha, & \tilde{\Gamma}_{yy}^y &= -\alpha^2 y. \end{aligned}$$

У соответствующего тензора кривизны со всеми опущенными индексами только одна компонента отлична от нуля:

$$\tilde{R}_{xyxy} = \alpha^2,$$

что можно переписать в ковариантном виде

$$\tilde{R}_{ABCD} = \alpha^2 (g_{AC}g_{BD} - g_{AD}g_{BC}).$$

Тем самым групповое многообразие $\mathbb{G} \approx \mathbb{R}^2$ с метрикой (8.60) становится пространством постоянной кривизны. Соответствующий тензор Риччи и скалярная кривизна нетривиальны:

$$\begin{aligned} \tilde{R}_{xx} &= \alpha^2(1 + \alpha^2 y^2), \\ \tilde{R}_{xy} &= \tilde{R}_{yx} = -\alpha^3 y, \\ \tilde{R}_{yy} &= \alpha^2, \\ \tilde{R} &= -2K = 2\alpha^2, \end{aligned}$$

где $K = -\alpha^2$ – гауссова кривизна двумерной группы Ли $\mathbb{G}$. Как многообразие, это – поверхность постоянной отрицательной кривизны (одна компонента двуполостного гиперболоида), которая рассмотрена в разделе 29.2.

8.8 Гомоморфизмы групп Ли

Определение. Отображение $f : \mathbb{G} \rightarrow \mathbb{H}$ называется *гомоморфизмом групп Ли*, если f гладко и в то же время является гомоморфизмом групп. Если, кроме того, f является диффеоморфизмом, то отображение f называется *изоморфизмом групп Ли*. Изоморфизм группы Ли на себя называется ее *автоморфизмом*. Ядром гомоморфного отображения называется то множество элементов группы $\mathbb{G}$, которые отображаются в единичный элемент группы $\mathbb{H}$. $\square$

Из определения следует, что при гомоморфизме групп единичный элемент группы $\mathbb{G}$ всегда отображается в единичный элемент группы $\mathbb{H}$.

Можно доказать, что из непрерывности гомоморфизма $f : \mathbb{G} \rightarrow \mathbb{H}$ следует его вещественная аналитичность [42]. Поэтому в определении гомоморфизма достаточно потребовать его непрерывность.

Теорема 8.8.1. Ядром гомоморфного отображения $\mathbb{G}_1 \rightarrow \mathbb{G}_2$ всегда является инвариантная подгруппа $\mathbb{H}_1$ группы $\mathbb{G}_1$. При этом имеет место изоморфизм

$$\frac{\mathbb{G}_1}{\mathbb{H}_1} \simeq \mathbb{G}_2.$$

Доказательство. См., например, [61], глава 1, § 3, теорема 1. $\square$

Определение. Гомоморфизм $\rho : \mathbb{G} \rightarrow \mathbb{H}$ называется *представлением* группы Ли $\mathbb{G}$, если $\mathbb{H} = \text{aut}(\mathbb{V})$ для некоторого векторного пространства $\mathbb{V}$, или $\mathbb{H} = \text{GL}(n, \mathbb{C})$, или $\mathbb{H} = \text{GL}(n, \mathbb{R})$. Если гомоморфизм $\rho : \mathbb{G} \rightarrow \rho(\mathbb{G}) \subset \mathbb{H}$ является изоморфизмом, то представление называется *точным*. Представление ρ группы Ли $\mathbb{G}$ в векторном пространстве $\mathbb{V}$ мы обозначаем парой $(\rho, \mathbb{V})$. $\square$

В определении представления во всех трех случаях каждому элементу группы $\mathbb{G}$ ставится в соответствие матрица. Если элементами матриц являются вещественные или комплексные числа, то говорят, соответственно, о *вещественном* или *комплексном* представлении. При этом все матрицы имеют отличный от нуля определитель, т.к. у каждого элемента группы должен быть обратный элемент. Следовательно, каждой матрице представления соответствует некоторый автоморфизм векторного пространства. Для точного представления группы образ $\rho(\mathbb{G}) \subset \mathbb{H}$ совсем не обязательно совпадает со всем $\mathbb{H}$.

Пусть $f : \mathbb{G} \rightarrow \mathbb{H}$ – гомоморфизм групп Ли. Так как под действием отображения f единица группы $\mathbb{G}$ переходит в единицу группы $\mathbb{H}$, то дифференциал отображения отображает касательные пространства к единицам групп:

$$f_* : T_e(\mathbb{G}) \rightarrow T_e(\mathbb{H}).$$

Замечание. Единицы разных групп мы обозначаем одним и тем же символом e , т.к. это не приводит к путанице. $\square$

Поскольку касательные пространства к единицам изоморфны алгебрам Ли, то дифференциал отображения индуцирует отображение алгебр, которое мы также будем обозначать f_* . Таким образом,

$$f_* : \mathfrak{g} \rightarrow \mathfrak{h}, \quad (8.61)$$

где образом $f_*(X)$ для $X \in \mathfrak{g}$ является единственное левоинвариантное векторное поле на $\mathbb{H}$, такое, что в единице группы $e \in \mathbb{H}$ оно определено соотношением

$$f_*(X)(e) := f_*(X(e)).$$

Предложение 8.8.1. Пусть $\mathbb{G}$ и $\mathbb{H}$ – группы Ли с алгебрами Ли $\mathfrak{g}$ и $\mathfrak{h}$, соответственно, и пусть $f : \mathbb{G} \rightarrow \mathbb{H}$ – гомоморфизм групп Ли. Тогда отображение (8.61) является гомоморфизмом алгебр Ли.

Доказательство. Дифференциал отображения линеен и сохраняет скобку Ли (2.74). $\square$

Пусть $f : \mathbb{G} \rightarrow \mathbb{H}$ – гомоморфизм групп Ли. Тогда возврат отображения f^* отображает левоинвариантные 1-формы на $\mathbb{H}$ в левоинвариантные 1-формы на $\mathbb{G}$, как отображение сопряженного пространства к $\mathfrak{h}$ в сопряженное пространство к $\mathfrak{g}$.

Пример 8.8.1 (Обозначения). Запишем вектор из векторного пространства $\mathbb{V}$ в виде $X = X^i e_i$, где e_i , $i = 1, \dots, \dim \mathbb{V}$, – некоторый фиксированный базис. Если задано представление $(\rho, \mathbb{V})$ группы Ли $\mathbb{G}$, то мы пишем

$$\mathbb{G} \ni a : \quad \mathbb{V} \ni \quad X = X^i e_i \mapsto X' = X^i S_i^j e_j \quad \in \mathbb{V}.$$

То есть каждому элементу группы $a \in \mathbb{G}$ соответствует некоторая матрица $S_i^j(a)$. По определению, все элементы матрицы гладко зависят от a . Мы пишем

$$\rho : \mathbb{G} \ni a \mapsto S_i^j(a) \in \text{aut } \mathbb{V}.$$

Ясно, что единичному элементу группы соответствует единичная матрица,

$$S_i^j(e) = \delta_i^j,$$

а обратному элементу группы – обратная матрица,

$$S_i^j(a^{-1}) = S_j^{-1j}(a).$$

Произведению двух элементов группы ab соответствует матрица

$$S_i^j(ab) = S_i^k(b) S_k^j(a),$$

которая равна произведению матриц для каждого элемента, взятых в обратном порядке. Порядок матриц является следствием принятого нами соглашения в теории групп преобразований, где элемент группы, по определению, действует на точку многообразия справа (см. раздел 9).

Дифференциал отображения ρ_* отображает алгебры Ли. В частности, генераторы алгебры Ли отображаются в матрицы,

$$\rho_* : \mathfrak{g} \ni L_A \mapsto L_{Ai}^j \in \text{end } \mathbb{V}.$$

Матрицы представления L_{Ai}^j генераторов L_A удовлетворяют тем же коммутационным соотношениям, что и левоинвариантный базис векторных полей на группе Ли:

$$[L_A, L_B]_i^j = f_{AB}^c L_{Ci}^j, \quad (8.62)$$

где f_{AB}^c – структурные константы группы Ли $\mathbb{G}$, определенные в (8.16).

Можно доказать, что матрицы представления группы Ли $\mathbb{G}$ вблизи единицы группы в линейном приближении имеют вид

$$S_i^j(da) = \delta_i^j - da^A L_{Ai}^j + \dots, \quad (8.63)$$

где $da \in \mathbb{G}$ – элемент группы, близкий к единице.

Найдем правила дифференцирования матриц представления. С одной стороны, используя разложение (8.63), имеем равенство

$$S_i^j(f(a, da)) = S_i^k(da) S_k^j(a) = (\delta_i^k - da^A L_{Ai}^k) S_k^j(a). \quad (8.64)$$

С другой стороны, матрицы представления можно разложить в ряд Тейлора:

$$S_i^j(f(a, da)) = S_i^j(a) + da^A L_A S_i^j, \quad (8.65)$$

где левоинвариантные векторные поля L_A действуют на матрицы представления, как дифференцирования, и мы воспользовались равенством

$$f^A(a, da) = a^A + da^B L_B^A + \dots$$

Сравнивая формулы (8.64) и (8.65), получаем правило дифференцирования матриц представления

$$L_A S_i^j = -L_{Ai}^k S_k^j. \quad (8.66)$$

Дифференцируя равенство $S^{-1}S = \mathbb{1}$, получаем правило дифференцирования обратных матриц представления

$$L_A S_i^{-1j} = S_i^{-1k} L_{Ak}{}^j. \quad (8.67)$$

Каждому автоморфизму $\rho(a)$ векторного пространства $\mathbb{V}$ соответствует гомоморфизм алгебр $\rho_*(a)$. Поскольку гомоморфизм алгебр сохраняет скобку Ли, то матрицы представления генераторов группы инвариантны:

$$L_{Ai}{}^j = S_A^{-1B} S_i^{-1k} L_{Bk}{}^l S_l^j, \quad (8.68)$$

где S_A^B – матрица присоединенного представления. $\square$

Продолжим общее рассмотрение.

Определение. Пара $(f, \mathbb{H})$ называется *подгруппой Ли* группы Ли $\mathbb{G}$, если выполнены следующие условия:

- 1) $\mathbb{H}$ – группа Ли;
- 2) $(f, \mathbb{H})$ – подмногообразие в $\mathbb{G}$;
- 3) $f : \mathbb{H} \rightarrow \mathbb{G}$ – групповой гомоморфизм.

Пара $(f, \mathbb{H})$ называется *замкнутой подгруппой Ли* в $\mathbb{G}$, если к тому же $f(\mathbb{H})$ – замкнутое подмногообразие в $\mathbb{G}$. $\square$

Определение замкнутого подмногообразия было дано в разделе 2.10. Если группа Ли $\mathbb{G}$ связна и $\mathbb{H}$ не изоморфна $\mathbb{G}$, то размерность замкнутой подгруппы Ли всегда меньше размерности самой группы.

Пусть $(f, \mathbb{H})$ – подгруппа Ли в $\mathbb{G}$, и $\mathfrak{h}$ и $\mathfrak{g}$ – соответствующие алгебры Ли. Тогда дифференциал отображения f_* задает изоморфизм $\mathfrak{h}$ с подалгеброй $f_*(\mathfrak{h})$ в алгебре $\mathfrak{g}$.

Гомоморфизмы групп можно описать следующим образом. Пусть $f : \mathbb{G} \rightarrow \mathbb{H}$ – сюръективное отображение. Если это отображение является взаимно однозначным, то мы имеем изоморфизм групп, $\mathbb{G} \simeq \mathbb{H}$. Если отображение f не является взаимно однозначным, то ядро этого отображения $\ker f$ представляет собой нормальную подгруппу в $\mathbb{G}$. Тогда фактор группа $\mathbb{G}/\ker f$ изоморфна $\mathbb{H}$. Если отображение f не является сюръективным, то группа $\mathbb{G}$ сюръективно отображается на какую то подгруппу $\mathbb{H}_1 \subset \mathbb{H}$. Если отображение $\mathbb{G}$ на $\mathbb{H}_1$ взаимно однозначно, то отображение f является мономорфизмом. В противном случае, $\mathbb{G}/\ker f \simeq \mathbb{H}_1$.

Напомним, что подалгеброй Ли $\tilde{\mathfrak{h}}$ в $\mathfrak{g}$ называется линейное подпространство в $\mathfrak{g}$, для которого выполнено включение $[\tilde{\mathfrak{h}}, \tilde{\mathfrak{h}}] \subset \tilde{\mathfrak{h}}$.

Теорема 8.8.2. Пусть $\mathbb{G}$ – группа Ли с алгеброй Ли $\mathfrak{g}$, и пусть $\tilde{\mathfrak{h}}$ – подалгебра Ли в $\mathfrak{g}$. Тогда существует единственная связная подгруппа Ли $(f, \mathbb{H})$ группы $\mathbb{G}$ с алгеброй $\tilde{\mathfrak{h}}$, такая, что $f_*(\tilde{\mathfrak{h}}) = \tilde{\mathfrak{h}}$.

Доказательство. Основано на теореме Фробениуса [42]. Левоинвариантные векторные поля из подалгебры $\tilde{\mathfrak{h}}$, согласно определению, находятся в инволюции. Из теоремы Фробениуса следует, что в каждой точке группы Ли они задают интегральные подмногообразия. Максимальное интегральное подмногообразие, проходящее через единицу группы, и есть подгруппа Ли $(f, \mathbb{H})$. Если $\dim \tilde{\mathfrak{h}} < \dim \mathfrak{g}$, то подгруппа Ли $\mathbb{H}$ обязательно является замкнутым подмножеством в $\mathbb{G}$. $\square$

Следствие. Существует взаимно однозначное соответствие между связными подгруппами Ли группы Ли и подалгебрами ее алгебры Ли. $\square$

Теорема 8.8.3. Пусть $(f, \mathbb{H})$ – подгруппа Ли группы Ли $\mathbb{G}$. Тогда отображение f является регулярным вложением тогда и только тогда, когда $(f, \mathbb{H})$ – замкнутая подгруппа в $\mathbb{G}$.

Доказательство. См., например, [42], теорема 3.21. $\square$

8.9 Экспоненциальное отображение для групп Ли

Определение. Гомоморфизм $f : \mathbb{R} \rightarrow \mathbb{G}$ вещественной прямой, рассматриваемой как группа Ли по сложению, в группу Ли называется *однопараметрической подгруппой* группы Ли $\mathbb{G}$. $\square$

Подчеркнем, что однопараметрическая подгруппа Ли является не подмножеством в $\mathbb{G}$, а отображением.

Рассмотрим отображение соответствующих алгебр Ли. Элемент алгебры Ли $\mathfrak{r}$ вещественной прямой $t \in \mathbb{R}$ имеет вид $s\partial_t \in \mathfrak{r}$, где $s \in \mathbb{R}$ и ∂_t – базисный вектор касательного пространства к прямой $\mathbb{R}$ в нуле. Пусть $\mathfrak{g}$ – алгебра Ли группы Ли $\mathbb{G}$, которую в данном случае удобнее отождествить с касательным пространством к единице группы. Рассмотрим произвольный элемент алгебры Ли $X \in \mathfrak{g}$. Тогда отображение

$$\mathfrak{r} \ni s\partial_t \mapsto sX \in \mathfrak{g}$$

является гомоморфизмом алгебры Ли $\mathfrak{r}$ в алгебру Ли $\mathfrak{g}$. Так как вещественная прямая односвязна, то по утверждению теоремы 8.11.8 существует единственная однопараметрическая подгруппа, для которой мы введем специальное обозначение

$$\exp_X : \mathbb{R} \rightarrow \mathbb{G}, \quad (8.69)$$

такая, что дифференциал соответствующего отображения отображает элементы алгебры Ли по следующему правилу

$$(\exp_X)_*(s\partial_t) = sX. \quad (8.70)$$

Определение. Экспоненциальным отображением алгебры Ли $\mathfrak{g}$ в группу Ли $\mathbb{G}$ называется отображение

$$\exp : \mathfrak{g} \ni X \mapsto \exp(X) := \exp_X(1) \in \mathbb{G}. \quad \square \quad (8.71)$$

Происхождение такой терминологии будет ясно из дальнейшего, когда будет показано, что для линейной группы $\mathrm{GL}(n, \mathbb{R})$ экспоненциальное отображение задается показательной функцией от матрицы.

Рассмотрим экспоненциальное отображение (8.71) с точки зрения интегральных кривых для векторных полей, введенных в разделе 2.6.5. Обозначим координаты в окрестности единицы группы через a^A , $A = 1, \dots, N$, такие, что единице группы $e \in \mathbb{G}$ соответствует начало координат. Тогда отображение (8.69) в координатах задается набором функций $a^A(t)$, которые определяются системой обыкновенных дифференциальных уравнений

$$\dot{a}^A = X^A,$$

где X^A – компоненты левоинвариантного векторного поля X , с начальными условиями $a^A(0) = 0$. Другими словами, отображение $t \mapsto \exp_X(t)$ является единственной однопараметрической подгруппой в $\mathbb{G}$ с касательным вектором $X(0)$ в нуле.

Между интегральными кривыми левоинвариантных векторных полей и операцией коммутирования в алгебре Ли существует связь.

Предложение 8.9.1. Пусть задано два левоинвариантных векторных поля $X, Y \in \mathfrak{g}$. Обозначим через $a(t)$ и $b(t)$ интегральные кривые этих векторных полей, которые проходят через единицу группы. Тогда кривая

$$c(t) := (\sqrt{t})b(\sqrt{t})a^{-1}(\sqrt{t})b^{-t}(\sqrt{t})$$

касается векторного поля $[X, Y] \in \mathfrak{g}$ в единице группы.

Доказательство. Прямая проверка. □

Свойства экспоненциального отображения дает следующая

Теорема 8.9.1. Рассмотрим произвольный элемент алгебры Ли $X \in \mathfrak{g}$ на группе Ли $\mathbb{G}$. Тогда:

- 1) $\exp(tX) = \exp_X(t)$, $\forall t \in \mathbb{R}$;
- 2) $\exp[(t_1 + t_2)X] = \exp(t_1X) \exp(t_2X)$, $\forall t_1, t_2 \in \mathbb{R}$;
- 3) $\exp(-tX) = [\exp(tX)]^{-1}$, $\forall t \in \mathbb{R}$;
- 4) отображение $\exp : \mathfrak{g} \rightarrow \mathbb{G}$ гладкое, и дифференциал отображения $\exp_* : T_0(\mathfrak{g}) \rightarrow T_e(\mathbb{G})$ – тождественное отображение (при обычных отождествлениях), так что $\exp$ задает диффеоморфизм окрестности нуля в алгебре $\mathfrak{g}$ на окрестность единицы в группе $\mathbb{G}$;
- 5) одномерное подмногообразие, полученное с помощью действия левых сдвигов $l_a \circ \exp_X \in \mathbb{G}$, – это единственная интегральная кривая левоинвариантного векторного поля X , проходящая при $t = 0$ через точку $a \in \mathbb{G}$;
- 6) однопараметрическая группа диффеоморфизмов X_t , порожденная левоинвариантным векторным полем X , задается правыми сдвигами

$$X_t = r_{\exp_X(t)}.$$

Доказательство. См., например, [42], теорема 3.31. □

Из свойства 5), в частности, следует, что все левоинвариантные векторные поля полны.

Замечание. Согласно теореме 2.6.6, векторное поле на компактном многообразии, не обращающееся в нуль, является полным. В то же время существует много примеров неполных векторных полей на некомпактных многообразиях. В случае групп Ли каждое право- или левоинвариантное векторное поле полно, даже если группа Ли некомпактна. □

Для двух коммутирующих элементов алгебры Ли,

$$[X, Y] = 0, \quad X, Y \in \mathfrak{g},$$

справедлива формула

$$\exp(X + Y) = \exp X \exp Y.$$

Экспоненциальное отображение согласовано с гомоморфизмом групп.

Теорема 8.9.2. Пусть задан гомоморфизм групп Ли $f : \mathbb{H} \rightarrow \mathbb{G}$, тогда следующая диаграмма коммутативна

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{f} & \mathbb{G} \\ \exp \uparrow & & \uparrow \exp \\ \mathfrak{h} & \xrightarrow{f_*} & \mathfrak{g} \end{array}$$

Доказательство. См., например, [42], теорема 3.32. $\square$

Следствие. Любая однопараметрическая подгруппа в $\mathbb{G}$ имеет вид экспоненциального отображения $t \mapsto \exp_X(t)$ для некоторого $X \in \mathfrak{g}$. $\square$

Если группа Ли $\mathbb{G}_0$ компактна и связна, то экспоненциальное отображение является сюръективным, т.е. отображает алгебру Ли на все многообразие группы. Как показывают явные формулы для групп $\mathrm{SO}(3)$ и $\mathrm{SU}(2)$ (см. раздел 1.8) экспоненциальное отображение не является взаимно однозначным: одному элементу группы Ли $\mathbb{G}_0$ в общем случае соответствует много различных элементов соответствующей алгебры Ли $\mathfrak{g}$. Поскольку n -мерная алгебра Ли как векторное пространство представляет собой вещественное n -мерное пространство $\mathbb{R}^n$, то при экспоненциальном отображении возникает параметризация связной компоненты единицы компактной группы Ли $\mathbb{G}_0$ точками евклидова пространства $\mathbb{R}^n$ с некоторым отношением эквивалентности. А именно, мы отождествляем те точки евклидова пространства, которые отображаются в один и тот же элемент группы. Экспоненциальное отображение является взаимно однозначным только в некоторой окрестности единицы группы $e \in \mathbb{G}_0$.

Если связная группа Ли некомпактна, то экспоненциальное отображение в общем случае не является сюръективным. Пример приведен в [64].

8.10 Интегрирование на группах Ли

Рассмотрим группу Ли $\mathbb{G}$ размерности n . В разделе 8.2 было показано, что любая группа Ли является ориентируемым многообразием. Зафиксируем некоторую ориентацию на $\mathbb{G}$ и выберем левоинвариантную дифференциальную n -форму ω на $\mathbb{G}$, согласованную с ориентацией, которая нигде не обращается в нуль. Для этого достаточно выбрать ненулевую n -форму правильной ориентации, например, в единице группы, и разнести ее по групповому многообразию с помощью действия группы слева. Напомним, что действие группы и слева, и справа сохраняет ориентацию выбранного базиса.

Запишем условие левой инвариантности формы объема ω с помощью возврата отображения. Пусть задано левое действие группы

$$l_a : \mathbb{G} \ni b \mapsto l_a b := ab \in \mathbb{G}.$$

Тогда условие левой инвариантности принимает вид

$$\omega(b) = l_a^* \omega(ab),$$

где l_a^* – возврат отображения l_a .

Пусть на группе Ли $\mathbb{G}$ задана непрерывная функция f с компактным носителем. Тогда определен интеграл

$$\int_{\mathbb{G}} \omega(b) f(b), \quad (8.72)$$

где аргумент b соответствует переменной интегрирования. Этот интеграл зависит от формы ω , которую можно умножить на постоянный положительный множитель. В случае компактных групп Ли форму ω можно однозначно фиксировать с помощью дополнительного условия

$$\int_{\mathbb{G}} \omega(b) = 1.$$

Заметим, что если ω является формой объема, то ωf – это n -форма, и на нее также действует возврат отображения.

Поскольку левое действие группы является диффеоморфизмом, то справедлива формула замены переменных интегрирования (3.74). В рассматриваемом случае она принимает вид

$$\int_{a\mathbb{G}} \omega(ab) f(ab) = \int_{\mathbb{G}} l_a^*(\omega(ab) f(ab)),$$

где в первом интеграле интегрирование ведется по ab , а во втором – по b . Поскольку $l_a^*(\omega(ab) f(ab)) = (l_a^* \omega(ab)) f(ab)$, то из равенства $a\mathbb{G} = \mathbb{G}$ и левой инвариантности формы объема вытекает равенство интегралов

$$\int_{\mathbb{G}} \omega(b) f(b) = \int_{\mathbb{G}} \omega(b) f(ab).$$

Это означает что интеграл (8.72) от функции f является левоинвариантным.

Поставим вопрос о том, когда интеграл (8.72) является одновременно и правоинвариантным, т.е. выполнено равенство

$$\int_{\mathbb{G}} \omega(b) f(b) = \int_{\mathbb{G}} \omega(b) f(ba).$$

Форма $r_a^* \omega$ по-прежнему левоинвариантна, потому что левые и правые сдвиги коммутируют между собой. Следовательно, форма $r_a^* \omega$ отличается от ω только постоянным положительным множителем, который может зависеть от a . Это значит, что на $\mathbb{G}$ существует положительная вещественная функция $\lambda(a) > 0$ такая, что

$$r_a^* \omega(ba) = \lambda(a) \omega(b).$$

Предложение 8.10.1. *Функция $\lambda(a)$ задает гомоморфизм группы Ли $\mathbb{G}$ в мультипликативную группу положительных чисел:*

$$\lambda(ab) = \lambda(a) \lambda(b). \quad (8.73)$$

Доказательство. Вычислим действие двух возвратов отображений

$$r_a^* r_b^* \omega(cab) = r_a^* \lambda(b) \omega(ca) = \lambda(a) \lambda(b) \omega(c).$$

С другой стороны, для возвратов отображений выполнено равенство

$$r_a^* r_b^* = (r_b r_a)^* = r_{ab}^*,$$

где мы воспользовались формулой (2.80). Поэтому

$$r_a^* r_b^* \omega(cab) = r_{ab}^* \omega(cab) = \lambda(ab) \omega(c).$$

Следовательно, равенство (8.73) выполнено. $\square$

Определение. Функция $\lambda(a)$ называется *модулярной функцией*. $\square$

Предложение 8.10.2. Модулярная функция связана с присоединенным представлением следующим соотношением

$$\lambda(a) = |\det \operatorname{ad}(a)|, \quad a \in \mathbb{G},$$

где $\operatorname{ad}(a)$ – матрица присоединенного представления элемента группы a .

Поскольку групповое действие является диффеоморфизмом, то справедливы равенства:

$$\int_{\mathbb{G}_a} \omega(ba) f(ba) = \int_{\mathbb{G}} r_a^*(\omega(ba) f(ba)) = \int_{\mathbb{G}} \omega(b) \lambda(a) f(ba).$$

Отсюда следует, что левоинвариантный интеграл (8.72) является также правоинвариантным тогда и только тогда, когда $\lambda(a) = 1$ для всех $a \in \mathbb{G}$. Группа Ли $\mathbb{G}$, для которой $\lambda(a) = 1$ называется *унимодулярной*.

Предложение 8.10.3. Каждая компактная группа Ли является унимодулярной.

Доказательство. Поскольку левоинвариантная форма объема на компактной группе Ли нормирована, то справедливы равенства:

$$1 = \int_{\mathbb{G}_a} \omega(ba) = \int_{\mathbb{G}} r_a^* \omega(ba) = \lambda(a) \int_{\mathbb{G}} \omega(b) = \lambda(a).$$

Отсюда следует утверждение предложения. $\square$

Следовательно, интеграл от функции (8.72) по компактной группе Ли является одновременно и право-, и левоинвариантным.

Компактные группы Ли не исчерпывают весь класс унимодулярных групп.

Предложение 8.10.4. Любая полупростая группа Ли является унимодулярной.

Доказательство. Для любой полупростой группы Ли форма Киллинга–Картана невырождена и определяет двусторонне инвариантную метрику (см. раздел 8.5) и, следовательно, меру на групповом многообразии. $\square$

Существуют также другие унимодулярные группы.

Определение. Мера интегрирования ω на группе Ли называется *левоинвариантной мерой Хаара*. В случае унимодулярной группы Ли мера ω называется *двусторонне инвариантной мерой Хаара*. $\square$

Теорема 8.10.1. Всякая локально компактная группа обладает левоинвариантной мерой Хаара ω . Если ω' – любая другая нетривиальная левоинвариантная мера Хаара, то $\omega' = c\omega$ для некоторого положительного числа c .

Доказательство. см., например, [65], глава IV, §15. $\square$

Можно проверить, что из существования левоинвариантной меры Хаара вытекает существование правоинвариантной меры Хаара и наоборот (см., например, [32], глава 2, §3).

Поскольку левые и правые сдвиги на группе Ли коммутируют, то всякая абелева локально компактная группа Ли унимодулярна.

Пример 8.10.1. Рассмотрим вещественную прямую $x \in \mathbb{R}$ как группу сдвигов. Эта группа абелева, и ее двусторонне инвариантная мера Хаара имеет стандартный вид $\omega = dx$. Условие левой и правой инвариантности интеграла от непрерывной функции с компактным носителем хорошо известно:

$$\int_{\mathbb{R}} dx f(x) = \int_{\mathbb{R}} dx f(x+h),$$

где $h \in \mathbb{R}$. □

Пример 8.10.2. Рассмотрим группу комплексных 2×2 матриц с единичным определителем $a, b \in \mathrm{SL}(2, \mathbb{C})$. Эта группа проста и на ней существует двусторонне инвариантная мера Хаара. Найдем ее явный вид. Пусть задана матрица

$$b = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C}), \quad (\alpha, \beta, \gamma, \delta) \in \mathbb{C}^4.$$

Многообразие группы Ли $\mathrm{SL}(2, \mathbb{C})$ – это гиперповерхность в $\mathbb{C}^4$, которая определяется уравнением $\alpha\delta - \beta\gamma = 1$. Определим на $\mathbb{C}^4$ 3-форму μ (вещественная размерность) следующим равенством

$$d\alpha \wedge d\beta \wedge d\gamma \wedge d\delta =: d(\alpha\delta - \beta\gamma) \wedge \mu,$$

где

$$\mu := \frac{1}{\delta} d\beta \wedge d\gamma \wedge d\delta.$$

При левом сдвиге $b \mapsto ab$ на фиксированный элемент $a \in \mathrm{SL}(2, \mathbb{C})$ форма $d\alpha \wedge d\beta \wedge d\gamma \wedge d\delta$, так же как и определитель матрицы b , сохраняется. Следовательно, 3-форма μ также сохраняется: $\mu(b) = \mu(ab)$. То же самое справедливо и для правых сдвигов $b \mapsto ba$. Таким образом, положительная форма

$$\omega(b) := \mu(b) \wedge \bar{\mu}(b) = \frac{1}{\delta\bar{\delta}} d\beta \wedge d\gamma \wedge d\delta \wedge d\bar{\beta} \wedge d\bar{\gamma} \wedge d\bar{\delta} \quad (8.74)$$

задает двусторонне инвариантную меру:

$$\omega(b) = \omega(ab) = \omega(ba).$$

Поэтому формула (8.74) задает меру Хаара на группе $\mathrm{SL}(2, \mathbb{C})$. □

Пример 8.10.3 (Группа Гейзенберга). Группа Гейзенберга $\mathbb{H}_3(\mathbb{R})$ состоит из верхне треугольных матриц вида

$$g(a, b, c) := \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}, \quad a, b, c \in \mathbb{R},$$

с обычным правилом умножения. Эта группа трехмерна некомпактна и, как многообразие, диффеоморфна евклидову пространству $\mathbb{R}^3$. Следовательно, она связна и односвязна.

Функция композиции для группы Гейзенберга записывается в явном виде:

$$(a, b, c)(a', b', c') = (a + a', b + b', c + c' + ab').$$

Отсюда видно, что группа Гейзенберга является неабелевой.

Алгебра Ли группы Гейзенберга состоит из верхне треугольных матриц вида

$$g(a, b, c) := \begin{pmatrix} 0 & \alpha & \gamma \\ 0 & 0 & \beta \\ 0 & 0 & 0 \end{pmatrix}, \quad \alpha, \beta, \gamma \in \mathbb{R},$$

Нетрудно проверить, что якобиан отображений $g \mapsto g'g$ и $g \mapsto gg'$ равен единице. Поэтому евклидова мера в $\mathbb{R}^3$, заданная формулой

$$dg(a, b, c) := da db dc,$$

является двусторонне инвариантной мерой Хаара. Следовательно, группа Гейзенберга унимодулярна.

В. Гейзенберг использовал эту группу в квантовой механике для описания одномерных квантово механических систем. $\square$

Предложение 8.10.5. *Для унимодулярной группы справедливо равенство*

$$\int_{\mathbb{G}} \omega(b) f(b) = \int_{\mathbb{G}} \omega(b) f(b^{-1}). \quad (8.75)$$

Доказательство. См., например, [32], глава II, §3, утверждение 3. $\square$

Пример 8.10.4. Рассмотрим положительные числа $x \in \mathbb{R}_+$ как абелеву группу по умножению. Каждому элементу $y \in \mathbb{R}_+$ соответствует диффеоморфизм (действие группы слева):

$$\mathbb{R}_+ \ni x \mapsto yx \in \mathbb{R}_+.$$

Отсюда вытекает выражение для меры Хаара $\omega = dx/x$. Пусть $f(x)$ – непрерывная функция с компактным носителем в $\mathbb{R}^\times$. Тогда

$$\int_0^\infty \frac{dx}{x} f(x) = - \int_\infty^0 \frac{dy}{y} f\left(\frac{1}{y}\right) = \int_0^\infty \frac{dy}{y} f\left(\frac{1}{y}\right),$$

где $x := 1/y$. Эта инвариантность интеграла соответствует формуле (8.75). $\square$

Типичное применение интеграла по компактной группе Ли состоит в следующем.

Определение. Рассмотрим представление $\rho : \mathbb{G} \rightarrow \text{aut}(\mathbb{V})$ группы Ли $\mathbb{G}$ в группе автоморфизмов вещественного или комплексного векторного пространства $\mathbb{V}$, наделенного скалярным произведением. Представление ρ называется *унитарным* (когда $\mathbb{V}$ комплексное) или *ортогональным* (когда $\mathbb{V}$ вещественное), если для всех $u, v \in \mathbb{V}$ и всех $a \in \mathbb{G}$ выполнено равенство

$$(\rho(a)u, \rho(a)v) = (u, v). \quad \square \quad (8.76)$$

Теорема 8.10.2. *Пусть $\mathbb{G}$ – компактная группа Ли, и $\mathbb{V}$ – комплексное (вещественное) векторное пространство, в котором задано представление ρ группы $\mathbb{G}$. Тогда на $\mathbb{V}$ существует скалярное произведение, относительно которого представление ρ унитарно (ортогонально).*

Доказательство. Рассмотрим произвольное скалярное произведение $\langle \cdot, \cdot \rangle$ на $\mathbb{V}$. С его помощью определим новое скалярное произведение следующей формулой

$$(u, v) := \int_{\mathbb{G}} \omega(a) \langle \rho(a)u, \rho(a)v \rangle,$$

где интегрирование ведется по $a \in \mathbb{G}$ с двусторонне инвариантной мерой Хаара ω . Используя правую инвариантность интеграла, нетрудно убедиться в справедливости равенства (8.76). $\square$

8.11 Некоторые общие свойства групп Ли

Существует определенная связь между группами Ли, локальными группами Ли и алгебрами Ли. Сейчас мы приведем несколько утверждений об общем устройстве групп Ли и о связи между группами Ли и их алгебрами Ли.

Теорема 8.11.1. *Пусть $\mathbb{G}$ – связная группа Ли, и $\mathbb{U}$ – некоторая окрестность единицы. Тогда*

$$\mathbb{G} = \bigcup_{k=1}^{\infty} \mathbb{U}^k, \quad (8.77)$$

где множество элементов $\mathbb{U}^k$ состоит из всех возможных произведений k элементов из $\mathbb{U}$.

Доказательство. См., например, [42], предложение 3.18. $\square$

Основываясь на этой теореме, иногда говорят, что произвольная окрестность единицы группы Ли (соответствующая локальная группа Ли) порождает всю связную компоненту единицы. Это не так. Знание локальной группы Ли недостаточно для построения всей связной компоненты единицы группы. Например, группы Ли $\mathrm{SO}(3)$ и $\mathrm{SU}(2)$ имеют одинаковые локальные группы Ли, однако не изоморфны. Дело в том, что в формуле (8.77) есть произведения большого числа элементов локальной группы Ли, которые лежат вне $\mathbb{U}$. То есть для построения связной компоненты единицы по формуле (8.77) необходимо знание правила умножения не только элементов локальной группы Ли, но и всех других элементов группы.

В то же время локальная группа Ли, как мы увидим ниже, однозначно с точностью до изоморфизма определяет связную и односвязную группу Ли (универсальную накрывающую).

Теорема 8.11.2 (Э. Картан). *Каждая алгебра Ли $\mathfrak{g}$ является алгеброй Ли некоторой группы Ли $\mathbb{G}$.*

Доказательство. См., например, [61], глава 10, теорема 96. $\square$

Важность понятия локальной группы Ли заключается в следующем утверждении, которое является следствием теоремы Картана.

Теорема 8.11.3. *Каждая алгебра Ли является алгеброй Ли некоторой локальной группы Ли. Локальные группы Ли изоморфны тогда и только тогда, когда изоморфны их алгебры Ли.*

Доказательство. Приведено в разделе 8.2. $\square$

Другими словами, каждая алгебра Ли определяет локальную группу Ли с точностью до изоморфизма, но ни в коем случае группу Ли в целом.

В общем случае справедлива следующая

Теорема 8.11.4. *Каждой алгебре Ли $\mathfrak{g}$ соответствует единственная, с точностью до изоморфизма, связная односвязная группа Ли $\tilde{\mathbb{G}}$ (универсальная накрывающая), для которой $\mathfrak{g}$ является алгеброй Ли. Все связные группы Ли, для которых $\mathfrak{g}$ является алгеброй Ли имеют вид $\tilde{\mathbb{G}}/\mathbb{D}$, где $\mathbb{D}$ – дискретный нормальный делитель, лежащий в центре группы Ли $\tilde{\mathbb{G}}$.*

Доказательство. См., например, [10], §6, раздел 6.3, теорема 2. $\square$

Напомним, что центром группы $\mathbb{G}$ называется множество всех элементов группы, которые перестановочны со всеми элементами из $\mathbb{G}$. Слово дискретный в сформулированной теореме означает, что подмножество $\mathbb{D} \subset \mathbb{G}$ состоит из отдельных точек.

Таким образом, существует взаимно однозначное соответствие между классами изоморфных алгебр Ли и классами изоморфных связных и односвязных групп Ли.

При описании связи между группами Ли, локальными группами Ли и алгебрами Ли используется

Теорема 8.11.5 (о монодромии). Пусть $\mathbb{G}$ – связная односвязная группа Ли, и $\mathbb{H}$ – произвольная группа Ли. Тогда всякий локальный гомоморфизм $\mathbb{G}$ в $\mathbb{H}$ (т.е. гомоморфизм соответствующих локальных групп), однозначно продолжается до глобального гомоморфизма $\mathbb{G}$ в $\mathbb{H}$.

Доказательство. См., например, [10], §6, раздел 6.3. $\square$

В теореме 8.11.4 мы упомянули универсальную накрывающую группы Ли. Общая теория накрытий для многообразий рассмотрена далее в главе 11. Ниже мы сформулируем несколько утверждений, касающихся накрытий групп Ли.

Определение. Непрерывное отображение топологических пространств $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ называется *накрытием*, если выполнены следующие условия:

- 1) p сюръективно;
- 2) для любого $x \in \mathbb{M}$ найдется открытая окрестность $\mathbb{U}_x$ точки x такая, что $p^{-1}(\mathbb{U}_x) = \bigcup_{j \in \mathcal{J}} \mathbb{U}_j$ для некоторого семейства открытых подмножеств $\mathbb{U}_j \subset \tilde{\mathbb{M}}$, удовлетворяющих условиям $\mathbb{U}_j \cap \mathbb{U}_k = \emptyset$ при $j \neq k$ и сужение отображения $p|_{\mathbb{U}_j} : \mathbb{U}_j \rightarrow \mathbb{U}_x$ – гомеоморфизм для всех $j \in \mathcal{J}$.

Топологическое пространство $\mathbb{M}$ называется *базой* накрытия, а $\tilde{\mathbb{M}}$ – *накрывающим пространством*. Если топологическое пространство $\tilde{\mathbb{M}}$ является односвязным, то накрытие называется *универсальным*. $\square$

В рассматриваемом случае базой накрытия является группа Ли $\mathbb{M} = \mathbb{G}$. Если накрывающее пространство односвязно, то на нем можно индуцировать групповую структуру, которая превращает накрывающее пространство $\tilde{\mathbb{M}} = \mathbb{G}$ в группу Ли, а накрывающее отображение p в гомоморфизм групп. Доказательство следующих трех теорем приведено в [42].

Теорема 8.11.6. Каждая связная группа Ли $\mathbb{G}$ обладает универсальным накрывающим пространством $\tilde{\mathbb{G}}$, которое само является группой Ли, причем накрывающее отображение – гомоморфизм групп Ли.

Теорема 8.11.7. Пусть $\mathbb{G}$ и $\mathbb{H}$ – связные группы Ли, и $f : \mathbb{G} \rightarrow \mathbb{H}$ – гомоморфизм. Отображение f является накрытием тогда и только тогда, когда дифференциал отображения $f_* : T_e(\mathbb{G}) \rightarrow T_e(\mathbb{H})$ является изоморфизмом касательных пространств к единицам групп.

Теорема 8.11.8. Пусть $\mathbb{G}$ и $\mathbb{H}$ – связные группы Ли с алгебрами Ли $\mathfrak{g}$ и $\mathfrak{h}$, и пусть группа Ли $\mathbb{G}$ является односвязной. Пусть $\varphi : \mathfrak{g} \rightarrow \mathfrak{h}$ – гомоморфизм алгебр Ли. Тогда существует единственный гомоморфизм групп Ли $f : \mathbb{G} \rightarrow \mathbb{H}$, такой, что $f_* = \varphi$.

Следствие. Если алгебры Ли связных и односвязных групп Ли $\mathbb{G}$ и $\mathbb{H}$ изоморфны, то сами группы $\mathbb{G}$ и $\mathbb{H}$ изоморфны. $\square$

В сформулированных утверждениях важно, чтобы группы Ли были связны, т.к. в противном случае теоремы не верны.

Пусть $e \in \mathbb{G}$ – единица группы Ли. Согласно определению накрытия, множество прообразов $p^{-1}(e) \in \tilde{\mathbb{G}}$ является конечным или счетным. В случае гомоморфизма групп $\tilde{\mathbb{G}} \rightarrow \mathbb{G}$ справедливо равенство

$$p(\tilde{a}\tilde{b}) = p(\tilde{a})p(\tilde{b}), \quad \forall \tilde{a}, \tilde{b} \in \tilde{\mathbb{G}}.$$

Из этого равенства немедленно следует, что элементы прообраза $p^{-1}(e)$ образуют абелеву подгруппу в $\tilde{\mathbb{G}}$. Более того, эта абелева подгруппа лежит в центре группы Ли $\tilde{\mathbb{G}}$. Введем для нее обозначение $\mathbb{D} := p^{-1}(e)$. Поскольку $\mathbb{D}$ – нормальная подгруппа, то универсальное накрытие групп Ли всегда можно представить в виде фактор группы

$$\mathbb{G} = \frac{\tilde{\mathbb{G}}}{\mathbb{D}},$$

где $\tilde{\mathbb{G}}$ – универсальная накрывающая группы Ли $\mathbb{G}$, а $\mathbb{D}$ – дискретный нормальный делитель, лежащий в центре группы Ли $\mathbb{G}$.

Забегая вперед, отметим, что из теоремы 11.1.4 немедленно вытекает

Теорема 8.11.9. Для любой связной группы Ли $\mathbb{G}$ ее фундаментальная группа $\pi_1(\mathbb{G})$ изоморфна группе $\mathbb{D}$ и, следовательно, абелева и лежит в центре универсальной накрывающей группы $\tilde{\mathbb{G}}$.

Следующая теорема говорит о том как устроены некоторые некомпактные группы Ли с топологической точки зрения.

Теорема 8.11.10. Пусть $\mathbb{G}$ – связная некомпактная полупростая группа Ли $\mathbb{G}$ и $\mathbb{K} \subset \mathbb{G}$ – ее некоторая максимальная компактная подгруппа. Тогда существует такое подмногообразие $\mathbb{U}$ в $\mathbb{G}$, диффеоморфное евклидову пространству, что отображение

$$\mathbb{K} \times \mathbb{U} \ni (k, u) \mapsto ku \in \mathbb{G}$$

является диффеоморфизмом прямого произведения $\mathbb{K} \times \mathbb{U}$ на $\mathbb{G}$.

Доказательство. См., например, [66], глава VI, теорема 2.2. $\square$

Данная теорема означает, что, как многообразие, любая связная некомпактная полупростая группа Ли представляет собой прямое произведение

$$\mathbb{G} \approx \mathbb{K} \times \mathbb{R}^{\dim \mathbb{G} - \dim \mathbb{K}} \quad (8.78)$$

некоторого компактного многообразия и евклидова пространства.

Пример 8.11.1. Рассмотрим произвольную $n \times n$ матрицу $M \in \mathrm{SL}(n, \mathbb{R})$ с единичным определителем. Согласно теореме о полярном разложении матриц 37.1.5, она единственным образом представляется в виде произведения $M = SA$, где $S \in \mathrm{SO}(n)$ – некоторая специальная ортогональная матрица и A – симметричная положительно определенная матрица с единичным определителем, $\det A = 1$. Группа $\mathrm{SL}(n, \mathbb{R})$ связна, некомпактна и проста. Группа $\mathrm{SO}(n)$ связна, компактна и является подгруппой в

$\mathrm{SL}(n, \mathbb{R})$. Ее размерность равна $n(n-1)/2$. Множество симметричных положительно определенных матриц с единичным определителем диффеоморфно евклидову пространству $\mathbb{R}^{\frac{n(n+1)}{2}-1}$. Таким образом, как многообразие группа $\mathrm{SL}(n, \mathbb{R})$ диффеоморфна прямому произведению

$$\mathrm{SL}(n, \mathbb{R}) \approx \mathrm{SO}(n) \times \mathbb{R}^{\frac{n(n+1)}{2}-1}.$$

□

8.12 Полупрямое произведение групп

Введем важное понятие полупрямого произведения двух групп, которое обобщает прямое произведение групп, определенное в разделе 1.2. Для этого нам понадобятся несколько предварительных определений и утверждений. Напомним, что сюръективное гомоморфное отображение группы (в том числе группы Ли) на себя называется автоморфизмом. При этом единица группы отображается в себя.

Пример 8.12.1. Рассмотрим группу $\mathbb{G}$ и произвольный элемент $b \in \mathbb{G}$. Тогда преобразование подобия

$$a \mapsto bab^{-1},$$

как легко проверить, задает автоморфизм. Такие автоморфизмы называются *внутренними*. Все другие виды автоморфизмов называются *внешними*. □

Теорема 8.12.1. *Множество автоморфизмов группы $\mathbb{G}$ само образуют группу, которая называется группой автоморфизмов группы $\mathbb{G}$ и обозначается $\mathrm{aut} \mathbb{G}$.*

Доказательство. Проводится простой проверкой групповых аксиом, причем единицей группы $\mathrm{aut} \mathbb{G}$ служит тождественное отображение $\mathrm{id} : \mathbb{G} \rightarrow \mathbb{G}$. □

Поскольку внутренние автоморфизмы сами образуют группу, то они образуют подгруппу группы $\mathrm{aut} \mathbb{G}$. Если группа $\mathbb{G}$ допускает гомоморфное отображение на свою подгруппу $\mathbb{H} \subset \mathbb{G}$, то говорят, что она допускает *эндоморфизм*.

Прямое произведение групп (см. раздел 1.2), было определено для произвольных групп. Полупрямое произведение двух групп $\mathbb{G}$ и $\mathbb{H}$ определено только в том случае, когда $\mathbb{G}$ является группой автоморфизмов $\mathbb{H}$.

Определение. Рассмотрим две группы $\mathbb{H}$ и $\mathbb{G} \subset \mathrm{aut} \mathbb{H}$. Обозначим их элементы через $a, b, \dots$ и $\alpha, \beta, \dots$, соответственно. Тогда *полупрямым произведением* $\mathbb{G} \ltimes \mathbb{H}$ групп $\mathbb{G}$ и $\mathbb{H}$ называются все возможные упорядоченные пары (α, a) со следующим законом композиции

$$(\mathbb{G} \ltimes \mathbb{H}) \times (\mathbb{G} \ltimes \mathbb{H}) \ni (\alpha, a) \ltimes (\beta, b) := (\alpha\beta, \alpha(b)a) \in \mathbb{G} \ltimes \mathbb{H},$$

где $\alpha(b)$ – образ элемента b при автоморфизме, соответствующем элементу α . □

Нетрудно проверить групповые аксиомы. Прямые вычисления показывают, что полупрямое произведение трех элементов ассоциативно:

$$(\alpha, a) \ltimes (\beta, b) \ltimes (\gamma, c) = (\alpha\beta\gamma, \alpha(c)\beta(c)\alpha(b)a).$$

Единица полупрямого произведения определяется единицами в группах $\mathbb{G}$ и $\mathbb{H}$:

$$\mathbb{G} \ltimes \mathbb{H} \ni e = (e_{\mathbb{G}}, e_{\mathbb{H}}), \quad e_{\mathbb{G}} \in \mathbb{G}, \quad e_{\mathbb{H}} \in \mathbb{H}.$$

Обратный элемент имеет вид

$$(\alpha, a)^{-1} = (\alpha^{-1}, \alpha^{-1}(a^{-1})).$$

Перечислим некоторые свойства полупрямого произведения.

1). Группы $\mathbb{G}$ и $\mathbb{H}$ естественно вложены в $\mathbb{G} \ltimes \mathbb{H}$. Это вложение задается отображениями $\alpha \mapsto (\alpha, e_{\mathbb{H}})$ и $a \mapsto (e_{\mathbb{G}}, a)$. При этом группа $\mathbb{H}$ является нормальной подгруппой в $\mathbb{G} \ltimes \mathbb{H}$, и существует изоморфизм

$$\mathbb{G} \simeq \frac{\mathbb{G} \ltimes \mathbb{H}}{\mathbb{H}}.$$

2). Каждый элемент полупрямого произведения групп $g \in \mathbb{G} \ltimes \mathbb{H}$ взаимно однозначно представим в виде $g = (\alpha, a)$, где $\alpha \in \mathbb{G}$ и $a \in \mathbb{H}$. Это свойство оправдывает название “полупрямое произведение”.

3). Заданное действие группы $\mathbb{G}$ в группе $\mathbb{H}$ с помощью автоморфизмов $\text{aut } \mathbb{H}$ совпадает с действием $\mathbb{G}$ в $\mathbb{H}$, которое определяется сужением внутренних автоморфизмов в полупрямом произведении $\mathbb{G} \ltimes \mathbb{H}$ на группу $\mathbb{H}$.

Можно доказать, что любая группа, обладающая свойствами 1)–3), изоморфна полупрямому произведению некоторых подгрупп (свойство универсальности полупрямого произведения).

Введенное выше полупрямое произведение относится к произвольным группам, в том числе и к группам Ли.

Заметим, что прямое произведение является частным случаем полупрямого произведения групп. Оно возникает, если гомоморфизм $\mathbb{G} \rightarrow \text{aut } \mathbb{H}$ тривиален, т.е. каждому элементу группы $\mathbb{G}$ соответствует тождественное преобразование группы $\mathbb{H}$.

Пример 8.12.2. Группа Ли аффинных преобразований прямой, рассмотренная в разделе 8.7, является полупрямым произведением группы дилатаций (группа $\mathbb{G}$) на группу сдвигов (группа $\mathbb{H}$). Действительно, совершим два преобразования с параметрами (β, b) и (α, a) :

$$x \mapsto \beta x + b \mapsto \alpha(\beta x + b) + a = \alpha\beta x + \alpha b + a,$$

где греческими буквами обозначены параметры дилатаций. Тогда закон умножения в группе запишется в виде

$$(\alpha, a) \ltimes (\beta, b) = (\alpha\beta, \alpha b + a),$$

что соответствует полупрямому произведению. $\square$

Пример 8.12.3. В физике важную роль играет группа Пуанкаре $\mathbb{IO}(1, n-1)$, $n \geq 2$, рассмотренная в разделе 1.9.1 и равная полупрямому произведению группы Лоренца на группу сдвигов. При этом каждому элементу группы Лоренца соответствует вращение (автоморфизм) пространства Минковского, которое рассматривается как группа сдвигов.

Ограничимся собственной ортохронной группой Лоренца $\mathbb{SO}_0(1, n-1)$ (связной компонентой единицы). Тогда группа Пуанкаре является связной как полупрямое произведение связных подгрупп. Она не является односвязной, поскольку собственная ортохронная группа Лоренца не является односвязной.

Алгебра Ли группы Пуанкаре $\mathfrak{l} + \mathfrak{t}$ состоит из преобразований Лоренца $\mathfrak{l}$ и сдвигов $\mathfrak{t}$ со следующими коммутационными соотношениями:

$$[\mathfrak{l}, \mathfrak{l}] = \mathfrak{l}, \quad [\mathfrak{l}, \mathfrak{t}] = \mathfrak{t}, \quad [\mathfrak{t}, \mathfrak{t}] = 0.$$

Мы видим, что алгебра Ли группы Пуанкаре имеет две подалгебры $\mathfrak{l}$ и $\mathfrak{t}$. При этом подгруппа сдвигов является инвариантной, и ей соответствует идеал $\mathfrak{t}$. Факторалгебра Ли $(\mathfrak{l} + \mathfrak{t})/\mathfrak{t}$ изоморфна алгебре Лоренца $\mathfrak{so}(1, n-1)$. Линейное факторпространство $(\mathfrak{l} + \mathfrak{t})/\mathfrak{l}$ изоморфно пространству Минковского, но не является факторалгеброй Ли. $\square$

8.13 Алгебры Ли

Алгебры Ли уже давно стали важным самостоятельным разделом математики. В настоящем разделе мы приведем определения и кратко рассмотрим основные свойства алгебр Ли.

Определение. Пусть $\mathfrak{g}$ – конечномерное векторное (линейное) пространство над полем вещественных $\mathbb{R}$ или комплексных $\mathbb{C}$ чисел. Если в $\mathfrak{g}$ задана билинейная операция (коммутатор), которую мы обозначим квадратными скобками:

$$\mathfrak{g} \times \mathfrak{g} \ni X, Y \mapsto Z := [X, Y] \in \mathfrak{g}, \quad (8.79)$$

такая, что выполнены два условия:

- 1) $[X, Y] = -[Y, X]$ – антисимметрия,
- 2) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ – тождества Якоби,

то множество векторов $\mathfrak{g}$ называется вещественной или комплексной *алгеброй Ли*. *Размерностью* алгебры Ли называется ее размерность как векторного пространства. Алгебра Ли называется *абелевой* или *коммутативной*, если $[X, Y] = 0$ для всех $X, Y \in \mathfrak{g}$. $\square$

Мы уже достаточно подробно рассмотрели алгебру Ли векторных полей. Ниже мы рассматриваем алгебры Ли с абстрактной точки зрения.

В приложениях, как правило, используются вещественные алгебры и группы Ли. Поскольку алгебры Ли являются векторными пространствами, то они допускают естественную комплексификацию, которая будет описана ниже. Оказывается, что простые комплексные алгебры Ли допускают классификацию, что приводит к классификации простых вещественных алгебр Ли. Поэтому в настоящем разделе мы рассматриваем алгебры Ли над полем как вещественных, так и комплексных чисел. При этом комплексные алгебры Ли следует рассматривать, в первую очередь, как средство для изучения вещественных алгебр Ли.

Как многообразие алгебра Ли диффеоморфна евклидову пространству $\mathbb{R}^N$ и поэтому некомпактна. С другой стороны, каждая алгебра Ли является алгеброй Ли некоторой группы Ли. Поэтому введем

Определение. Алгебра Ли, изоморфная алгебре Ли некоторой компактной группы Ли, называется *компактной*. $\square$

Пример 8.13.1. Рассмотрим абелеву группу Ли $\mathbb{R}^N$ по сложению. Эта группа некомпактна. С другой стороны, N -мерный тор $\mathbb{T}^N = \mathbb{U}(1) \times \dots \times \mathbb{U}(1)$ также является абелевой группой, но уже компактной. Их алгебры Ли изоморфны и, согласно определению, компактны. $\square$

Таким образом, компактная алгебра Ли может быть также алгеброй Ли некомпактной группы. Это связано с наличием абелевых инвариантных подгрупп.

Теорема 8.13.1. *Всякая связная группа Ли с компактной полупростой алгеброй Ли компактна.*

Доказательство. См., например, [61], глава 11, §61. $\square$

Пусть e_A , $A = 1, \dots, N$, – базис в алгебре Ли $\mathfrak{g}$, $\dim_{\mathbb{C}} \mathfrak{g} = N$ или $\dim_{\mathbb{R}} \mathfrak{g} = N$, рассматриваемой как векторное пространство. Число N является (комплексной) размерностью алгебры Ли. Любой элемент алгебры Ли представим виде $X = X^A e_A \in \mathfrak{g}$, где X^A вещественные или комплексные компоненты вектора X . Коммутатор двух базисных векторов всегда можно разложить по базису:

$$[e_A, e_B] = f_{AB}^C e_C, \quad (8.80)$$

где f_{AB}^C – структурные константы алгебры Ли. Из антисимметрии коммутатора и тождеств Якоби следует, что структурные константы удовлетворяют следующим тождествам:

$$\begin{aligned} f_{AB}^C &= -f_{BA}^C, \\ f_{AB}^D f_{CD}^E + f_{BC}^D f_{AD}^E + f_{CA}^D f_{BD}^E &= 0. \end{aligned} \quad (8.81)$$

Равенство $Z = [X, Y]$ в компонентах принимает вид

$$Z^A = [X, Y]^A = X^B Y^C f_{BC}^A.$$

Несмотря на свое название, структурные константы не являются постоянными. При изменении базиса в алгебре Ли $\mathfrak{g}$ они преобразуются как компоненты тензора третьего ранга с одним контравариантным и двумя ковариантными индексами.

По определению, структурные константы антисимметричны по двум индексам. Следовательно, они могут иметь не более $N^2(N-1)/2$ линейно независимых компонент. Тождества Якоби (8.81) антисимметричны по трем индексам A, B и C , и поэтому представляют собой систему из

$$NC_N^3 = \frac{N^2(N-1)(N-2)}{6}$$

алгебраических уравнений на структурные константы. Кроме того, свобода в выборе базиса определяется невырожденной матрицей, содержащей N^2 элементов, что также приводит к уменьшению числа независимых компонент структурных констант. В низших размерностях, $N \leq 3$, все возможные структурные константы можно достаточно просто классифицировать, и, следовательно, перечислить все независимые алгебры Ли.

Алгебра Ли $\mathfrak{g}$ является кольцом (со сложением и умножением) и одновременно векторным пространством над полем вещественных или комплексных чисел. Нулевой элемент алгебры Ли $0 \in \mathfrak{g}$ является единичным элементом по отношению к сложению векторов в $\mathfrak{g}$. Однако он не является единичным элементом по отношению к умножению (коммутатору), поскольку $[X, 0] = 0 \neq X$ для всех отличных от нуля элементов $X \in \mathfrak{g}$. Следовательно, единица в алгебре Ли $\mathfrak{g}$, которую мы рассматриваем в данном случае как кольцо, отсутствует.

В алгебре Ли $\mathfrak{g}$ умножение задается коммутатором, и поэтому в общем случае она является некоммутативной алгеброй. Ясно также, что умножение в алгебре Ли, которое задается коммутатором, неассоциативно. Условие ассоциативности при этом заменяется на тождества Якоби.

Если A – произвольная конечномерная ассоциативная алгебра с законом умножения

$$A \times A \ni X, Y \mapsto X \cdot Y \in A,$$

то ее можно снабдить также структурой алгебры Ли, если положить

$$[X, Y] := X \cdot Y - Y \cdot X.$$

Обратное утверждение, вообще говоря, неверно. Не каждую алгебру Ли можно снабдить структурой ассоциативной алгебры. Примером является алгебра Ли векторных полей.

Пусть $\mathfrak{h} \subset \mathfrak{g}$ и $\mathfrak{f} \subset \mathfrak{g}$ – два линейных подпространства в $\mathfrak{g}$. Обозначим через $\mathfrak{h} + \mathfrak{f}$ и $[\mathfrak{h}, \mathfrak{f}]$ множество всех векторов вида $X + Y$ и $[X, Y]$, где $X \in \mathfrak{h}$ и $Y \in \mathfrak{f}$.

Если алгебра Ли $\mathfrak{g}$ абелева, то $[\mathfrak{g}, \mathfrak{g}] = 0$.

Каждая алгебра Ли содержит нетривиальную абелеву подалгебру $\mathfrak{h} \subset \mathfrak{g}$. Например, все элементы вида $X = aX_0$, где $a \in \mathbb{R}$ или $a \in \mathbb{C}$, которые пропорциональны некоторому фиксированному элементу $X_0 \in \mathfrak{g}$, образуют абелеву подалгебру $X \in \mathfrak{h}$, поскольку $[\mathfrak{h}, \mathfrak{h}] = 0$.

Пусть $\mathfrak{h}$ и $\mathfrak{f}$ – линейные подпространства в алгебре Ли $\mathfrak{g}$, тогда нетрудно проверить следующие соотношения:

$$[\mathfrak{h}_1 + \mathfrak{h}_2, \mathfrak{f}] \subset [\mathfrak{h}_1, \mathfrak{f}] + [\mathfrak{h}_2, \mathfrak{f}], \quad (8.82)$$

$$[\mathfrak{h}, \mathfrak{f}] \subset [\mathfrak{f}, \mathfrak{h}], \quad (8.83)$$

$$[[\mathfrak{h}, \mathfrak{f}], \mathfrak{g}] \subset [[\mathfrak{f}, \mathfrak{g}], \mathfrak{h}] + [[\mathfrak{g}, \mathfrak{h}], \mathfrak{f}]. \quad (8.84)$$

Второе включение является следствием того обстоятельства, что если $X \in \mathfrak{h}$, то также $-X \in \mathfrak{h}$. Последнее включение является следствием тождеств Якоби.

Определение. Линейное подпространство $\mathfrak{h} \subset \mathfrak{g}$ называется *подалгеброй Ли*, если выполнено включение $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$. То есть оно само является алгеброй Ли. Линейное подпространство $\mathfrak{a} \subset \mathfrak{g}$ называется *идеалом*, если выполнено включение $[\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a}$. $\square$

Замечание. Пусть $\mathfrak{a} \ni X$ является левым идеалом в алгебре Ли $\mathfrak{g}$, т.е. $[\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a}$. Тогда все элементы вида $-X$ также лежат в $\mathfrak{a}$, т.к. элементы идеала должны быть подгруппой по отношению к сложению. Поскольку для левого идеала $[\mathfrak{g}, \mathfrak{a}] = [\mathfrak{a}, \mathfrak{g}] \subset \mathfrak{a}$, то каждый левый идеал совпадает с правым и наоборот. Поэтому в алгебре Ли все идеалы являются двусторонними. $\square$

Каждая алгебра Ли $\mathfrak{g}$ содержит по крайней мере два идеала, которыми являются нулевой элемент 0 и вся алгебра Ли $\mathfrak{g}$.

Пусть $\mathfrak{a} \subset \mathfrak{g}$ и $\mathfrak{b} \subset \mathfrak{g}$ – два идеала в $\mathfrak{g}$. Тогда из включения (8.82) следует, что их сумма $\mathfrak{a} + \mathfrak{b}$ также является идеалом в $\mathfrak{g}$. Аналогично, из формулы (8.84) вытекает, что коммутатор двух идеалов $[\mathfrak{a}, \mathfrak{b}]$ также является идеалом.

Пусть $\mathfrak{h}$ – подалгебра в некоторой алгебре Ли $\mathfrak{g}$. Введем в пространстве $\mathfrak{g}$ отношение эквивалентности

$$X \sim Y \pmod{\mathfrak{h}},$$

если $X - Y \in \mathfrak{h}$. Тогда вся алгебра $\mathfrak{g}$ разлагается в непересекающиеся классы эквивалентности $[X] := X + \mathfrak{h}$. В общем случае множество классов эквивалентности $[X]$ не образует алгебры Ли. Действительно, если

$$X_1 \sim Y_1 \pmod{\mathfrak{h}}, \quad \text{т.е.} \quad X_1 = Y_1 + H_1,$$

$$X_2 \sim Y_2 \pmod{\mathfrak{h}}, \quad \text{т.е.} \quad X_2 = Y_2 + H_2,$$

где $H_{1,2} \in \mathfrak{h}$, то

$$[X_1, X_2] = [Y_1, Y_2] + [Y_1, H_2] + [H_1, Y_2] + [H_1, H_2]. \quad (8.85)$$

Отсюда следует, что отношение эквивалентности для коммутаторов

$$[X_1, X_2] \sim [Y_1, Y_2] \pmod{\mathfrak{h}} \quad (8.86)$$

в общем случае не выполняется. Однако, если подалгебра является идеалом, $\mathfrak{h} = \mathfrak{a}$, то последние три слагаемых содержатся в $\mathfrak{a}$ и условие (8.86) выполнено. Тогда множество классов эквивалентности представляет собой алгебру Ли. В этом случае множество классов эквивалентности $[X]$ называется *факторалгеброй Ли* и обозначается $\mathfrak{g}/\mathfrak{a}$.

8.13.1 Операции над алгебрами Ли

Определение. Пусть $\mathfrak{g}$ и $\mathfrak{h}$ – алгебры Ли. Тогда отображение $\varphi : \mathfrak{g} \rightarrow \mathfrak{h}$ называется *гомоморфизмом алгебр Ли*, если оно линейно и сохраняет скобку Ли, т.е. $\varphi([X, Y]) = [\varphi(X), \varphi(Y)] \in \mathfrak{h}$ для всех $X, Y \in \mathfrak{g}$. Если, кроме того, φ является взаимно однозначным отображением “на” (биекцией), то φ называется *изоморфизмом алгебр Ли*. Изоморфизм алгебры Ли на себя называется ее *автоморфизмом*. Автоморфизм φ алгебры Ли $\mathfrak{g}$ называется *инволютивным*, если $\varphi^2 = \text{id}$.

Множество векторов

$$\mathfrak{n} := \{X \in \mathfrak{g} : \varphi(X) = 0\}$$

называется *ядром* гомоморфизма φ . □

В определении гомоморфизма алгебр Ли мы не требуем гладкости отображения φ , т.к. алгебры Ли рассматриваются, как векторные пространства, а не многообразия. С другой стороны, на любом векторном пространстве можно ввести гладкую структуру многообразия, тогда из линейности отображения φ будет следовать его гладкость. Более того, из непрерывности гомоморфизма следует его вещественная аналитичность [66], глава II, теорема 2.6.

Предложение 8.13.1. Ядро $\mathfrak{n}$ всякого гомоморфизма алгебр Ли $\varphi : \mathfrak{g} \rightarrow \mathfrak{h}$ является идеалом в $\mathfrak{g}$. При этом факторалгебра $\mathfrak{g}/\mathfrak{n}$ изоморфна $\varphi(\mathfrak{g})$.

Доказательство. Пусть $X \in \mathfrak{g}$ и $Y \in \mathfrak{h}$. Тогда справедливы равенства

$$\varphi([X, Y]_{\mathfrak{g}}) = [\varphi(X), 0]_{\mathfrak{h}} = 0.$$

Следовательно, $[X, Y] \in \mathfrak{n}$. Изоморфизм алгебр просто проверяется. □

Определение. Гомоморфизм $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$ называется *представлением алгебры Ли $\mathfrak{g}$* , если $\mathfrak{h} = \text{end}(\mathbb{V})$ для некоторого векторного пространства $\mathbb{V}$, или $\mathfrak{h} = \mathfrak{gl}(n, \mathbb{C})$, или $\mathfrak{h} = \mathfrak{gl}(n, \mathbb{R})$. Если гомоморфизм $\phi : \mathfrak{g} \rightarrow \varphi(\mathfrak{g}) \subset \mathfrak{h}$ является изоморфизмом, то представление называется *точным*. □

Как и в случае представления группы каждый элемент представления алгебры является матрицей. Однако эти матрицы могут быть вырождены и, соответственно, задавать только эндоморфизм векторного пространства.

Определение. Отображение $\sigma : \mathfrak{g} \rightarrow \mathfrak{g}$ комплексной алгебры Ли $\mathfrak{g}$ на себя, удовлетворяющее условиям:

$$\begin{aligned}\sigma(aX + bY) &:= \bar{a}\sigma(X) + \bar{b}\sigma(Y), \\ \sigma([X, Y]) &:= [\sigma(X), \sigma(Y)],\end{aligned}$$

где $a, b \in \mathbb{C}$ и черта обозначает комплексное сопряжение, такое, что $\sigma^2 = \text{id}$, называется *сопряжением* в алгебре Ли. $\square$

Заметим, что отображение σ не является автоморфизмом, т.к. оно антилинейно.

Пример 8.13.2. Пусть ${}^{\mathbb{C}}\mathfrak{g}$ – комплексификация вещественной алгебры Ли $\mathfrak{g}$. Тогда отображение

$${}^{\mathbb{C}}\mathfrak{g} \ni X + iY \mapsto \sigma(X + iY) := X - iY \in {}^{\mathbb{C}}\mathfrak{g},$$

где $X, Y \in \mathfrak{g}$, является сопряжением в алгебре Ли ${}^{\mathbb{C}}\mathfrak{g}$. $\square$

Определение. Дифференцированием D в алгебре Ли $\mathfrak{g}$ называется линейное отображение алгебры Ли в себя, удовлетворяющее условию (правило Лейбница)

$$D[X, Y] = [DX, Y] + [X, DY],$$

для всех $X, Y \in \mathfrak{g}$. $\square$

Пусть в алгебре Ли $\mathfrak{g}$ задано два дифференцирования D_1 и D_2 . Тогда их линейная комбинация $aD_1 + bD_2$, как легко проверить, также будет дифференцированием в $\mathfrak{g}$. Кроме этого справедливо равенство

$$\begin{aligned}D_1D_2[X, Y] &= D_1([D_2X, Y] + [X, D_2Y]) = \\ &= [D_1D_2X, Y] + [D_2X, D_1Y] + [D_1X, D_2Y] + [X, D_1D_2Y].\end{aligned}\quad (8.87)$$

Если из этого равенства вычесть такое же равенство с заменой индексов $1 \leftrightarrow 2$, то получим соотношение

$$[D_1, D_2][X, Y] = [[D_1, D_2]X, Y] + [X, [D_1, D_2]Y].$$

То есть коммутатор двух дифференцирований снова является дифференцированием в алгебре Ли $\mathfrak{g}$. Таким образом, множество всех дифференцирований алгебры Ли $\mathfrak{g}$ само образует алгебру Ли, которую мы обозначим $\mathfrak{g}_A$.

Предложение 8.13.2. Алгебра дифференцирований $\mathfrak{g}_A$ является алгеброй Ли группы Ли $\mathbb{G}_A$ всех автоморфизмов исходной алгебры Ли $\mathfrak{g}$.

Доказательство. Пусть $\varphi_t := \exp(At)$, – однопараметрическая группа автоморфизмов $\mathfrak{g}$. Тогда выполнено равенство

$$\varphi_t([X, Y]) = [\varphi_t(X), \varphi_t(Y)].\quad (8.88)$$

Дифференцирование этого соотношения по t при $t = 0$ приводит к равенству

$$A[X, Y] = [AX, Y] + [X, AY].$$

Это означает, что генератор A однопараметрической группы диффеоморфизмов является дифференцированием в алгебре Ли $\mathfrak{g}$, т.е. $A \in \mathfrak{g}_A$.

Обратно. Пусть $A \in \mathfrak{g}_A$ – дифференцирование. Тогда можно доказать, что соответствующая однопараметрическая группа преобразований удовлетворяет равенству (8.88). $\square$

Определение. Пусть $\mathfrak{g}$ – алгебра Ли. Каждому элементу алгебры Ли $X \in \mathfrak{g}$ поставим в соответствие отображение по следующему правилу

$$\mathfrak{g} \ni Y \mapsto \text{ad } X(Y) := [X, Y] \in \mathfrak{g}. \quad (8.89)$$

Следовательно, определено отображение

$$\text{ad} : \mathfrak{g} \ni X \mapsto \text{ad } X \in \text{end } \mathfrak{g} \quad (8.90)$$

Это отображение называется *присоединенным* представлением алгебры Ли $\mathfrak{g}$. $\square$

Проверим, что данное отображение действительно является представлением. Из тождеств Якоби следует равенство

$$\text{ad } [X, Y](Z) = \text{ad } X(\text{ad } Y(Z)) - \text{ad } Y(\text{ad } X(Z)).$$

Это означает, что

$$\text{ad } [X, Y] = \text{ad } X(\text{ad } Y) - \text{ad } Y(\text{ad } X),$$

и отображение $X \mapsto \text{ad } X$ – действительно представление.

Если в алгебре Ли $\mathfrak{g}$ выбран некоторый базис e_A , $A = 1, \dots, N$, то каждое отображение $\text{ad } X$ будет задаваться некоторой $N \times N$ матрицей.

С другой стороны, тождества Якоби можно переписать в виде равенства

$$\text{ad } X([Y, Z]) = [\text{ad } X(Y), Z] + [Y, \text{ad } X(Z)].$$

Это означает, что присоединенное представление алгебры Ли является дифференцированием в $\mathfrak{g}$. Таким образом, множество элементов присоединенного представления образует некоторую подалгебру в алгебре всех дифференцирований $\mathfrak{g}_A$.

Определение. Множество элементов присоединенного представления

$$\mathfrak{g}_{\text{ad}} := \{ \text{ad } X, \quad X \in \mathfrak{g} \}$$

называется *присоединенной* алгеброй Ли. $\square$

Отображение

$$\psi : \mathfrak{g} \ni X \mapsto \text{ad } X \in \mathfrak{g}_{\text{ad}}$$

представляет собой гомоморфизм алгебр Ли. Ядро этого гомоморфизма является центром алгебры Ли $\mathfrak{g}$.

Если $\varphi : \mathfrak{g} \rightarrow \mathfrak{g}$ – произвольный автоморфизм алгебры Ли, то справедлива следующая цепочка равенств:

$$\text{ad } \varphi(X)(Y) = [\varphi(X), Y] = \varphi[X, \varphi^{-1}(Y)] = \varphi(\text{ad } X(\varphi^{-1}(Y))).$$

Это означает, что

$$\text{ad } \varphi(X) = \varphi \text{ad } X \varphi^{-1}, \quad (8.91)$$

как и положено присоединенному представлению.

Если у алгебры Ли $\mathfrak{g}$ нулевой центр (т.е. единственный элемент, коммутирующий со всеми элементами алгебры – это нуль), то гомоморфизм (8.90) инъективен, и присоединенное представление является точным представлением алгебры Ли $\mathfrak{g}$ в $\text{end } \mathfrak{g}$.

Пример 8.13.3. Рассмотрим трехмерную алгебру Ли $\mathfrak{su}(2)$. Выберем в ней базис

$$e_i := \frac{i}{2} \sigma_i \in \mathfrak{su}(2),$$

где σ_i , $i = 1, 2, 3$, – матрицы Паули. Этот базис удовлетворяет коммутационным соотношениям

$$[e_i, e_j] = -\varepsilon_{ijk} e^k,$$

где ε_{ijk} – полностью антисимметричный тензор третьего ранга и подъем индексов осуществляется с помощью евклидовой метрики: $e^k := \delta^{ki} e_i$. Согласно определению, $\text{ad } e_1(e_1) = 0$, $\text{ad } e_1(e_2) = -e_3$ и $\text{ad } e_1(e_3) = e_2$. Аналогичные соотношения можно написать для $\text{ad } e_2$ и $\text{ad } e_3$. Следовательно, присоединенная алгебра $\mathfrak{su}_{\text{ad}}(2)$ трехмерна, и присоединенное представление базисных векторов имеет вид

$$\text{ad } e_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \text{ad } e_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{ad } e_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Можно проверить, что эти матрицы удовлетворяют тем же коммутационным соотношениям

$$[\text{ad } e_i, \text{ad } e_j] = -\varepsilon_{ijk} \text{ad } e^k.$$

Заметим, что присоединенное представление алгебры $\mathfrak{su}(2)$ совпадает с фундаментальным представлением алгебры Ли группы вращений $\mathfrak{so}(3)$, состоящим из всех антисимметричных 3×3 матриц. $\square$

Пример 8.13.4. Рассмотрим вещественную алгебру Ли $\mathfrak{g}$ с базисом e_A , $A = 1, \dots, N$. Базис алгебры Ли удовлетворяет коммутационным соотношениям (8.80) с некоторыми вещественными структурными константами f_{AB}^C . Рассмотрим отображение

$$\mathfrak{g} \ni e_A \mapsto E_A \in \mathfrak{gl}(N, \mathbb{R}),$$

где матрицы E_A определяются структурными константами: $E_{AB}^C := -f_{AB}^C$. При этом индексы B, C рассматриваются как матричные. Тогда из тождеств Якоби для структурных констант (8.81) следует следующее правило коммутации

$$[E_A, E_B] = f_{AB}^C E_C.$$

Это означает, что построенное отображение является представлением алгебры Ли:

$$\text{ad} : \mathfrak{g} \ni X = X^A e_A \mapsto \text{ad } X = X^A E_A \in \mathfrak{gl}(N, \mathbb{R}).$$

Нетрудно проверить, что это действительно присоединенное представление. $\square$

Теперь введем новое понятие для двух алгебр Ли.

Определение. Пусть $\mathfrak{g}$ и $\mathfrak{h}$ – две алгебры Ли. Рассмотрим их прямую сумму как векторных пространств $\mathfrak{g} \oplus \mathfrak{h}$. Введем в этом пространстве операцию коммутирования с помощью коммутирований, определенных в $\mathfrak{g}$ и $\mathfrak{h}$:

$$[X_1 \oplus Y_1, X_2 \oplus Y_2] := [X_1, X_2] \oplus [Y_1, Y_2],$$

для всех $X_{1,2} \in \mathfrak{g}$ и $Y_{1,2} \in \mathfrak{h}$. Тогда алгебра Ли $\mathfrak{g} \oplus \mathfrak{h}$ называется прямой суммой алгебр Ли $\mathfrak{g}$ и $\mathfrak{h}$. $\square$

Корректность определения коммутатора в прямой сумме, т.е. билинейность, антисимметрия и тождества Якоби, легко проверяется.

Определение. Если алгебру Ли $\mathfrak{g}$ как векторное пространство можно представить в виде прямой суммы векторных подпространств

$$\mathfrak{g} = \bigoplus_{i=1}^k \mathfrak{g}_i$$

и, кроме того, справедливы включения

$$[\mathfrak{g}_i, \mathfrak{g}_i] \subset \mathfrak{g}_i, \quad [\mathfrak{g}_i, \mathfrak{g}_j] = 0, \quad i \neq j,$$

то мы говорим, что алгебра Ли $\mathfrak{g}$ разлагается в прямую сумму подалгебр $\mathfrak{g}_i$. $\square$

Ясно, что все подалгебры Ли $\mathfrak{g}_i$ являются идеалами в $\mathfrak{g}$, потому что

$$[\mathfrak{g}, \mathfrak{g}_i] = [\mathfrak{g}_i, \mathfrak{g}_i] \subset \mathfrak{g}_i.$$

Кроме того, если $\mathfrak{a}$ – идеал в одной из подалгебр $\mathfrak{g}_i$, то он также является идеалом во всей алгебре Ли $\mathfrak{g}$.

Рассмотрим более сложную конструкцию.

Определение. Пусть $\mathfrak{g}$ и $\mathfrak{t}$ – две алгебры Ли, и пусть $D : \mathfrak{g} \rightarrow \mathfrak{t}_A$ – гомоморфизм алгебры Ли $\mathfrak{g}$ в алгебру дифференцирований $\mathfrak{t}_A$, т.е. $D(X) \in \mathfrak{t}_A$ для всех $X \in \mathfrak{g}$. Возьмем прямую сумму векторных пространств $\mathfrak{g} \dot{+} \mathfrak{t}$ и снабдим ее структурой алгебры Ли с помощью следующего коммутатора

$$[X_1 \dot{+} T_1, X_2 \dot{+} T_2] := [X_1, X_2] \dot{+} ([T_1, T_2] + D(X_1)T_2 - D(X_2)T_1), \quad (8.92)$$

для всех $X_{1,2} \in \mathfrak{g}$ и $T_{1,2} \in \mathfrak{t}$. Тогда $\mathfrak{g} \dot{+} \mathfrak{t}$ называется *полупрямой суммой* алгебр Ли $\mathfrak{g}$ и $\mathfrak{t}$. $\square$

Если дифференцирование тривиально, т.е. $D(X) = 0$ для всех $X \in \mathfrak{g}$, то полупрямая сумма алгебр Ли сводится к прямой сумме.

Проверим корректность данного определения. Билинейность и антисимметрия коммутатора очевидны. Необходимо проверить только тождества Якоби. Если все три вектора имеют вид $X \dot{+} 0$ или $0 \dot{+} T$, где $X \in \mathfrak{g}$ и $T \in \mathfrak{t}$, то тождества Якоби выполняются, т.к. они выполнены в каждом слагаемом $\mathfrak{g}$ и $\mathfrak{t}$. Следовательно, ввиду линейности коммутатора, достаточно проверить тождества Якоби для двух троек векторов: $\{X \dot{+} 0, Y \dot{+} 0, 0 \dot{+} T\}$ и $\{X \dot{+} 0, 0 \dot{+} T, 0 \dot{+} P\}$. Двойной коммутатор для каждой тройки будет иметь вид $0 \dot{+} *$, где $*$ – некоторый элемент из $\mathfrak{t}$. Поэтому для упрощения записи в следующих формулах мы опустим два первых символа 0 и $\dot{+}$. Для первой тройки векторов $\{X, Y, T\}$ тождества Якоби принимают вид

$$\begin{aligned} [[X, Y], T] + [[Y, T], X] + [[T, X], Y] &= \\ &= D([X, Y])T - D(X)D(Y)T + D(Y)D(X)T = 0. \end{aligned}$$

Равенство нулю следует из того, что $D(X)$ – это дифференцирование в $\mathfrak{t}$. Аналогично, для второй тройки векторов $\{X, T, P\}$ тождества Якоби принимают вид

$$\begin{aligned} [[X, T], P] + [[T, P], X] + [[P, X], T] &= \\ &= [D(X)T, P] - D(X)[T, P] + [T, D(X)P] = 0. \end{aligned}$$

Равенство нулю здесь также является следствием свойств дифференцирования.

Полупрямая сумма двух алгебр $\mathfrak{g} \dot{+} \mathfrak{t}$ содержит по крайней мере две подалгебры: $\mathfrak{g}$ и $\mathfrak{t}$. Более того, подалгебра $\mathfrak{t}$ является идеалом в $\mathfrak{g} \dot{+} \mathfrak{t}$, что является следствием определения коммутатора (8.92).

8.13.2 Простые и полупростые алгебры Ли

Простые алгебры Ли наиболее часто встречаются в математической физике и поэтому представляют особенный интерес. В настоящее время все они классифицированы. Ниже мы дадим необходимые определения и приведем некоторые свойства простых и полупростых алгебр Ли.

Начнем с необходимой конструкции. Напомним, что алгебра Ли $\mathfrak{g}$ является идеалом в себе самой и коммутатор двух идеалов также является идеалом. Поэтому формулы

$$\mathfrak{g}^{(1)} := \mathfrak{g}, \quad \mathfrak{g}^{(2)} := [\mathfrak{g}^{(1)}, \mathfrak{g}^{(1)}], \quad \dots, \quad \mathfrak{g}^{(k)} := [\mathfrak{g}^{(k-1)}, \mathfrak{g}^{(k-1)}], \quad \dots \quad (8.93)$$

определяют в алгебре Ли $\mathfrak{g}$ невозрастающую последовательность идеалов:

$$\mathfrak{g} = \mathfrak{g}^{(1)} \supset \mathfrak{g}^{(2)} \supset \dots \supset \mathfrak{g}^{(k)} \supset \dots$$

Определение. Алгебра Ли $\mathfrak{g}$ называется *разрешимой*, если существует такое число $k \geq 1$, что $\mathfrak{g}^{(k)} = 0$. $\square$

Пример 8.13.5. Рассмотрим алгебру Ли $\mathfrak{io}(2)$ неоднородной группы вращений евклидовой плоскости $\mathbb{R}^2$. Она состоит из вращений $L = x\partial_y - y\partial_x$ и сдвигов $P_x = \partial_x$, $P_y = \partial_y$. Нетривиальные коммутационные соотношения имеют вид

$$[L, P_x] = -P_y, \quad [L, P_y] = P_x.$$

Все остальные коммутаторы равны нулю. Поэтому последовательность идеалов (8.93) обрывается после второго шага:

$$\mathfrak{g}^{(1)} = \mathfrak{io}(2), \quad \mathfrak{g}^{(2)} = \mathfrak{p}, \quad \mathfrak{g}^{(3)} = 0,$$

где $\mathfrak{p}$ – двумерная абелева алгебра Ли с образующими P_x и P_y (подалгебра сдвигов). Следовательно, алгебра неоднородных вращений плоскости $\mathfrak{io}(2)$ разрешима. $\square$

Пример 8.13.6. Рассмотрим алгебру Ли трехмерных вращений $\mathfrak{so}(3)$ с образующими e_i , $i = 1, 2, 3$. Поскольку коммутационные соотношения имеют вид

$$[e_i, e_j] = -\varepsilon_{ijk}e_k,$$

где ε_{ijk} – полностью антисимметричный тензор третьего ранга, то последовательность идеалов (8.93) никогда не оборвется:

$$\mathfrak{g}^{(1)} = \mathfrak{g}^{(2)} = \mathfrak{g}^{(3)} = \dots = \mathfrak{so}(3).$$

Следовательно, алгебра трехмерных вращений $\mathfrak{so}(3)$ неразрешима. $\square$

Теперь введем другую последовательность идеалов, которую пронумеруем нижними индексами:

$$\mathfrak{g}_{(1)} := \mathfrak{g}, \quad \mathfrak{g}_{(2)} := [\mathfrak{g}_{(1)}, \mathfrak{g}], \quad \dots, \quad \mathfrak{g}_{(k)} := [\mathfrak{g}_{(k-1)}, \mathfrak{g}], \quad \dots \quad (8.94)$$

Она также является невозрастающей:

$$\mathfrak{g} = \mathfrak{g}_{(1)} \supset \mathfrak{g}_{(2)} \supset \dots \supset \mathfrak{g}_{(k)} \supset \dots$$

Определение. Алгебра Ли $\mathfrak{g}$ называется *нильпотентной*, если существует такое число $k \geq 1$, что $\mathfrak{g}_{(k)} = 0$. $\square$

Ясно, что любая абелева алгебра Ли является и разрешимой, и нильпотентной.

Предложение 8.13.3. $\mathfrak{g}^{(n)} \subset \mathfrak{g}_{(n)}$.

Доказательство. Проверяем по индукции. В самом деле $\mathfrak{g}^{(1)} = \mathfrak{g}_{(1)}$. Если $\mathfrak{g}^{(n)} \subset \mathfrak{g}_{(n)}$, то

$$\mathfrak{g}^{(n+1)} = [\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}] \subset [\mathfrak{g}_{(n)}, \mathfrak{g}] \subset \mathfrak{g}_{(n+1)}. \quad \square$$

Следствие. Всякая нильпотентная алгебра разрешима. $\square$

Обратное утверждение неверно.

Пример 8.13.7. Рассмотрим двумерную неабелеву алгебру Ли $\mathfrak{g}$ из раздела 8.7, соответствующую аффинным преобразованиям прямой. Ее алгебра Ли имеет две образующие L_x, L_y с коммутационными соотношениями

$$[L_x, L_y] = L_y, \quad [L_x, L_x] = [L_y, L_y] = 0.$$

Легко вычислить последовательности идеалов

$$\begin{aligned} \mathfrak{g}^{(1)} &= \mathfrak{g}, & \mathfrak{g}^{(2)} &= \{L_y\}, & \mathfrak{g}^{(3)} &= 0, \\ \mathfrak{g}_{(1)} &= \mathfrak{g}, & \mathfrak{g}_{(2)} &= \mathfrak{g}_{(3)} = \dots = \{L_y\}, \end{aligned}$$

где $\{L_y\}$ – одномерная абелева алгебра Ли с образующей L_y . Таким образом, алгебра Ли $\mathfrak{g}$ разрешима, но не нильпотентна. $\square$

Пример 8.13.8. Вычислим последовательность идеалов (8.94) для алгебры Ли $\mathfrak{io}(2)$ из примера 8.13.5:

$$\mathfrak{g}^{(1)} = \mathfrak{io}(2), \quad \mathfrak{g}^{(2)} = \mathfrak{g}^{(3)} = \dots = \mathfrak{p}.$$

Поэтому алгебра Ли $\mathfrak{io}(2)$ разрешима, но не нильпотентна. $\square$

Теорема 8.13.2 (Энгель). Алгебра Ли нильпотентна тогда и только тогда, когда ее форма Киллинга–Картана тождественно равна нулю.

Доказательство. См., например, [10], §6, задача 13. $\square$

Напомним, что для любых двух идеалов $\mathfrak{a} \subset \mathfrak{g}$ и $\mathfrak{b} \subset \mathfrak{g}$ их сумма $\mathfrak{a} + \mathfrak{b}$ также является идеалом в $\mathfrak{g}$.

Предложение 8.13.4. Пусть $\mathfrak{a}$ и $\mathfrak{b}$ – два идеала в алгебре Ли $\mathfrak{g}$. Тогда факторалгебра $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b}$ изоморфна фактор алгебре $\mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$ и поэтому разрешима, если идеал $\mathfrak{a}$ разрешим.

Доказательство. Пусть φ – естественный гомоморфизм алгебры $\mathfrak{a} + \mathfrak{b}$ на факторалгебру $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b}$. Тогда $\varphi(\mathfrak{a}) = (\mathfrak{a} + \mathfrak{b})/\mathfrak{b}$. Ядром гомоморфизма φ алгебры $\mathfrak{a}$ является пересечение $\mathfrak{a} \cap \mathfrak{b}$. Поэтому из предложения 8.13.1 следует изоморфизм

$$(\mathfrak{a} + \mathfrak{b})/\mathfrak{b} \simeq \mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b}). \quad \square$$

Следствие. Если идеал $\mathfrak{b}$ также разрешим, то разрешим и идеал $\mathfrak{a} + \mathfrak{b}$. $\square$

Таким образом, сумма $\mathfrak{a} + \mathfrak{b}$ двух разрешимых идеалов $\mathfrak{a}$ и $\mathfrak{b}$ является разрешимым идеалом. Поэтому в любой конечномерной алгебре Ли $\mathfrak{g}$ существует наибольший разрешимый идеал $\mathfrak{r}$, который называется *радикалом*, содержащий все разрешимые идеалы: им является сумма всех разрешимых идеалов.

Определение. Алгебра Ли $\mathfrak{g}$ называется полупростой, если ее радикал тривиален $\mathfrak{r} = 0$. $\square$

Другими словами, полупростая алгебра Ли не содержит разрешимых идеалов, кроме тривиального $\mathfrak{r} = 0$.

Теорема 8.13.3. Пусть $\mathfrak{r}$ – радикал алгебры Ли $\mathfrak{g}$, тогда фактор алгебра Ли $\mathfrak{g}/\mathfrak{r}$ полупроста.

Теорема 8.13.4 (Э. Картан). Для того, чтобы алгебра Ли $\mathfrak{g}$ была полупростой, необходимо и достаточно, чтобы она не содержала абелевых идеалов, отличных от 0.

Доказательство. См., например, [67], лекция 17. $\square$

Конечно, каждая полупростая алгебра Ли является алгеброй Ли некоторой полупростой группы Ли, и любая полупростая группа Ли имеет полупростую алгебру Ли.

Определение. Центром алгебры Ли $\mathfrak{g}$ называется ее аннулятор, т.е. наибольшее подпространство $\mathfrak{z} \subset \mathfrak{g}$, для которого $[\mathfrak{z}, \mathfrak{g}] = 0$. $\square$

У каждой алгебры Ли $\mathfrak{g}$ есть нулевой элемент $0 \in \mathfrak{g}$, для которого $[0, \mathfrak{g}] = 0$. Поэтому $0 \in \mathfrak{z}$. Очевидно, что центр алгебры Ли $\mathfrak{g}$ является абелевым идеалом, и алгебра Ли $\mathfrak{g}$ абелева тогда и только тогда, когда ее центр совпадает со всей алгеброй $\mathfrak{z} = \mathfrak{g}$. Из теоремы 8.13.4 следует, что алгебра Ли является полупростой тогда и только тогда, когда ее центр равен нулю.

Введем понятие простой алгебры и группы Ли.

Определение. Алгебра Ли $\mathfrak{g}$ называется *простой*, если выполнены следующие два условия: алгебра $\mathfrak{g}$ неабелева, и единственные ее идеалы есть 0 и $\mathfrak{g}$.

Группа Ли $\mathbb{G}$ называется *простой*, если она неабелева и не имеет инвариантных подгрупп, отличных от единицы и всей группы. $\square$

Условие неабелевости алгебры Ли $\mathfrak{g}$ эквивалентно условию $\mathfrak{g}^{(2)} \neq 0$. Это условие исключает алгебры Ли размерности 1. Алгебры Ли размерности 1 без условия неабелевости являлись бы в этом случае простыми, но не полупростыми.

Всякие простые алгебры и группы Ли являются также полупростыми. Обратное утверждение неверно, но справедливы следующие теоремы.

Теорема 8.13.5. Полупростая алгебра Ли $\mathfrak{g}$ представима одним и только одним способом в виде прямой суммы конечного числа простых идеалов

$$\mathfrak{g} = \bigoplus_{i=1}^k \mathfrak{s}_i$$

для некоторого k . При этом каждый идеал алгебры Ли $\mathfrak{g}$ является прямой суммой некоторых идеалов $\mathfrak{s}_i$.

Теорема 8.13.6. Полупростая группа Ли $\mathbb{G}$ представима одним и только одним способом в виде прямого произведения конечного числа простых групп Ли

$$\mathbb{G} = \bigotimes_{i=1}^k \mathbb{G}_i.$$

При этом каждая инвариантная подгруппа группы Ли $\mathbb{G}$ является прямым произведением некоторых подгрупп $\mathbb{G}_i$.

8.13.3 Квадратичные формы

Поскольку алгебра Ли является векторным пространством, то на ней можно задать билинейную симметричную квадратичную форму

$$\mathfrak{g} \times \mathfrak{g} \ni X, Y \mapsto (X, Y) \in \mathbb{R}, \mathbb{C}.$$

Если e_A , $A = 1, \dots, N$ – базис алгебры Ли, то в компонентах эта форма имеет вид

$$(X, Y) = X^A Y^B \eta_{AB},$$

где η_{AB} – некоторая матрица.

В разделе 8.13.1 было введено присоединенное представление $Z \mapsto \text{ad } Z$ для любого элемента алгебры Ли $Z \in \mathfrak{g}$. Это позволяет дать следующее

Определение. Квадратичная форма (X, Y) называется *инвариантной*, если выполнено равенство

$$(\text{ad } Z(X), Y) + (X, \text{ad } Z(Y)) = 0$$

для всех $X, Y, Z \in \mathfrak{g}$. □

Поскольку алгебра Ли изоморфна касательному пространству в произвольной точке группового многообразия, то задание невырожденной симметричной квадратичной формы в алгебре Ли однозначно определяет метрику на соответствующей группе Ли. Эта метрика в левоинвариантном базисе (см. раздел 8.2) имеет постоянные компоненты η_{AB} и, по построению, является левоинвариантной. В общем случае метрика на групповом многообразии в левоинвариантном базисе имеет компоненты $g_{AB}(a)$ и зависит от точки $a \in \mathbb{G}$. То есть каждой точке $a \in \mathbb{G}$ соответствует своя метрика в алгебре Ли. Задание инвариантной квадратичной формы в алгебре Ли соответствует заданию двусторонне инвариантной метрики на групповом многообразии.

Присоединенное представление алгебры Ли позволяет определить следующую симметричную квадратичную форму.

Определение. Симметричная билинейная квадратичная форма в алгебре Ли

$$(X, Y) := -\text{tr}(\text{ad } X \text{ ad } Y) \tag{8.95}$$

называется *формой Киллинга–Картана* алгебры Ли $\mathfrak{g}$. □

В компонентах форма Киллинга–Картана имеет вид

$$\eta_{AB} = -f_{AC}^D f_{BD}^C.$$

Форма Киллинга–Картана играет фундаментальную роль в теории алгебр Ли и их представлений.

Замечание. Знак минус в определении формы Киллинга–Картана необходим для того, чтобы метрика на соответствующих полупростых вещественных компактных группах Ли была положительно, а не отрицательно определена. Например, для группы $\mathrm{SO}(3)$, структурными константами которой является полностью антисимметричный тензор третьего ранга, $f_{ab}{}^c \mapsto \varepsilon_{ij}{}^k$, форма Киллинга–Картана имеет вид

$$\eta_{ij} = -\varepsilon_{ik}{}^l \varepsilon_{jl}{}^k = 2\delta_{ij},$$

где мы воспользовались формулами Приложения 37.67. $\square$

В общем случае форма Киллинга–Картана может быть вырождена или невырождена. Это зависит от алгебры Ли. Для абелевых алгебр Ли форма Киллинга–Картана равна нулю.

Предложение 8.13.5. *Форма Киллинга–Картана (8.95) инвариантна.*

Доказательство. Следствие симметрии следа матриц:

$$\mathrm{tr}(\mathrm{ad} X \mathrm{ad} Y \mathrm{ad} Z) = \mathrm{tr}(\mathrm{ad} Z \mathrm{ad} X \mathrm{ad} Y). \quad \square$$

Форма Киллинга–Картана инвариантна относительно произвольных автоморфизмов алгебры Ли $\varphi : \mathfrak{g} \rightarrow \mathfrak{g}$. Действительно, поскольку при автоморфизме матрица присоединенного представления преобразуется по правилу (8.91), то форма Киллинга–Картана инвариантна:

$$(\varphi(X), \varphi(Y)) = (X, Y).$$

Форма Киллинга–Картана позволяет сформулировать важный критерий.

Теорема 8.13.7 (Э. Картан). *Алгебра Ли $\mathfrak{g}$ полупроста тогда и только тогда, когда ее форма Киллинга–Картана невырождена.*

Доказательство. См., например, [68], теорема VI. $\square$

В заключение приведем критерий компактности алгебр Ли.

Теорема 8.13.8. *Алгебра Ли $\mathfrak{g}$ компактна тогда и только тогда, когда в $\mathfrak{g}$ существует положительно определенная инвариантная квадратичная форма.*

Доказательство. См., например, [61], глава 11, теорема 103. $\square$

8.14 Группа Ли $\mathrm{GL}(n, \mathbb{C})$

В настоящем разделе мы приведем некоторые сведения из теории линейных групп Ли $\mathrm{GL}(n, \mathbb{C})$ преобразований (автоморфизмов) комплексного n -мерного пространства $\mathbb{C}^n$ и их алгебр Ли $\mathfrak{gl}(n, \mathbb{C})$, которые важны с точки зрения дифференциальной геометрии и физических приложений. Напомним, что при общих преобразованиях координат в касательных пространствах к точкам многообразия на компоненты тензоров действует матрица Якоби (1.68), которая является элементом группы $\mathrm{GL}(n, \mathbb{R}) \subset \mathrm{GL}(n, \mathbb{C})$. Кроме того, удобным выбором независимых переменных в аффинной геометрии являются переменные Картана: репер и линейная или $\mathrm{GL}(n, \mathbb{R})$ связность (см. раздел 5.4).

Сначала мы рассмотрим алгебры Ли, как более простые объекты, а затем перейдем к изучению соответствующих групп Ли.

8.14.1 Алгебра Ли $\mathfrak{gl}(n, \mathbb{C})$

Обозначим множество всех квадратных $n \times n$ матриц с комплексными элементами через $\mathfrak{gl}(n, \mathbb{C})$. Это обозначение обосновано в дальнейшем тем, что множество комплексных квадратных матриц можно рассматривать, как алгебру Ли группы Ли $\mathrm{GL}(n, \mathbb{C})$. Множество матриц $\mathfrak{gl}(n, \mathbb{C})$ с обычным умножением матриц на числа и сложением можно рассматривать как векторное пространство размерности n^2 над полем комплексных чисел (комплексная размерность) или как вещественное векторное пространство размерности $2n^2$ (вещественная размерность). Чтобы превратить это множество в алгебру необходимо ввести дополнительную бинарную операцию – умножение. Если в качестве умножения рассматривать обычное умножение матриц, то мы получим ассоциативную алгебру матриц над полем комплексных чисел. Однако это не единственная возможность. В качестве алгебраической операции мы будем рассматривать коммутатор матриц, т.е. любым двум матрицам $X, Y \in \mathfrak{gl}(n, \mathbb{C})$ мы ставим в соответствие их коммутатор:

$$\mathfrak{gl}(n, \mathbb{C}) \times \mathfrak{gl}(n, \mathbb{C}) \ni X, Y \mapsto [X, Y] := XY - YX \in \mathfrak{gl}(n, \mathbb{C}).$$

Эта операция антисимметрична, $[X, Y] = -[Y, X]$, и для нее выполняется тождество Якоби

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0.$$

Тем самым множество всех матриц, включая вырожденные, становится алгеброй Ли $\mathfrak{gl}(n, \mathbb{C})$. Эта алгебра неассоциативна, что следует из тождеств Якоби.

В дальнейшем мы, как правило, будем рассматривать множество комплексных $n \times n$ матриц, как алгебру Ли $\mathfrak{gl}(n, \mathbb{C})$.

Множество матриц можно рассматривать, как множество операторов, действующих в комплексном векторном пространстве $\mathbb{V}$. Обозначим элементы векторного пространства через $x = x^a e_a \in \mathbb{V}$, где e_a , $a = 1, \dots, n$ – некоторый фиксированный базис, и $x^a \in \mathbb{C}$ – комплексные компоненты вектора. Комплексная размерность этого векторного пространства равна n , а вещественная – $2n$. При фиксированном базисе векторное пространство $\mathbb{V}$ естественным образом отождествляется с комплексным евклидовым пространством $\mathbb{C}^n$ и вещественным евклидовым пространством удвоенной размерности $\mathbb{R}^{2n}$. Мы предполагаем, что топология $\mathbb{V}$ индуцируется взаимно однозначным отображением $\mathbb{V} \leftrightarrow \mathbb{R}^{2n}$.

Преобразование элементов векторного пространства $x = x^a e_a \in \mathbb{C}^n$ мы записываем в виде

$$x^a \mapsto x^b X_b^a, \quad X = (X_b^a) \in \mathfrak{gl}(n, \mathbb{C}),$$

т.е. вектор-строка (x^a) умножается справа на матрицу преобразований $X = (X_b^a)$. Такая запись вызвана принятыми ранее правилами: компоненты вектора мы нумеруем верхним индексом и придерживаемся правила записи индексов суммирования “с десяти до четырех”.

Базис алгебры Ли $\mathfrak{gl}(n, \mathbb{C})$ можно выбрать из квадратных матриц e_a^b , имеющих один ненулевой элемент, который равен единице и находится на a -той строке и в b -том столбце. Очевидно, что любую матрицу можно записать в виде

$$X = X_a^b e_b^a \in \mathfrak{gl}(n, \mathbb{C}), \quad X_a^b \in \mathbb{C}.$$

Выбранный базис удовлетворяет соотношениям коммутации

$$[e_a^b, e_c^d] = \delta_c^b e_a^d - \delta_a^d e_c^b, \quad (8.96)$$

что проверяется прямой проверкой.

Рассмотренная параметризация естественным образом отождествляет алгебру Ли $\mathfrak{gl}(n, \mathbb{C})$ с комплексным евклидовым пространством $\mathbb{C}^{n^2}$ и вещественным евклидовым пространством $\mathbb{R}^{2n^2}$. Мы предполагаем, что топология алгебры Ли $\mathfrak{gl}(n, \mathbb{C})$ индуцирована вложением $\mathfrak{gl}(n, \mathbb{C}) \hookrightarrow \mathbb{R}^{2n^2}$.

Базис e_a^b алгебры Ли $\mathfrak{gl}(n, \mathbb{C})$ можно также представить в виде дифференциальных операторов

$$e_a^b = -x^b \partial_a,$$

действующих на компоненты векторов из $\mathbb{V}$. Это представление определяет действие генераторов алгебры Ли $\mathfrak{gl}(n, \mathbb{C})$ на любую дифференцируемую функцию $f(x) \in C^1(\mathbb{C}^n)$ от декартовых координат точки $\mathbb{C}^n$.

Определение. Матрица X называется *невыврожденной* или *регулярной*, если для нее существует обратная матрица X^{-1} , т.е. выполнено равенство $XX^{-1} = X^{-1}X = \mathbb{1}$. $\square$

Для того, чтобы матрица X была регулярной необходимо и достаточно, чтобы ее определитель был отличен от нуля, $\det X \neq 0$. Если эндоморфизм X векторного пространства $\mathbb{C}^n$ отображает $\mathbb{C}^n$ на себя (сюръективен), а не на какое-нибудь подпространство меньшей размерности, то соответствующая матрица X регулярна, и существует обратный эндоморфизм X^{-1} . В этом случае эндоморфизм X является автоморфизмом.

Пусть $X = (X_a^b)$ – квадратная $n \times n$ матрица, тогда для определителя справедливо разложение, в частности, по первой строке

$$\det X = \sum_{\sigma} X_1^{\sigma(1)} \dots X_n^{\sigma(n)} \operatorname{sgn} \sigma,$$

где сумма берется по всем перестановкам $\sigma(1, \dots, n)$ и множитель $\operatorname{sgn} \sigma$ обозначает знак перестановки.

Перепишем соотношения коммутации (8.96) в виде

$$[e_A, e_B] = f_{AB}^C e_C,$$

где пару индексов мы для краткости обозначили одной буквой $e_A := e_a^b$, $A = 1, \dots, n^2$. Тогда

$$f_{AB}^C = f_a^b c^d e^f = \delta_{ace}^{fbd} - \delta_{ace}^{dfb}.$$

Простые вычисления приводят к следующей форме Киллинга–Картана

$$\eta_{AB} := -f_{AC}^D f_{BD}^C = \eta_a^b c^d = -2(n\delta_{ac}^{db} + \delta_{ac}^{bd}).$$

Форма Киллинга–Картана задает инвариантное скалярное произведение в алгебре Ли $\mathfrak{gl}(n, \mathbb{C})$:

$$(X, Y) = -2n \operatorname{tr}(XY) + 2 \operatorname{tr} X \operatorname{tr} Y. \quad (8.97)$$

Инвариантность в данном случае означает независимость результата скалярного произведения от преобразования подобия:

$$(X', Y') = (X, Y),$$

где

$$X' := SXS^{-1}, \quad Y' := SYS^{-1}, \quad \forall S \in \mathbb{GL}(n, \mathbb{C}).$$

Скалярное произведение (8.97) вырождено. Действительно, поскольку $\operatorname{tr} \mathbb{1} = n$, то скалярное произведение всех матриц, кратных единичной матрице $E = a\mathbb{1}$, $a \in \mathbb{C}$, равно нулю со всеми матрицами из алгебры Ли $\mathfrak{gl}(n, \mathbb{C})$,

$$(E, X) = a(\mathbb{1}, X) = -2an \operatorname{tr} X + 2an \operatorname{tr} X = 0, \quad \forall X \in \mathfrak{gl}(n, \mathbb{C}).$$

Отсюда следует, что форма Киллинга–Картана вырождена, $\det \eta_{\text{ав}} = 0$, и поэтому алгебра Ли $\mathfrak{gl}(n, \mathbb{C})$ не является полупростой. Это значит, что на групповом многообразии группы Ли $\operatorname{GL}(n, \mathbb{C})$ не существует двусторонне инвариантной метрики и, следовательно, возникают серьезные проблемы с построением инвариантов.

Поскольку единичная матрица коммутирует со всеми матрицами, то множество матриц, кратных единице, $\mathfrak{a} = \{E\}$ является центром и образует идеал в алгебре Ли $\mathfrak{gl}(n, \mathbb{C})$. Его вещественная размерность равна двум, $\dim_{\mathbb{R}} \mathfrak{a} = 2$.

Алгебра Ли $\mathfrak{gl}(n, \mathbb{C})$ содержит много подалгебр. Рассмотрим некоторые подалгебры, которые наиболее часто встречаются в приложениях. Нетрудно видеть, что множество всех квадратных вещественных $n \times n$ матриц также является алгеброй Ли, которую мы обозначим через $\mathfrak{gl}(n, \mathbb{R})$. Эта алгебра является подалгеброй в алгебре комплексных матриц: $\mathfrak{gl}(n, \mathbb{R}) \subset \mathfrak{gl}(n, \mathbb{C})$. Базис e_a^b в $\mathfrak{gl}(n, \mathbb{R})$ можно выбрать таким же, как и в случае $\mathfrak{gl}(n, \mathbb{C})$, только коэффициенты разложения теперь будут не комплексные, а вещественные числа. Таким образом, алгебра Ли $\mathfrak{gl}(n, \mathbb{R})$ естественным образом отождествляется с евклидовым пространством $\mathbb{R}^{n^2}$. Формы Киллинга–Картана для $\mathfrak{gl}(n, \mathbb{C})$ и $\mathfrak{gl}(n, \mathbb{R})$ совпадают и, следовательно, алгебра Ли $\mathfrak{gl}(n, \mathbb{R})$ также не является полупростой.

В разделе 8.16 будет показано, что вещественная алгебра Ли $\mathfrak{gl}(n, \mathbb{R}) = {}^r\mathfrak{gl}(n, \mathbb{C})$ является вещественной формой комплексной алгебры Ли $\mathfrak{gl}(n, \mathbb{C})$ и, наоборот, комплексная алгебра Ли $\mathfrak{gl}(n, \mathbb{C}) = {}^{\mathbb{C}}\mathfrak{gl}(n, \mathbb{R})$ является комплексификацией вещественной алгебры Ли $\mathfrak{gl}(n, \mathbb{R})$.

Максимальные полупростые подалгебры Ли в $\mathfrak{gl}(n, \mathbb{C})$ и $\mathfrak{gl}(n, \mathbb{R})$ – это алгебры комплексных и вещественных матриц с нулевым следом:

$$\mathfrak{sl}(n, \mathbb{C}) = \{X \in \mathfrak{gl}(n, \mathbb{C}) : \operatorname{tr} X = 0\}, \quad (8.98)$$

$$\mathfrak{sl}(n, \mathbb{R}) = \{X \in \mathfrak{gl}(n, \mathbb{R}) : \operatorname{tr} X = 0\}. \quad (8.99)$$

Они имеют следующие вещественные размерности:

$$\dim_{\mathbb{R}} \mathfrak{sl}(n, \mathbb{C}) = 2n^2 - 2, \quad \dim \mathfrak{sl}_{\mathbb{R}}(n, \mathbb{R}) = n^2 - 1.$$

Каждую матрицу можно однозначно представить в виде

$$X = \tilde{X} + \frac{1}{n} \operatorname{tr} X \mathbb{1}, \quad \operatorname{tr} \tilde{X} = 0,$$

выделив из нее след. Будем считать две матрицы X и Y эквивалентными, если равны их бесследовые части, $\tilde{X} = \tilde{Y}$. Такие матрицы связаны соотношением

$$X = Y + \frac{1}{n} (\operatorname{tr} X - \operatorname{tr} Y) \mathbb{1}.$$

Поэтому

$$\mathfrak{sl}(n, \mathbb{C}) \simeq \frac{\mathfrak{gl}(n, \mathbb{C})}{\mathfrak{a}},$$

где $\mathfrak{a}$ – идеал в алгебре Ли $\mathfrak{gl}(n, \mathbb{C})$, состоящий из матриц, кратных единичной.

Чтобы доказать полупростоту этих подалгебр, вычислим форму Киллинга–Картана. Выберем базис в алгебрах Ли $\mathfrak{sl}(n, \mathbb{C})$ и $\mathfrak{sl}(n, \mathbb{R})$

$$\tilde{e}_a^b := e_a^b - \delta_a^b e_n^n, \quad a, b = 1, \dots, n, \quad (\text{суммирования по } n \text{ нет}),$$

который удовлетворяет условию $\mathrm{tr} \tilde{e}_a^b = 0$. Это соотношение определяет только $n^2 - 1$ базисных векторов, поскольку $e_n^n = 0$. Коммутационные соотношения принимают вид

$$[\tilde{e}_a^b, \tilde{e}_c^d] = \delta_c^b \tilde{e}_a^d - \delta_a^d \tilde{e}_c^b - \delta_{ac}^{bn} \tilde{e}_n^d + \delta_{an}^{bd} \tilde{e}_c^n - \delta_{cn}^{db} \tilde{e}_a^n + \delta_{ca}^{dn} \tilde{e}_n^b.$$

Теперь нетрудно вычислить форму Киллинга–Картана

$$\eta_{ab} = \eta_a^b{}_c^d = -2n (\delta_{ac}^{db} + \delta_{ac}^{bd} - \delta_{acn}^{bnd} - \delta_{can}^{dnb}).$$

Отсюда следует, что, если хотя бы одна из пар индексов имеет вид $(a, b) = (n, n)$ или $(c, d) = (n, n)$, то

$$\eta_a^b{}_n^n = \eta_n^n{}_c^d = 0.$$

Форма Киллинга–Картана задает инвариантное скалярное произведение в алгебрах Ли:

$$(X, Y) = -2n \mathrm{tr}(XY), \quad X, Y \in \mathfrak{sl}(n, \mathbb{C}) \quad \text{или} \quad X, Y \in \mathfrak{sl}(n, \mathbb{R}).$$

Эта форма Киллинга–Картана невырождена, и, значит, алгебры матриц с нулевым следом полупросты.

Более того, алгебра Ли $\mathfrak{sl}(n, \mathbb{C})$ проста.

В настоящее время все простые комплексные алгебры Ли классифицированы. Согласно теореме Адо классификация алгебр Ли сводится к классификации матричных алгебр Ли. И эта задача решена. Существуют четыре классические бесконечные серии:

$$\mathfrak{a}_n \ (n \geq 1), \quad \mathfrak{b}_n \ (n \geq 2), \quad \mathfrak{c}_n \ (n \geq 3), \quad \mathfrak{d}_n \ (n \geq 4) \quad (8.100)$$

и пять исключительных алгебр

$$\mathfrak{g}_2, \quad \mathfrak{f}_4, \quad \mathfrak{e}_6, \quad \mathfrak{e}_7, \quad \mathfrak{e}_8. \quad (8.101)$$

Теорема 8.14.1. Любая простая комплексная алгебра Ли изоморфна одной из алгебр (8.100), (8.101). Алгебры (8.100) и (8.101) между собой попарно не изоморфны.

Доказательство. См., например, [61], глава 11. □

Нижний индекс в используемых обозначениях для простых алгебр Ли имеет особый смысл: он равен комплексной размерности максимальной коммутативной подалгебры.

Алгебры Ли комплексных матриц с нулевым следом представляют собой первую из классических серий (8.100),

$$\mathfrak{a}_n := \mathfrak{sl}(n+1, \mathbb{C}),$$

и, следовательно, просты.

Три оставшиеся серии классических комплексных простых алгебр Ли строятся с помощью квадратичных форм.

Пусть в комплексном векторном пространстве $\mathbb{V}$ задана симметричная билинейная форма

$$\mathbb{V} \times \mathbb{V} \ni \quad x, y \mapsto (x, y) \quad \in \mathbb{V}. \quad (8.102)$$

Рассмотрим линейные преобразования X пространства $\mathbb{V}$, которые сохраняют заданную квадратичную форму в следующем смысле:

$$(Xx, y) + (x, Xy) = 0, \quad \forall x, y \in \mathbb{V}. \quad (8.103)$$

Множество таких преобразований образуют алгебру Ли с коммутатором $[X, Y] := XY - YX$. Действительно, если для двух преобразований X и Y выполнено равенство (8.103), то оно будет выполнено и для их коммутатора:

$$([X, Y]x, y) = (XYx, y) - (YXx, y) = -(x, [X, Y]y).$$

Алгебры Ли преобразований векторного пространства $\mathbb{V}$, сохраняющих билинейную форму, определяются выбором этой формы.

Пусть комплексная размерность векторного пространства равна $\dim_{\mathbb{C}} \mathbb{V} = m$. Зададим в $\mathbb{V}$ невырожденную билинейную симметричную положительно определенную (в вещественном подпространстве) квадратичную форму

$$(x, y) := (y, x).$$

Тогда в пространстве $\mathbb{V}$ существует базис e_a , $a = 1, \dots, m$, такой, что квадратичная форма имеет вид

$$(x, y) = x^a y^b \delta_{ab},$$

где $x = x^a e_a$, $y = y^a e_a$ и $\delta_{ab} := \text{diag}(+ \dots +)$ – обычная евклидова метрика. Алгебры Ли, сохраняющие эту форму называются *ортгоналными* и обозначаются $\mathfrak{o}(m, \mathbb{C})$.

Вторая и четвертая из классических серий (8.100) определяются следующими равенствами:

$$\begin{aligned} \mathfrak{b}_n &:= \mathfrak{o}(2n + 1, \mathbb{C}), \\ \mathfrak{d}_n &:= \mathfrak{o}(2n, \mathbb{C}). \end{aligned}$$

Пусть в векторном пространстве $\mathbb{V}$ задана невырожденная антисимметричная билинейная квадратичная форма, $(x, y) = -(y, x)$. В этом случае размерность векторного пространства должна быть четной $\dim_{\mathbb{C}} \mathbb{V} = 2n$ для невырожденности. Тогда в пространстве $\mathbb{V}$ существует такой базис e_a , где $a = 1, \dots, 2n$, что квадратичная форма имеет вид

$$(x, y) = x^a y^b \varpi_{ab},$$

где

$$\varpi = (\varpi_{ab}) := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$$

– каноническая симплектическая форма. Комплексные алгебры Ли, сохраняющие эту квадратичную форму, называются *симплектическими* и обозначаются $\mathfrak{sp}(n, \mathbb{C})$. Они дают третью классическую серию

$$\mathfrak{c}_n := \mathfrak{sp}(n, \mathbb{C}).$$

Описание комплексных исключительных алгебр Ли (8.101) довольно сложно. Интересующийся читатель может найти его, например, в [61].

Теперь обсудим более элементарные свойства матриц.

Определение. Пусть $A \in \mathfrak{gl}(n, \mathbb{C})$ и $\lambda \in \mathbb{C}$. Алгебраическое уравнение

$$\det(A - \lambda \mathbb{1}) = 0 \quad (8.104)$$

n -того порядка относительно λ называется *характеристическим* или *вековым* уравнением. Решения этого уравнения называются *собственными числами* матрицы A . $\square$

Согласно основной теореме алгебры каждая матрица A имеет в точности n собственных чисел $\lambda_1, \lambda_2, \dots, \lambda_n$ с учетом их кратности. В общем случае собственные числа комплексны даже для вещественной матрицы $A \in \mathfrak{gl}(n, \mathbb{R})$.

Матрицы A и A' называются *подобными* или *сопряженными*, если существует невырожденная матрица B такая, что $A' = BAB^{-1}$. Из уравнения (8.104) следует, что собственные числа подобных матриц совпадают.

Опишем экспоненциальное отображение матриц.

Определение. Пусть $A \in \mathfrak{gl}(n, \mathbb{C})$ – произвольная матрица с ограниченными элементами. Тогда ряд

$$e^A = \exp A := \mathbb{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

равномерно сходится, если A остается в ограниченной области пространства $\mathfrak{gl}(n, \mathbb{C}) \approx \mathbb{R}^{2n^2}$, т.е. каждый элемент матрицы ограничен. Этот ряд называется *экспоненциалом* матрицы A . $\square$

Функция $\exp A$ определена и непрерывна на $\mathfrak{gl}(n, \mathbb{C})$ и отображает $\mathfrak{gl}(n, \mathbb{C})$ в себя.

Сформулируем некоторые свойства экспоненциала матрицы (см., например, [50], глава 1, §2).

Предложение 8.14.1. Пусть B – невырожденная $n \times n$ матрица. Тогда

$$e^{BAB^{-1}} = B e^A B^{-1}.$$

Предложение 8.14.2. Для произвольной матрицы $A \in \mathfrak{gl}(n, \mathbb{C})$ существует такая невырожденная матрица B , что матрица BAB^{-1} является верхнетреугольной. Тогда матрица $e^{BAB^{-1}}$ также является верхнетреугольной. При этом на диагонали матрицы BAB^{-1} стоят собственные числа $\lambda_1, \dots, \lambda_n$ матрицы A , а на диагонали матрицы $e^{BAB^{-1}}$ стоят собственные числа $e^{\lambda_1}, \dots, e^{\lambda_n}$ матрицы e^A .

Следствие. Для любой матрицы $A \in \mathfrak{gl}(n, \mathbb{C})$ справедливо равенство

$$\det(\exp A) = \exp(\operatorname{tr} A) \quad \Leftrightarrow \quad \det B = \exp(\operatorname{tr} \ln B), \quad (8.105)$$

где $B := e^A$. $\square$

Отсюда следует, в частности, что $\det(\exp A) \neq 0$. То есть экспоненциал произвольной матрицы является невырожденной (регулярной) матрицей и, следовательно, принадлежит группе общих преобразований $\exp A \in \mathbb{GL}(n, \mathbb{C})$ (см. следующий раздел).

Предложение 8.14.3. Если матрицы A, B коммутируют, то

$$e^{A+B} = e^A e^B.$$

Следствие. Экспоненциальное отображение $t \mapsto e^{tA}$, где $t \in \mathbb{R}$ и $A \in \mathfrak{gl}(n, \mathbb{C})$ – фиксированная матрица, есть гладкое гомоморфное отображение аддитивной группы вещественных чисел в группу $\mathrm{GL}(n, \mathbb{C})$. $\square$

Касательным вектором к отображению $t \mapsto e^{tA}$ в точке $t = 0$ является матрица A (достаточно почленно продифференцировать степенной ряд). Экспоненциальное отображение $\mathbb{R} \rightarrow \mathrm{GL}(n, \mathbb{C})$ является единственной однопараметрической подгруппой в $\mathrm{GL}(n, \mathbb{C})$ с касательным вектором A в нуле.

Приведем несколько очевидных формул для экспоненциального отображения матриц

$$\begin{aligned} \exp A^T &= (\exp A)^T, \\ \exp A^\dagger &= (\exp A)^\dagger, \\ \exp(-A) &= (\exp A)^{-1}, \end{aligned}$$

где индексы T и $\dagger$ обозначают транспонирование и эрмитово сопряжение матриц.

Экспоненциальное отображение матриц аналогично экспоненциальному отображению, которое генерируется векторными полями (см. раздел 8.9). Оно задает отображение алгебры Ли $\mathfrak{gl}(n, \mathbb{C})$ в группу Ли $\mathrm{GL}(n, \mathbb{C})$.

Предложение 8.14.4. В алгебре Ли $\mathfrak{gl}(n, \mathbb{C})$ существует окрестность $\mathcal{U}$ нулевой матрицы $0 \in \mathcal{U} \subset \mathfrak{gl}(n, \mathbb{C})$, которая удовлетворяет следующим условиям:

1) при экспоненциальном отображении $A \mapsto \exp A$, где $A \in \mathcal{U}$, окрестность $\mathcal{U}$ непрерывно отображается на некоторую окрестность единичной матрицы $\mathbb{1} \in \mathrm{GL}(n, \mathbb{C})$;

2) $|\mathrm{tr} A| < 2\pi$;

3) при $A \in \mathcal{U}$ справедливы включения $-A, A^T, A^\dagger \in \mathcal{U}$.

При построении экспоненциального отображения поле комплексных чисел можно заменить на поле вещественных чисел. При этом конструкция не изменится, если собственные числа матрицы с вещественными элементами также вещественны. (В общем случае это, конечно, не так.)

Экспоненциальное отображение задает отображение множества эндоморфизмов векторного пространства в множество его автоморфизмов

$$\exp : \quad \mathrm{end}(\mathbb{V}) \rightarrow \mathrm{aut}(\mathbb{V}),$$

где $\mathbb{V}$ – произвольное векторное пространство над полем вещественных или комплексных чисел.

Можно доказать [42], что экспоненциальное отображение для комплексной группы $\mathrm{GL}(n, \mathbb{C})$ является сюръективным. В то же время для группы общих линейных преобразований над полем вещественных чисел $\mathrm{GL}(n, \mathbb{R})$ это не так.

Пример 8.14.1. Матрицу

$$\begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}$$

нельзя представить в виде e^A ни для какой вещественной матрицы $A \in \mathfrak{gl}(2, \mathbb{R})$. $\square$

В алгебре Ли операцией умножения является коммутатор. Посмотрим, что ему соответствует в группе Ли.

Предложение 8.14.5. *В окрестности единицы группы справедливо равенство*

$$e^{tA} e^{tB} e^{-tA} e^{-tB} = [A, B]t^2 + O(t^3), \quad t \rightarrow 0.$$

Доказательство. Прямая проверка. □

Аналогичное утверждение справедливо не только для матричных, но и произвольных групп Ли, которое можно также проверить прямыми вычислениями.

8.14.2 Группа Ли $\mathrm{GL}(n, \mathbb{C})$

Определение. Группой *общих линейных преобразований* $\mathrm{GL}(n, \mathbb{C})$ n -мерного векторного пространства над полем комплексных чисел называется множество всех невырожденных квадратных $n \times n$ матриц с комплексными элементами и обычным правилом умножения. Множество всех невырожденных $n \times n$ матриц с вещественными элементами образует группу общих линейных преобразований n -мерного вещественного пространства $\mathrm{GL}(n, \mathbb{R})$. □

Единственное условие, которое накладывается на матрицы – это отличие от нуля их определителей, которое необходимо и достаточно для существования обратных матриц.

Умножение компонент векторов, которые мы записываем в виде строки, на матрицы сохраняет линейную структуру векторного пространства $\mathbb{V}$ и поэтому является автоморфизмом $\mathbb{V}$, т.к. нулевой вектор остается неподвижным.

Множество всех, в том числе вырожденных, комплексных $n \times n$ матриц, как многообразие, представляет собой евклидово пространство $\mathbb{R}^{2n^2}$, т.к. параметризуется n^2 комплексными числами, на которые не наложено никаких ограничений. Поскольку определитель матрицы является непрерывной функцией от матрицы, то множество матриц с нулевым определителем представляет собой замкнутое подмножество в $\mathbb{R}^{2n^2}$, поскольку 0 является замкнутым подмножеством в $\mathbb{R}$. Поэтому множество невырожденных матриц $\mathrm{GL}(n, \mathbb{C})$ как дополнение замкнутого подмножества образует в евклидовом пространстве $\mathbb{R}^{2n^2}$ открытое подмногообразие и, следовательно, имеет ту же размерность $2n^2$, что и евклидово пространство.

Нетрудно проверить, что групповая операция является гладкой, и, следовательно, группа $\mathrm{GL}(n, \mathbb{C})$ является группой Ли. Эта группа связна некомпактна, но локально компактна.

Аналогично, группа всех невырожденных вещественных $n \times n$ матриц $\mathrm{GL}(n, \mathbb{R})$, как многообразие, представляет собой открытое подмногообразие в евклидовом пространстве $\mathbb{R}^{n^2}$ с гладкой групповой операцией. В отличие от группы комплексных матриц $\mathrm{GL}(n, \mathbb{C})$, группа вещественных матриц не является связной. Она состоит из двух связных компонент: матриц с положительным и отрицательным определителем. Множество матриц с положительным определителем $\mathrm{GL}_+(n, \mathbb{R})$ само является группой Ли и представляет собой связную компоненту единицы в $\mathrm{GL}(n, \mathbb{R})$. Вторая компонента представляет собой ее смежный класс и состоит из матриц с отрицательным определителем.

Это различие легко понять. В комплексном случае матрицу A с вещественным положительным определителем, $\det A > 0$, можно непрерывно трансформировать в матрицу B с вещественным отрицательным определителем, $\det B < 0$, вдоль пути,

целиком лежащим в $\mathbb{GL}(n, \mathbb{C})$ и состоящим из матриц с комплексным определителем. Здесь прослеживается связь с вещественной прямой и комплексной плоскостью. Если из вещественной прямой удалить точку 0, то она распадается на два несвязных многообразия. В то же время, удаление начала координат из комплексной плоскости оставляет ее связной.

Очевидно что группа вещественных матриц $\mathbb{GL}(n, \mathbb{R})$ является подгруппой в $\mathbb{GL}(n, \mathbb{C})$. Поскольку размерность $\dim \mathbb{GL}(n, \mathbb{R}) < \dim \mathbb{GL}(n, \mathbb{C})$, то согласно теоремам 2.10.1 и 2.10.2 эта подгруппа замкнута.

Теперь опишем несколько других подгрупп в $\mathbb{GL}(n, \mathbb{C})$.

Произвольную невырожденную матрицу $A = (A_a^b)$, $a, b, \dots = 1, 2, \dots, n$, можно представить в виде

$$A = a^{1/n} B, \quad \text{где } a := \det A.$$

Тогда

$$\det B = 1.$$

Определение. Совокупность всех квадратных $n \times n$ матриц с комплексными элементами, определитель которых равен единице, является группой, которая называется группой *специальных линейных преобразований* n -мерного векторного пространства над полем комплексных чисел и обозначается $\mathbb{SL}(n, \mathbb{C})$. Если элементы матриц вещественны, то соответствующая группа обозначается $\mathbb{SL}(n, \mathbb{R})$. Она является подгруппой в $\mathbb{SL}(n, \mathbb{C})$. $\square$

Очевидно, что $\mathbb{SL}(n, \mathbb{C}) \subset \mathbb{GL}(n, \mathbb{C})$. Это максимальная полупростая подгруппа группы $\mathbb{GL}(n, \mathbb{C})$, и ее размерность равна $2n^2 - 1$. Можно доказать, что она является простой. Соответствующая алгебра Ли $\mathfrak{sl}(n, \mathbb{C})$ изоморфна алгебре комплексных матриц с нулевым следом и была рассмотрена в предыдущем разделе.

Группа $\mathbb{SL}(n, \mathbb{R})$ является подгруппой группы общих линейных преобразований $\mathbb{GL}(n, \mathbb{R})$ и является связной некомпактной, но локально компактной группой Ли размерности $n^2 - 1$. Эта группа является максимальной полупростой подгруппой группы $\mathbb{GL}(n, \mathbb{R})$ (и простой). Алгебра Ли $\mathfrak{sl}(n, \mathbb{R})$ изоморфна алгебре вещественных матриц с нулевым следом.

В некоторой окрестности единичной матрицы $1 \in \mathbb{SL}(n, \mathbb{K})$, где $\mathbb{K} = \mathbb{R}$ или $\mathbb{K} = \mathbb{C}$, элемент группы $S \in \mathbb{SL}(n, \mathbb{K})$ взаимно однозначно представляется экспоненциальным отображением элемента алгебры $A \in \mathfrak{sl}(n, \mathbb{K})$:

$$S = e^A, \quad \text{tr } A = 0.$$

У группы общих линейных преобразований $\mathbb{GL}(n, \mathbb{C})$ есть много других подгрупп.

Пример 8.14.2. Группа вещественных ортогональных $n \times n$ матриц $\mathbb{O}(n)$ состоит из матриц $S \in \mathbb{O}(n) \subset \mathbb{GL}(n, \mathbb{R})$, удовлетворяющих условию ортогональности

$$SS^T = S^T S = 1,$$

где S^T обозначает транспонированную матрицу. След от этого равенства приводит к ограничению на матричные элементы $S_i^j S_j^i = n$. Это значит, что групповое многообразие является замкнутым подмножеством сферы $\mathbb{S}^{n^2-1} \subset \mathbb{R}^{n^2}$. Отсюда, по теореме Гейне–Бореля–Лебега, следует компактность группы вращений. Группа $\mathbb{O}(n)$ является компактной группой Ли размерности $n(n-1)/2$ и состоит из двух компонент связности. Матрицы с единичным определителем образуют связную компоненту единицы

$\mathrm{SO}(n)$ (собственные вращения). Вторую компоненту связности образует ее смежный класс, состоящий из несобственных ортогональных матриц с определителем, равным -1 .

Группы вращений $\mathrm{SO}(n)$ при $n \geq 2$ являются группами симметрий сфер $\mathbb{S}^{n-1}$, вложенных в евклидово пространство $\mathbb{R}^n$. Можно доказать, что фундаментальные группы групп вращений $\mathrm{SO}(n)$ при всех $n \geq 3$ равны $\mathbb{Z}_2$. Следовательно, они не являются односвязными.

Алгебры Ли групп $\mathcal{O}(n)$ и $\mathrm{SO}(n)$ совпадают и изоморфны алгебре антисимметричных $n \times n$ матриц $\mathfrak{so}(n)$. Вблизи единицы $1 \in \mathrm{SO}(n)$ матрица вращений $S \in \mathrm{SO}(n)$ взаимно однозначно представляется в виде экспоненциала отображения некоторого элемента алгебры $A \in \mathfrak{so}(n)$:

$$S = e^A, \quad \text{где } A = -A^T. \quad \square$$

Пример 8.14.3. Группа двумерных вращений $\mathrm{SO}(2)$ как многообразие представляет собой окружность $\mathbb{S}^1$. Ее фундаментальная группа изоморфна группе целых чисел $\mathbb{Z}$. Полная группа вращений $\mathcal{O}(2)$ состоит из двух компонент связности. Как многообразие она представляет собой два экземпляра окружности $\mathcal{O}(2) \approx \mathbb{S}^1 \times \mathbb{Z}_2$. Фундаментальная группа окружности совпадает с группой целых чисел по сложению $\mathbb{Z}$. $\square$

Пример 8.14.4. Группа трехмерных вращений $\mathrm{SO}(3)$ является группой симметрии сферы $\mathbb{S}^2$, вложенной в трехмерное евклидово пространство $\mathbb{R}^3$. Каждое вращение можно параметризовать вектором $\vec{n}$, направленным вдоль оси вращения, длина которого равна углу поворота против часовой стрелки $|\vec{n}| \leq \pi$. Множество векторов поворота заполняет шар $\mathbb{B}_\pi^3 \subset \mathbb{R}^3$. При этом противоположные точки граничной сферы $\mathbb{S}_\pi^2$ радиуса π необходимо отождествить, т.к. поворот на угол π совпадает с поворотом на угол $-\pi$. Таким образом, как многообразие группа $\mathrm{SO}(3)$ представляет собой трехмерный шар с отождествленными противоположными точками граничной сферы $\mathbb{S}_\pi^2$. Фундаментальная группа группы $\mathrm{SO}(3)$ является группа $\mathbb{Z}_2$. Полная группа трехмерных вращений $\mathcal{O}(3)$ как многообразие состоит из двух диффеоморфных компонент связности. Фундаментальная группа каждой компоненты связности равна $\mathbb{Z}_2$. $\square$

Пример 8.14.5. Пусть в декартовых координатах евклидова пространства $\mathbb{R}^n$ задана метрика, представляющая собой диагональную матрицу, у которой на диагонали стоят p положительных и q отрицательных единиц

$$\eta_{ab} = \mathrm{diag}(\underbrace{+ \dots +}_p \underbrace{- \dots -}_q), \quad p + q = n. \quad (8.106)$$

Для определенности мы считаем, что все положительные единицы идут вначале. Сигнатура этой метрики равна p, q . Матрицы преобразований евклидова пространства $S_a{}^b$, сохраняющие метрику (8.106),

$$\eta_{ab} = S_a{}^c S_b{}^d \eta_{cd}, \quad (8.107)$$

образуют группу, которая называется группой вращений $\mathcal{O}(p, q)$. В частности, при $q = 0$ группа $\mathcal{O}(n, 0) = \mathcal{O}(n)$, а при $p = 1$ она становится группой Лоренца $\mathcal{O}(1, n-1)$. Из определения (8.107) следует, что $\det S = \pm 1$. То есть группа вращений $\mathcal{O}(p, q)$ состоит по крайней мере из двух компонент связности: матриц с положительным

и отрицательным определителем. Можно показать, что группа Лоренца, состоит из четырех компонент связности. При $q \neq 0$, n группа вращений $\mathbb{O}(p, q)$ является некомпактной.

Группа вращений $\mathbb{O}(p, q)$ содержит подгруппу, состоящую из матриц с единичным определителем

$$\mathbb{SO}(p, q) := \mathbb{O}(p, q) \cap \mathbb{SL}(n, \mathbb{R}) \subset \mathbb{O}(p, q).$$

Группа вращений $\mathbb{O}(p, q)$ является группой симметрии пространств постоянной кривизны: гиперboloидов $\mathbb{H}$ размерности $\dim \mathbb{H} = n - 1$, вложенных в псевдоевклидово пространство $\mathbb{R}^{p, q}$ уравнением

$$(x^1)^2 + \dots + (x^p)^2 - (x^{p+1})^2 - \dots - (x^n)^2 = \pm r^2, \quad r = \text{const.}$$

Это утверждение является следствием определения (8.107). В зависимости от знака правой части мы получаем различные гиперboloиды, которые в общем случае состоят из нескольких компонент связности.

При рассмотрении алгебр Клиффорда (глава 26) важную роль играет комплексная группа вращений $\mathbb{O}(p, q, \mathbb{C})$. Она определяется тем же равенством (8.107), но для комплексных матриц S . $\square$

Пример 8.14.6. Рассмотрим комплексное векторное пространство $\mathbb{C}^n$ с базисом e_a , $a = 1, \dots, n$. Пусть $X = X^a e_a$ и $Y = Y^a e_a$ — два произвольных вектора из $\mathbb{C}^n$. Определим эрмитово скалярное произведение в $\mathbb{C}^n$ следующим соотношением

$$(X, Y) := X^a Y^{\dagger b} \delta_{ab}, \quad (8.108)$$

где символ $\dagger$ обозначает комплексное сопряжение. Это скалярное произведение линейно по первому аргументу и антилинейно по второму. Она называется *полуторалинейным* скалярным произведением. Из определения (8.108) следует, что $(X, X) \geq 0$ и $(X, X) = 0$ тогда и только тогда, когда $X = 0$. Вещественное число

$$\|X\| = \sqrt{(X, X)}$$

называется *длиной* вектора $X \in \mathbb{C}^n$.

Группа унитарных матриц $\mathbb{U}(n)$ состоит из комплексных $n \times n$ матриц $U \in \mathbb{U}(n)$, которые сохраняют эрмитово скалярное произведение (8.108). В матричных обозначениях это условие записывается в виде *условия унитарности*

$$UU^\dagger = U^\dagger U = 1,$$

где символ $\dagger$ обозначает эрмитово сопряжение (транспонирование и комплексное сопряжение всех элементов). Из условия унитарности следует, что определитель унитарной матрицы равен по модулю единице. Группа $\mathbb{U}(n)$ является связной группой Ли размерности $\dim \mathbb{U}(n) = n^2$. Так же как и группа вещественных ортогональных матриц эта группа компактна. Группа $\mathbb{U}(n)$ не является односвязной, и ее фундаментальная группа изоморфна группе целых чисел $\mathbb{Z}$ [67]. Как многообразие группа $\mathbb{U}(n)$ диффеоморфна прямому произведению окружности на группу специальных унитарных матриц с единичным определителем $\mathbb{U}(n) \approx \mathbb{S}^1 \times \mathbb{SU}(n)$, $n \geq 2$.

При $n = 1$ имеем абелеву группу $\mathbb{U}(1)$, которая состоит из комплексных чисел, равных по модулю единице, с обычным правилом умножения комплексных чисел.

Как многообразие эта группа представляет собой окружность, $\mathbb{U}(1) \approx \mathbb{S}^1$. Легко построить изоморфизм групп, $\mathbb{U}(1) \simeq \mathbb{SO}(2)$:

$$\mathbb{U}(1) \ni e^{i\alpha} \leftrightarrow \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \in \mathbb{SO}(2), \quad \alpha \in [0, 2\pi).$$

Алгебра Ли $\mathfrak{u}(n)$ изоморфна алгебре антиэрмитовых $n \times n$ матриц. Вблизи единицы $\mathbb{1} \in \mathbb{U}(n)$ произвольная унитарная матрица $U \in \mathbb{U}(n)$ взаимно однозначно представляется в виде экспоненциала от некоторой антиэрмитовой матрицы $A \in \mathfrak{u}(n)$

$$U = e^A, \quad A = -A^\dagger. \quad \square$$

Пример 8.14.7. Группа специальных унитарных матриц

$$\mathbb{SU}(n) := \mathbb{U}(n) \cap \mathbb{SL}(n, \mathbb{C}) \subset \mathbb{U}(n), \quad (8.109)$$

состоит из унитарных $n \times n$ матриц $U \in \mathbb{SU}(n)$ с единичным определителем:

$$UU^\dagger = U^\dagger U = 1, \quad \det U = 1.$$

Эта группа является связной компактной группой Ли размерности $\dim \mathbb{SU}(n) = n^2 - 1$. Она образует замкнутое подмножество в $\mathbb{U}(n)$. Группа $\mathbb{SU}(n)$ является также односвязной. Доказательство этого утверждения сложно, и интересующийся читатель может найти его в [67].

Группа $\mathbb{SU}(1)$ тривиальна и состоит из единственного элемента – единицы.

Алгебра Ли $\mathfrak{su}(n)$ изоморфна алгебре антиэрмитовых $n \times n$ матриц с нулевым следом. Вблизи единицы $\mathbb{1} \in \mathbb{SU}(n)$ каждая унитарная матрица с единичным определителем $U \in \mathbb{SU}(n)$ взаимно однозначно представляется экспоненциальным отображением элемента алгебры $A \in \mathfrak{su}(n)$:

$$U = e^A, \quad A = -A^\dagger, \quad \mathrm{tr} A = 0. \quad \square$$

Пример 8.14.8. Заменяем в определении эрмитова скалярного произведения (8.108) евклидову метрику δ_{ab} на метрику $\eta_{ab} := \mathrm{diag}(\underbrace{+\dots+}_p, \underbrace{-\dots-}_q)$, т.е.

$$(X, Y) := X^a Y^{\dagger b} \eta_{ab}.$$

Нетрудно проверить, что матрицы, сохраняющие эту полуторалинейную форму, образуют группу. Условие сохранения метрики имеет вид $U_a^c U_b^{*d} \eta_{cd} = \eta_{ab}$ или в матричных обозначениях

$$U \eta U^\dagger = \eta \quad \Leftrightarrow \quad \eta U^\dagger \eta = U^{-1}, \quad (8.110)$$

поскольку $\eta^2 = \mathbb{1}$. Группа комплексных матриц, удовлетворяющих условию (8.110), называется *псевдоунитарной* и обозначается $\mathbb{U}(p, q)$. Группа $\mathbb{U}(p, q)$ при всех p и q имеет подгруппу специальных псевдоунитарных матриц $\mathbb{SU}(p, q)$ с единичным определителем.

Алгебра Ли $\mathfrak{u}(p, q)$ состоит из матриц A , удовлетворяющих условию

$$A \eta = -\eta A^\dagger.$$

Эта система уравнений фиксирует $n + n(n-1) = n^2$ из $2n^2$ параметров. Следовательно,

$$\dim \mathbb{U}(p, q) = (p + q)^2.$$

Для алгебры Ли специальных псевдоунитарных матриц $\mathfrak{su}(p, q)$ на матрицы A добавляется еще одно условие $\mathrm{tr} A = 0$. Поэтому $\dim \mathbb{SU}(p, q) = (p + q)^2 - 1$. $\square$

Множество эрмитовых матриц играет важную роль и в математике, и в физике. Опишем их свойства подробнее. Напомним

Определение. Матрица $A \in \mathfrak{gl}(n, \mathbb{C})$ называется *эрмитовой*, если выполнено условие

$$A = A^\dagger. \quad \square \quad (8.111)$$

В вещественном случае множество эрмитовых матриц соответствует симметричным матрицам, а множество антиэрмитовых – антисимметричным.

Множество эрмитовых матриц $A, B, \dots$ в отличие от множества антиэрмитовых матриц не образует подалгебры Ли в $\mathfrak{gl}(n, \mathbb{C})$, т.к.

$$[A, B]^\dagger = [B, A] = -[A, B].$$

Заметим также, что отображение $A \mapsto A^\dagger$ не является автоморфизмом группы $\mathbb{GL}(n, \mathbb{C})$, т.к. при эрмитовом сопряжении меняется порядок матриц, и что эрмитовы матрицы не образуют в $\mathbb{GL}(n, \mathbb{C})$ подгруппы.

Сформулируем некоторые свойства эрмитовых матриц [50].

Предложение 8.14.6. Если A – эрмитова матрица, то UAU^{-1} – также эрмитова для любой унитарной матрицы $U \in \mathbb{U}(n)$. Далее, существует такая унитарная матрица U_0 (не единственная), что $U_0AU_0^{-1}$ является диагональной матрицей. Если матрица A вещественна, то матрицу U_0 можно выбрать ортогональной.

Предложение 8.14.7. Все собственные числа эрмитовой матрицы вещественны.

Предложение 8.14.8. Если A – эрмитова матрица, то в пространстве $\mathbb{C}^n$ существует ортонормальный базис, составленный из ее собственных векторов.

Эрмитова матрица называется *положительно определенной*, если все ее собственные числа положительны.

Предложение 8.14.9 (Полярное разложение матриц). Любая невырожденная матрица $A \in \mathbb{GL}(n, \mathbb{C})$ может быть записана, и притом лишь единственным способом, в виде произведения $A = UP$ или $A = PU$ унитарной матрицы $U \in \mathbb{U}(n)$ и положительно определенной матрицы P . Сомножители U и P являются непрерывными функциями A .

Поскольку множество всех положительно определенных матриц гомеоморфно $\mathbb{R}^{n^2}$, то группа $\mathbb{GL}(n, \mathbb{C})$, как многообразие, гомеоморфна прямому произведению $\mathbb{GL}(n, \mathbb{C}) \approx \mathbb{U}(n) \times \mathbb{R}^{n^2}$.

8.15 Универсальная накрывающая $\widetilde{\mathbb{SL}}(2, \mathbb{R})$

Построим универсальную накрывающую группу $\widetilde{\mathbb{SL}}(2, \mathbb{R})$ для группы $\mathbb{SL}(2, \mathbb{R})$. Это построение будет конструктивным, т.е. мы определим групповую операцию на связном и односвязном многообразии $\widetilde{\mathbb{SL}}(2, \mathbb{R})$, а затем построим гомоморфизм групп $\widetilde{\mathbb{SL}}(2, \mathbb{R}) \rightarrow \mathbb{SL}(2, \mathbb{R})$ (см., например, [67]).

Рассмотрим трехмерное многообразие $\widetilde{\mathbb{SL}}(2, \mathbb{R}) = \mathbb{R} \times \mathbb{D}$, которое представляет собой прямое произведение вещественной прямой $\mathbb{R}$ на единичный диск $\mathbb{D}$ плоскости комплексного переменного, состоящего из таких комплексных чисел z , что $|z| < 1$.

Поскольку каждый из сомножителей является связным и односвязным многообразием, то и их произведение также связно и односвязно.

Функция

$$w := \frac{1+z}{1+\bar{z}}$$

осуществляет непрерывное отображение диска $\mathbb{D}$ на окружность $|w| = 1$ с выколотой точкой $w = -1$. Поэтому для любого числа $z \in \mathbb{D}$ существует единственное число t , $-\pi < t < \pi$, для которого

$$e^{it} = \frac{1+z}{1+\bar{z}}.$$

Обозначим это число символом

$$t = \frac{1}{i} \ln \frac{1+z}{1+\bar{z}}.$$

Определим на многообразии $\widetilde{\mathbb{SL}}(2, \mathbb{R})$ умножение следующей формулой

$$(x, u)(y, v) = \left(x + y + t, \frac{u + e^{2iy}v}{e^{2iy} + u\bar{v}} \right), \quad x, y \in \mathbb{R}, \quad u, v \in \mathbb{D}, \quad (8.112)$$

где

$$t = \frac{1}{2i} \ln \frac{1 + e^{-2iy}u\bar{v}}{1 + e^{2iy}\bar{u}v}.$$

Так как

$$|e^{2iy} + u\bar{v}|^2 - |u + e^{2iy}v|^2 = (1 - |u|^2)(1 - |v|^2) > 0,$$

то

$$\left| \frac{u + e^{2iy}v}{e^{2iy} + u\bar{v}} \right| < 1.$$

Поэтому формула (8.112) корректно определяет в $\widetilde{\mathbb{SL}}(2, \mathbb{R}) = \mathbb{R} \times \mathbb{D}$ умножение.

Это умножение обладает единицей $(0, 0)$, и для любого элемента (x, u) существует обратный элемент

$$(x, u)^{-1} = (-x, -e^{2ix}u).$$

Кроме того, прямые вычисления показывают, что умножение (8.112) ассоциативно. Это умножение гладко, и тем самым доказано, что многообразие $\widetilde{\mathbb{SL}}(2, \mathbb{R}) = \mathbb{R} \times \mathbb{D}$ является группой Ли относительно умножения (8.112).

Центр этой группы состоит из таких элементов (x, u) , для которых, в частности, равенство

$$\frac{u + e^{2iy}v}{e^{2iy} + u\bar{v}} = \frac{v + e^{2ix}u}{e^{2ix} + v\bar{u}} \quad (8.113)$$

выполнено для всех $y \in \mathbb{R}$ и $v \in \mathbb{D}$. В частном случае при $v = 0$ должно выполняться равенство $u = e^{2iy}u$ для всех y . Это возможно только при $u = 0$. Тогда из (8.113) следует равенство $v = e^{2ix}v$, которое должно выполняться для всех v , что возможно только при $x = \pi n$, $n \in \mathbb{Z}$. Нетрудно проверить, что все элементы вида $(\pi n, 0)$ принадлежат центру группы $\widetilde{\mathbb{SL}}(2, \mathbb{R})$. Следовательно, центр группы $\widetilde{\mathbb{SL}}(2, \mathbb{R})$ бесконечен.

Теперь построим гомоморфизм $\widetilde{\mathbb{SL}}(2, \mathbb{R}) \rightarrow \mathbb{SL}(2, \mathbb{R})$. Легко проверить, что матрица

$$A = \frac{1}{\sqrt{1 - |u|^2}} \begin{pmatrix} \cos x + |u| \cos(x + \varphi) & |u| \sin(x + \varphi) - \sin x \\ |u| \sin(x + \varphi) + \sin x & \cos x - |u| \cos(x + \varphi) \end{pmatrix}, \quad (8.114)$$

где $\varphi = \arg u$, имеет единичный определитель. Прямые вычисления позволяют проверить, что отображение $\widetilde{\mathrm{SL}}(2, \mathbb{R}) \rightarrow \mathrm{SL}(2, \mathbb{R})$ определено корректно и является гомоморфизмом групп. Его ядро состоит из элементов вида $(2\pi n, 0)$ и, значит, дискретно. Поскольку размерности групп совпадают, то гомоморфизм $\widetilde{\mathrm{SL}}(2, \mathbb{R}) \rightarrow \mathrm{SL}(2, \mathbb{R})$ является групповым накрытием. Таким образом, мы построили универсальную накрывающую для группы $\mathrm{SL}(2, \mathbb{R})$:

$$\mathrm{SL}(2, \mathbb{R}) = \frac{\widetilde{\mathrm{SL}}(2, \mathbb{R})}{\mathbb{Z}},$$

где группа преобразований

$$\mathbb{Z} : \quad x \rightarrow x + 2\pi n$$

действует на прямой $\mathbb{R}$. С топологической точки зрения группа $\mathrm{SL}(2, \mathbb{R})$ является произведением окружности на диск

$$\mathrm{SL}(2, \mathbb{R}) \approx \mathbb{S}^1 \times \mathbb{D}.$$

Группа специальных матриц $\mathrm{SL}(2, \mathbb{R})$ дважды накрывает собственную ортохронную группу Лоренца $\mathrm{SO}_0(1, 2)$ (1.188). При этом одному лоренцеву вращению соответствует пара матриц (8.114), отличающихся знаком. Изменению знака матрицы A соответствует сдвиг координаты x на π . Поэтому

$$\mathrm{SO}_0(1, 2) = \frac{\mathrm{SL}(2, \mathbb{R})}{\mathbb{Z}_2} = \frac{\widetilde{\mathrm{SL}}(2, \mathbb{R})}{\mathbb{Z}},$$

где группа преобразований действует по следующему правилу:

$$\mathbb{Z} : \quad x \rightarrow x + \pi n.$$

Группа $\mathbb{Z}_2$ действует на окружности $\mathbb{S}^1$ путем отождествления противоположных точек и превращает окружность в проективное пространство $\mathbb{RP}^1$. Поскольку одномерное проективное пространство с топологической точки зрения является окружностью, то

$$\mathrm{SO}_0(1, 2) \approx \mathbb{S}^1 \times \mathbb{D}.$$

Поскольку диск диффеоморфен всей плоскости $\mathbb{D} \approx \mathbb{R}^2$, то

$$\mathrm{SO}_0(1, 2) \approx \mathbb{S}^1 \times \mathbb{R}^2.$$

Предложенная конструкция универсальной накрывающей имеет следующее объяснение. Согласно теореме о полярном разложении матриц любая унимодулярная матрица $A \in \mathrm{SL}(2, \mathbb{R})$ единственным образом раскладывается в произведение $A = UP$ некоторой матрицы вращения U и некоторой положительно определенной симметричной унимодулярной матрицы P . Матрица U задается углом поворота x , а матрица P , имеющая вид

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

где $a > 0$ и $ac - b^2 = 1$, задается двумя числами $a > 0$ и $b \in \mathbb{R}$, т.е. комплексным числом $z = a + ib$, принадлежащим правой полуплоскости $a > 0$. Переходя с помощью дробно-линейного преобразования от правой полуплоскости к единичному кругу, мы получаем, следовательно, что каждая матрица $A \in \mathrm{SL}(2, \mathbb{R})$ однозначно характеризуется парой (x, u) , где $x \in \mathbb{R}$ и $u \in \mathbb{D}$. При этом умножению матриц соответствует

умножение пар, которое задается формулой (8.112). Чтобы получить теперь универсальную накрывающую $\widetilde{\mathrm{SL}}(2, \mathbb{R})$ достаточно считать координату x не углом, а точкой вещественной прямой $\mathbb{R}$.

Теорема Адо, которая будет сформулирована в следующем разделе, утверждает, что любая алгебра Ли имеет точное матричное представление. То есть классификация матричных алгебр Ли влечет за собой классификацию всех алгебр Ли. С группами дело обстоит иначе.

Теорема 8.15.1. *Группа Ли $\widetilde{\mathrm{SL}}(2, \mathbb{R})$ не вкладывается ни в какую группу $\mathrm{GL}(n, \mathbb{R})$ и поэтому не является матричной группой Ли.*

Доказательство. См., например, [13], теорема 6.1. □

8.16 Классификация простых алгебр и групп Ли

Классификация простых алгебр Ли сводится к классификации матричных алгебр Ли, благодаря следующему утверждению.

Теорема 8.16.1 (Адо). *Каждая алгебра Ли $\mathfrak{g}$ над полем комплексных чисел $\mathbb{C}$ имеет точное (инъективное) представление в алгебре Ли $\mathfrak{gl}(n, \mathbb{R})$ для некоторого n .*

Доказательство. См. [69]. □

Следствие. Для каждой алгебры Ли $\mathfrak{g}$ существует группа Ли $\mathbb{G}$ (и, в частности, односвязная группа Ли) с данной алгеброй Ли. □

Теорема Адо важна, поскольку позволяет свести изучение абстрактных алгебр Ли к матричным.

Напомним, что комплексификацией вещественного векторного пространства $\mathbb{V}$ называется комплексное векторное пространство $\mathbb{V}^{\mathbb{C}}$, состоящее из векторов вида $Z = X + iY$, где $X, Y \in \mathbb{V}$ (см., раздел 1.3.3). При этом сложение векторов $Z_1 = X_1 + iY_1$ и $Z_2 = X_2 + iY_2$ задается формулой

$$Z_1 + Z_2 := X_1 + X_2 + i(Y_1 + Y_2).$$

Умножение вектора Z на комплексное число $c = a + ib$ определяется равенством

$$cZ := aX - bY + i(aY + bX).$$

Определение. Комплексификацией $\mathfrak{g}^{\mathbb{C}}$ вещественной алгебры Ли $\mathfrak{g}$ называется комплексная алгебра Ли, удовлетворяющая двум условиям:

- 1) $\mathfrak{g}^{\mathbb{C}}$ является комплексификацией вещественного векторного пространства $\mathfrak{g}$;
- 2) умножение в $\mathfrak{g}^{\mathbb{C}}$ задается формулой

$$[Z_1, Z_2] = [X_1 + iY_1, X_2 + iY_2] := [X_1, X_2] - [Y_1, Y_2] + i[X_1, Y_2] + i[Y_1, X_2]. \quad \square \quad (8.115)$$

Если $\dim_{\mathbb{R}} \mathfrak{g} = N$, то вещественная размерность ее комплексификации вдвое больше: $\dim_{\mathbb{R}} (\mathfrak{g}^{\mathbb{C}}) = 2N$. При этом комплексная размерность комплексифицированной алгебры Ли остается прежней $\dim_{\mathbb{C}} (\mathfrak{g}^{\mathbb{C}}) = \dim \mathfrak{g} = N$.

С другой стороны, комплексную алгебру Ли $\mathfrak{h}$ комплексной размерности N с базисом e_A , $A = 1, \dots, N$, можно рассматривать как вещественную алгебру Ли вещественной размерности $2N$ с базисом $e_1, ie_1, \dots, e_N, ie_N$. Это овеществление комплексной алгебры Ли обозначим $\mathfrak{h}^{\mathbb{R}}$. Можно поступить по-другому. Назовем вещественной формой комплексной алгебры Ли $\mathfrak{h}$ такую вещественную алгебру Ли $\mathfrak{h}^r$, которая после комплексификации дает $\mathfrak{h}$. В этом случае размерность вещественной формы комплексной алгебры Ли в два раза меньше вещественной размерности исходной алгебры Ли.

Пример 8.16.1. Рассмотрим вещественную алгебру Ли $\mathfrak{gl}(n, \mathbb{R})$, состоящую из вещественных $n \times n$ матриц с обычным правилом коммутирования. Тогда ее комплексификацией будет алгебра Ли $\mathfrak{gl}^{\mathbb{C}}(n, \mathbb{R}) = \mathfrak{gl}(n, \mathbb{C})$, состоящая из комплексных $n \times n$ матриц. Теперь рассмотрим комплексную алгебру Ли $\mathfrak{gl}(n, \mathbb{C})$. Ее овеществление будет изоморфно прямой сумме алгебр Ли: $\mathfrak{gl}^{\mathbb{R}}(n, \mathbb{C}) \simeq \mathfrak{gl}(n, \mathbb{R}) \oplus \mathfrak{gl}(n, \mathbb{R})$ той же вещественной размерности. В то же время вещественной формой алгебры Ли $\mathfrak{gl}(n, \mathbb{C})$ будет вещественная алгебра Ли $\mathfrak{gl}^r(n, \mathbb{C}) = \mathfrak{gl}(n, \mathbb{R})$ вдвое меньшей размерности. $\square$

В предыдущем разделе была приведена классификация всех простых комплексных алгебр Ли. Поэтому возникает вопрос следует ли отсюда классификация вещественных алгебр Ли. Ответ на этот вопрос положителен в силу следующего утверждения.

Теорема 8.16.2. *Все вещественные простые алгебры Ли получаются из простых комплексных алгебр Ли либо овеществлением, либо они являются вещественными формами простых комплексных алгебр Ли.*

Доказательство. См., например, [32], глава 1, §5, теорема 1. $\square$

Таким образом, для классификации простых вещественных алгебр Ли необходимо найти все вещественные неизоморфные между собой формы комплексных алгебр Ли. Этот вопрос сложен, но решен. Доказывается, что каждая комплексная простая алгебра имеет только конечное число вещественных форм, и все они найдены [70]. Среди этих вещественных форм только одна является компактной. Таким образом, существует взаимно однозначное соответствие между простыми комплексными алгебрами Ли и простыми вещественными компактными алгебрами Ли. Поэтому классификация простых вещественных компактных алгебр Ли сводится к классификации простых комплексных алгебр Ли.

Приведем классификацию всех простых вещественных алгебр Ли (доказательство можно найти, например, в [32]).

Вещественные формы алгебры Ли $\mathfrak{a}_n \simeq \mathfrak{sl}(n+1, \mathbb{C})$, $n \geq 1$

1. $\mathfrak{su}(n+1)$ – компактная алгебра Ли всех антиэрмитовых $(n+1) \times (n+1)$ матриц Z с нулевым следом, $\mathrm{tr} Z = 0$.
2. $\mathfrak{sl}(n+1, \mathbb{R})$ – некомпактная алгебра Ли всех вещественных $(n+1) \times (n+1)$ матриц X с нулевым следом, $\mathrm{tr} X = 0$.
3. $\mathfrak{su}(p, q)$, $p+q = n+1$, $p \geq q$, – некомпактная алгебра Ли всех матриц вида

$$\begin{pmatrix} Z_1 & Z_2 \\ Z_2^\dagger & Z_3 \end{pmatrix},$$

где Z_1, Z_3 – антиэрмитовы матрицы порядков p и q , соответственно, $\text{tr } Z_1 + \text{tr } Z_3 = 0$, матрица Z_2 произвольна.

4. $\mathfrak{su}^*(2(n+1))$ – некомпактная алгебра Ли всех $2(n+1) \times 2(n+1)$ матриц вида

$$\begin{pmatrix} Z_1 & Z_2 \\ -\bar{Z}_2 & \bar{Z}_1 \end{pmatrix},$$

где Z_1, Z_2 – комплексные $(n+1) \times (n+1)$ матрицы, $\text{tr } Z_1 + \text{tr } \bar{Z}_1 = 0$.

Вещественные формы алгебры Ли $\mathfrak{b}_n \simeq \mathfrak{so}(2n+1, \mathbb{C})$, $n \geq 2$

1. $\mathfrak{so}(2n+1)$ – компактная алгебра Ли всех вещественных антисимметричных $(2n+1) \times (2n+1)$ матриц.
2. $\mathfrak{so}(p, q)$, $p+q = 2n+1$, $p \geq q$, – некомпактная алгебра Ли всех вещественных $(2n+1) \times (2n+1)$ матриц вида

$$\begin{pmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{pmatrix},$$

где X_1, X_3 – антисимметричные матрицы порядков p и q , соответственно, матрица X_2 произвольна.

Вещественные формы алгебры Ли $\mathfrak{c}_n \simeq \mathfrak{sp}(n, \mathbb{C})$, $n \geq 3$

1. $\mathfrak{sp}(n)$ – компактная алгебра Ли всех антиэрмитовых бесследовых комплексных $2n \times 2n$ матриц вида

$$\begin{pmatrix} Z_1 & Z_2 \\ Z_3 & -Z_1^T \end{pmatrix},$$

где все матрицы $Z_{1,2,3}$ имеют порядок n , и матрицы Z_2 и Z_3 симметричные (т.е. $\mathfrak{sp}(n) = \mathfrak{sp}(n, \mathbb{C}) \cap \mathfrak{su}(2n)$).

2. $\mathfrak{sp}(n, \mathbb{R})$ – некомпактная алгебра Ли всех вещественных $2n \times 2n$ матриц вида

$$\begin{pmatrix} X_1 & X_2 \\ X_3 & -X_1^T \end{pmatrix},$$

где $X_{1,2,3}$ – вещественные $n \times n$ матрицы и $X_{2,3}$ – симметричные матрицы.

3. $\mathfrak{sp}(p, q)$, $p+q = n$, $p \geq q$, – некомпактная алгебра Ли всех комплексных $2n \times 2n$ матриц вида

$$\begin{pmatrix} Z_{11} & Z_{12} & Z_{13} & Z_{14} \\ Z_{12}^\dagger & Z_{22} & Z_{14}^T & Z_{24} \\ -\bar{Z}_{13} & \bar{Z}_{14} & \bar{Z}_{11} & -\bar{Z}_{12} \\ Z_{14}^\dagger & -\bar{Z}_{24} & -Z_{12}^T & \bar{Z}_{22} \end{pmatrix},$$

где Z_{11}, Z_{13} – матрицы порядка p , Z_{12} и Z_{14} – $p \times q$ матрицы, Z_{11} и Z_{22} – антиэрмитовы, Z_{13} и Z_{24} – симметричные.

Вещественные формы алгебры Ли $\mathfrak{d}_n \simeq \mathfrak{so}(2n, \mathbb{C})$, $n \geq 4$

1. $\mathfrak{so}(2n)$ – компактная алгебра Ли всех вещественных антисимметричных $2n \times 2n$ матриц.
2. $\mathfrak{so}(p, q)$, $p+q = 2n$, $p \geq q$, – некомпактная алгебра Ли всех вещественных $2n \times 2n$ матриц вида

$$\begin{pmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{pmatrix},$$

где X_1, X_3 – антисимметричные матрицы порядков p и q , соответственно, матрица X_2 произвольна.

3. $\mathfrak{so}^*(2n)$ – некомпактная алгебра Ли всех комплексных $2n \times 2n$ матриц вида

$$\begin{pmatrix} Z_1 & Z_2 \\ -\bar{Z}_2 & \bar{Z}_1 \end{pmatrix},$$

где Z_1, Z_2 – комплексные $n \times n$ матрицы, Z_1 – антисимметричная, Z_2 – эрмитова.

Вещественные формы алгебр Ли, приведенные выше, определены при всех $n \geq 1$. Ограничения на номера алгебр n возникают из-за наличия изоморфизмов между комплексными алгебрами Ли, которые индуцируют изоморфизмы их вещественных форм. Запишем эти изоморфизмы в виде таблицы. Случай $\mathfrak{d}_2 \simeq \mathfrak{a}_1 \oplus \mathfrak{a}_1$ включен для удобства. Отметим также, что алгебра $\mathfrak{d}_1$ не является полупростой.

Изоморфизмы комплексных алгебр Ли	Изоморфизмы вещественных форм алгебр Ли
$\mathfrak{a}_1 \simeq \mathfrak{b}_1 \simeq \mathfrak{c}_1$	$\mathfrak{su}(2) \simeq \mathfrak{so}(3) \simeq \mathfrak{sp}(1)$ $\mathfrak{sl}(2, \mathbb{R}) \simeq \mathfrak{su}(1, 1) \simeq \mathfrak{so}(2, 1) \simeq \mathfrak{sp}(1, \mathbb{R})$
$\mathfrak{b}_2 \simeq \mathfrak{c}_2$	$\mathfrak{so}(5) \simeq \mathfrak{sp}(2)$ $\mathfrak{so}(3, 2) \simeq \mathfrak{sp}(2, \mathbb{R})$ $\mathfrak{so}(4, 1) \simeq \mathfrak{sp}(1, 1)$
$\mathfrak{d}_2 \simeq \mathfrak{a}_1 \oplus \mathfrak{a}_1$	$\mathfrak{so}(4) \simeq \mathfrak{so}(3) \oplus \mathfrak{so}(3)$ $\mathfrak{so}(2, 2) \simeq \mathfrak{sl}(2, \mathbb{R}) \oplus \mathfrak{sl}(2, \mathbb{R})$ $\mathfrak{sl}(2, \mathbb{C}) \simeq \mathfrak{so}(3, 1)$ $\mathfrak{so}^*(4) \simeq \mathfrak{sl}(2, \mathbb{R}) \oplus \mathfrak{su}(2)$
$\mathfrak{a}_3 \simeq \mathfrak{d}_3$	$\mathfrak{su}(4) \simeq \mathfrak{so}(6)$ $\mathfrak{sl}(4, \mathbb{R}) \simeq \mathfrak{so}(3, 3)$ $\mathfrak{su}(2, 2) \simeq \mathfrak{so}(4, 2)$ $\mathfrak{su}(3, 1) \simeq \mathfrak{so}^*(6)$ $\mathfrak{su}^*(4) \simeq \mathfrak{so}(5, 1)$ $\mathfrak{so}^*(8) \simeq \mathfrak{so}(6, 2)$

Таблица 8.1: Изоморфизм комплексных и вещественных алгебр Ли

Вещественные формы исключительных простых алгебр Ли приведены, например, в [71].

Как уже отмечалось, каждой комплексной простой алгебре Ли соответствует только одна вещественная компактная форма.

Приведем также классификацию комплексных простых групп и их вещественных компактных форм (см., например, [66], глава X, §6). В таблице колонка простых

комплексных групп Ли обозначена $\mathbb{G}^{\mathbb{C}}$, а их вещественных компактных форм – $\mathbb{G}$. Универсальные накрывающие вещественных компактных групп Ли отмечены знаком тильды $\tilde{\mathbb{G}}$, а через $\mathbb{Z}(\tilde{\mathbb{G}})$ обозначен центр универсальной накрывающей. То есть $\mathbb{G} = \tilde{\mathbb{G}}/\mathbb{Z}(\tilde{\mathbb{G}})$. Приведены также размерности вещественных групп Ли.

$\mathfrak{g}$	$\mathbb{G}^{\mathbb{C}}$	$\mathbb{G}$	$\mathbb{Z}(\tilde{\mathbb{U}})$	$\dim \mathbb{U}$
$\mathfrak{a}_n \quad (n \geq 1)$	$\mathrm{SL}(n+1, \mathbb{C})$	$\mathrm{SU}(n+1)$	$\mathbb{Z}_{n+1}$	$n(n+2)$
$\mathfrak{b}_n \quad (n \geq 2)$	$\mathrm{SO}(2n+1, \mathbb{C})$	$\mathrm{SO}(2n+1)$	$\mathbb{Z}_2$	$n(2n+1)$
$\mathfrak{c}_n \quad (n \geq 3)$	$\mathrm{SP}(n, \mathbb{C})$	$\mathrm{SP}(n)$	$\mathbb{Z}_2$	$n(2n+1)$
$\mathfrak{d}_n \quad (n \geq 4)$	$\mathrm{SO}(2n, \mathbb{C})$	$\mathrm{SO}(2n)$	$\mathbb{Z}_4, \quad n \text{ нечетно}$ $\mathbb{Z}_2 \oplus \mathbb{Z}_2, \quad n \text{ четно}$	$n(2n-1)$
$\mathfrak{g}_2$	$\mathbb{G}_2^{\mathbb{C}}$	$\mathbb{G}_2$	$\mathbb{Z}_1$	14
$\mathfrak{f}_4$	$\mathbb{F}_4^{\mathbb{C}}$	$\mathbb{F}_4$	$\mathbb{Z}_1$	52
$\mathfrak{e}_6$	$\mathbb{E}_6^{\mathbb{C}}$	$\mathbb{E}_6$	$\mathbb{Z}_3$	78
$\mathfrak{e}_7$	$\mathbb{E}_7^{\mathbb{C}}$	$\mathbb{E}_7$	$\mathbb{Z}_2$	133
$\mathfrak{e}_8$	$\mathbb{E}_8^{\mathbb{C}}$	$\mathbb{E}_8$	$\mathbb{Z}_1$	248

Таблица 8.2: Простые группы Ли, соответствующие простым комплексным алгебрам Ли и их компактным вещественным формам

Глава 9

Группы преобразований

Теория групп имеет важные приложения в математической физике, и почти всегда группы появляются как группы преобразований чего-либо. Это может быть группа преобразований пространства-времени, группа симметрий изотопического пространства (например, $SU(3)$ симметрия в физике элементарных частиц), группа симметрий уравнений движения или что-нибудь более экзотическое. В настоящем разделе мы рассмотрим группы преобразований $\mathbb{G}$ многообразия $\mathbb{M}$. В общем случае группа $\mathbb{G}$ может быть конечной, счетной или группой Ли. В основном мы будем рассматривать группы Ли преобразований, хотя многие определения и утверждения справедливы и для произвольных групп.

9.1 Действие групп преобразований

Пусть на многообразии $\mathbb{M}$ задано *правое* действие группы $\mathbb{G}$, т.е. задано отображение

$$\mathbb{M} \times \mathbb{G} \ni x, a \mapsto xa \in \mathbb{M}, \quad (9.1)$$

которое предполагается достаточно гладким по x .

Определение. Пара $(\mathbb{M}, \mathbb{G})$ называется *группой правых преобразований* многообразия $\mathbb{M}$, если действие группы $\mathbb{G}$ на $\mathbb{M}$ справа удовлетворяет следующим условиям:

- 1) при любом $a \in \mathbb{G}$ отображение $\mathbb{M} \ni x \mapsto xa \in \mathbb{M}$ является диффеоморфизмом,
- 2) $(xa)b = x(ab)$ для всех $x \in \mathbb{M}$ и $a, b \in \mathbb{G}$.

В этом случае многообразие $\mathbb{M}$ называется $\mathbb{G}$ -многообразием. □

Если $\mathbb{G}$ – группа Ли, то отображение (9.1) предполагается достаточно гладким также по $a \in \mathbb{G}$ и, следовательно, непрерывным для всех $a \in \mathbb{G}$ и для всех $x \in \mathbb{M}$. Для каждой фиксированной точки $x \in \mathbb{M}$ мы имеем достаточно гладкое отображение $\mathbb{G} \rightarrow \mathbb{M}$ (орбита точки x).

Замечание. Иногда в качестве $\mathbb{M}$ мы будем рассматривать многообразия с краем. Под дифференцируемой функцией f на $\mathbb{M}$ в таком случае мы понимаем наличие всех частных производных у f во всех внутренних точках и существование предела этих производных на крае $\partial\mathbb{M}$. □

Каждому элементу группы $a \in \mathbb{G}$ соответствует достаточно гладкое отображение $\mathbb{M} \rightarrow \mathbb{M}$, заданное правилом $x \mapsto xa$. Поэтому группа преобразований определяет отображение $\mathbb{G} \rightarrow \text{diff}(\mathbb{M})$ группы $\mathbb{G}$ в группу диффеоморфизмов многообразия $\text{diff}(\mathbb{M})$.

Предложение 9.1.1. Пара $(\mathbb{M}, \mathbb{G})$ является группой преобразований многообразия $\mathbb{M}$ тогда и только тогда, когда отображение $\mathbb{G} \rightarrow \text{diff}(\mathbb{M})$ является гомоморфизмом групп и отображение (9.1) является достаточно гладким.

Доказательство. Условие 1) в определении группы преобразований означает, что отображение $x \mapsto xa$ принадлежит группе $\text{diff}(\mathbb{M})$. Групповая операция сохраняется при отображении в силу свойства 2). Осталось доказать, что единица $e \in \mathbb{G}$ отображается в тождественное преобразование многообразия $\mathbb{M}$. Действительно, отображение $x \mapsto xa$ является диффеоморфизмом и, следовательно, биекцией для всех $a \in \mathbb{G}$. В частности, для каждого $x \in \mathbb{M}$ найдется такая точка $y \in \mathbb{M}$, что $x = ye$. Умножив это уравнение справа на e , получаем цепочку равенств:

$$xe = ye^2 = ye = x, \quad \forall x \in \mathbb{M}.$$

То есть единичный элемент группы не сдвигает ни одну точку многообразия.

Обратно. Каждому элементу a соответствует некоторый диффеоморфизм, т.е. условие 1) выполнено. Поскольку отображение $\mathbb{G} \rightarrow \text{diff}(\mathbb{M})$ – гомоморфизм, то выполнено условие 2). $\square$

Аналогично определяется *левое* действие группы $\mathbb{G}$ на $\mathbb{M}$, т.е. отображение

$$\mathbb{G} \times \mathbb{M} \ni a, x \mapsto ax \in \mathbb{M}.$$

В дальнейшем, если не оговорено противное, под *группой преобразований* многообразия будем понимать группу правых преобразований и будем говорить, что $\mathbb{G}$ действует на $\mathbb{M}$, подразумевая действие группы справа.

Пример 9.1.1. Если задано представление группы $\mathbb{G}$, то задана группа линейных преобразований $(\mathbb{V}, \mathbb{G})$ векторного пространства $\mathbb{V}$ (автоморфизмов), которое снабжено естественной структурой многообразия. $\square$

Пример 9.1.2. Группа Ли является группой преобразований самой себя $\mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}$, которую можно рассматривать, как действующую слева или справа. $\square$

Определение. Группа $\mathbb{G}$ действует на $\mathbb{M}$ *тривиально*, если равенство $xa = x$ выполнено для всех $x \in \mathbb{M}$ и всех $a \in \mathbb{G}$. В этом случае группа преобразований превращается в тривиальное главное расслоение $\mathbb{M} \times \mathbb{G} \xrightarrow{\pi} \mathbb{M}$ с естественной проекцией $\pi : (x, a) \mapsto x$.

Говорят, что $\mathbb{G}$ действует *эффективно* на $\mathbb{M}$, если из равенства $xa = x$ для всех $x \in \mathbb{M}$ следует, что $a = e$. Если для любой точки $x \in \mathbb{M}$ уравнение $xa = x$ имеет единственное решение $a = e$, то говорят, что группа преобразований действует *свободно*. $\square$

Другими словами, действие группы $\mathbb{G}$ на $\mathbb{M}$ является эффективным, если любой элемент группы, отличный от единичного, перемещает хотя бы одну точку. Действие группы свободно, если любой элемент группы, отличный от единичного, перемещает все точки многообразия. Группа, действующая свободно, является эффективной. Обратное утверждение неверно, что показывает следующий

Пример 9.1.3. Группа вращений $\text{SO}(3)$ трехмерного евклидова пространства $\mathbb{R}^3$ действует эффективно, но не свободно, т.к. начало координат остается неподвижным. $\square$

Если действие группы преобразований свободно, то это значит, что для любой пары точек $x_1, x_2 \in \mathbb{M}$ либо не существует преобразования, связывающего эти точки, $x_2 = x_1 a$, либо элемент $a \in \mathbb{G}$ *единственен*.

Определение. Ядром $\mathbb{K}$ группы преобразований $(\mathbb{M}, \mathbb{G})$ называется множество элементов группы, которое действует тривиально на все точки многообразия,

$$\mathbb{K} := \{a \in \mathbb{G} : xa = x, \forall x \in \mathbb{M}\}. \quad \square$$

Действие группы эффективно, если $\mathbb{K} = \{e\}$.

Предложение 9.1.2. Ядро группы преобразований является нормальной подгруппой в $\mathbb{G}$. Кроме того, эта подгруппа замкнута в $\mathbb{G}$.

Доказательство. Отображение $\mathbb{G} \rightarrow \text{diff}(\mathbb{M})$ непрерывно и ядро $\mathbb{K}$ отображается в единицу группы диффеоморфизмов $\text{diff}(\mathbb{M})$. Поскольку единица замкнута в $\text{diff}(\mathbb{M})$, то ее прообраз $\mathbb{K}$ замкнут в $\mathbb{G}$ по критерию 1) теоремы 1.4.3. $\square$

Ядро группы преобразований представляет собой то множество элементов группы, которое вообще не действует на многообразии $\mathbb{M}$. Поэтому действие группы $\mathbb{G}$ естественным образом сводится к действию фактор группы $\mathbb{G}/\mathbb{K}$, которая действует эффективно.

Пример 9.1.4. Рассмотрим присоединенное действие группы Ли на себя $\mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}$, заданное преобразованием подобия $a \mapsto bab^{-1}$, для всех $a, b \in \mathbb{G}$. Ядром этого действия является центр группы $\mathbb{K} \subset \mathbb{G}$. Присоединенное действие группы Ли не является свободным, т.к. единичный элемент группы $e \in \mathbb{G}$ остается неподвижным, $aea^{-1} = e$, $\forall a \in \mathbb{G}$. Присоединенное действие фактор группы $\mathbb{G}/\mathbb{K}$ является эффективным. $\square$

Определение. Пусть точка $x \in \mathbb{M}$ фиксирована. Множество точек

$$x\mathbb{G} := \{xa \in \mathbb{M} : \forall a \in \mathbb{G}\}$$

называется *орбитой* точки x относительно действия группы $\mathbb{G}$. $\square$

Используя групповые свойства, нетрудно проверить, что орбиты двух точек $x_1, x_2 \in \mathbb{M}$ либо совпадают, либо не имеют общих точек. Орбитой произвольного подмножества $\mathbb{U} \subset \mathbb{M}$ будем называть объединение орбит всех точек из $\mathbb{U}$. Обозначим его $\mathbb{U}\mathbb{G}$.

Если группа $\mathbb{G}$ конечна, то орбита любой точки также конечна. Если группа преобразований счетна, то ее орбиты либо счетны, либо конечны.

Пример 9.1.5. Рассмотрим поворот евклидовой плоскости $\mathbb{R}^2$ на конечный угол θ . Начало координат неподвижно относительно вращений. Все остальные точки имеют нетривиальные орбиты, если угол поворота не кратен 2π . Пусть $\theta = 2\pi a$, где $a \in \mathbb{R}$. Если число a рационально, $a = m/n$, где m и n – взаимно простые натуральные числа, то орбиты состоят из n точек. Если a – иррациональное число, то орбита произвольной точки $x \neq 0$ счетна и являются всюду плотным подмножеством на окружности $\mathbb{S}^1 \hookrightarrow \mathbb{R}^2$, проходящей через точку x . $\square$

Если задана группа Ли преобразований $(\mathbb{M}, \mathbb{G})$, то орбита $x\mathbb{G}$ произвольной точки $x \in \mathbb{M}$ является подмногообразием в $\mathbb{M}$. Это следует из дифференцируемости отображения (9.1) для всех $x \in \mathbb{M}$ и групповых свойств $\mathbb{G}$. Размерность орбиты не превосходит размерности группы, $\dim(x\mathbb{G}) \leq \dim \mathbb{G}$.

Пример 9.1.6. Рассмотрим евклидово пространство $\mathbb{R}^n$ с декартовой системой координат x^α , $\alpha = 1, \dots, n$. Пусть группа вещественных чисел $\mathbb{R}$ по сложению действует на $\mathbb{R}^n$, как трансляции вдоль первой оси,

$$\mathbb{R}^n \times \mathbb{R} \ni (x^1, x^2, \dots, x^n), a \mapsto (x^1 + a, x^2, \dots, x^n) \in \mathbb{R}^n.$$

Тогда орбитой произвольной точки $x \in \mathbb{R}^n$ является прямая, проходящая через x и параллельная оси x^1 . $\square$

Предложение 9.1.3. Орбита $x\mathbb{G}$ произвольной точки $x \in \mathbb{M}$ является замкнутым подмногообразием в $\mathbb{M}$. Кроме того, если $\mathbb{G}$ – связная группа Ли, то орбита является связным подмногообразием.

Доказательство. Подмножество $\mathbb{M} \setminus x$ открыто в $\mathbb{M}$. Поэтому объединение открытых подмножеств $(\mathbb{M} \setminus x)\mathbb{G}$ открыто в $\mathbb{M}$. Следовательно, орбита произвольной точки замкнута, как дополнение открытого подмножества, $x\mathbb{G} = \mathbb{M} \setminus ((\mathbb{M} \setminus x)\mathbb{G})$. Поскольку отображение (9.1) является непрерывным для всех x , то связность сохраняется. $\square$

Определение. Обозначим через $\mathbb{M}/\mathbb{G}$ множество орбит $\check{x} := x\mathbb{G}$ относительно действия группы $\mathbb{G}$ на $\mathbb{M}$. То есть $\check{x}_1 = \check{x}_2$ тогда и только тогда, когда точки x_1 и x_2 лежат на одной орбите: $x_2 = x_1 a$ для некоторого $a \in \mathbb{G}$. Пусть $\pi : \mathbb{M} \rightarrow \mathbb{M}/\mathbb{G}$ – проекция, сопоставляющая каждой точке $x \in \mathbb{M}$ ее орбиту $\check{x}$. Тогда пространство орбит наделяется фактортопологией, т.е. множество $\mathbb{V} \in \mathbb{M}/\mathbb{G}$ является открытым тогда и только тогда, когда множество $\pi^{-1}(\mathbb{V})$ открыто в $\mathbb{M}$. Если $\mathbb{U} \subset \mathbb{M}$ – окрестность точки $x \in \mathbb{M}$, то множество орбит $\mathbb{V} = \mathbb{U}\mathbb{G}$ – окрестность точки $\check{x} \in \mathbb{M}/\mathbb{G}$. Множество орбит $\mathbb{M}/\mathbb{G}$ с фактортопологией называется *пространством орбит* многообразия $\mathbb{M}$ относительно действия группы $\mathbb{G}$ или *факторпространством*. $\square$

Хотя $\mathbb{M}$ и $\mathbb{G}$ – многообразия, пространство орбит с фактортопологией может оказаться весьма сложным. Оно, как покажут дальнейшие примеры, может не иметь структуры многообразия. В общем случае пространство орбит может не быть даже хаусдорфовым топологическим пространством.

Определение. Действие группы $\mathbb{G}$ на многообразии $\mathbb{M}$ называется *транзитивным*, если для любых двух точек $x_1, x_2 \in \mathbb{M}$ найдется такой элемент $a \in \mathbb{G}$, что $x_1 a = x_2$. $\square$

Если действие группы $\mathbb{G}$ на многообразии $\mathbb{M}$ транзитивно, то орбита произвольной точки $x \in \mathbb{M}$ совпадает со всем $\mathbb{M}$. Если на многообразии существует одна точка $x_0 \in \mathbb{M}$, которую можно преобразовать во все другие точки $\mathbb{M}$, то этого достаточно для того, чтобы действие группы $\mathbb{G}$ на $\mathbb{M}$ было транзитивным. Действительно, если $x_1 = x_0 a$ и $x_2 = x_0 b$, то $x_2 = x_1 a^{-1} b$.

Пример 9.1.7. Действие группы Ли $\mathbb{G}$ на себе является транзитивным и свободным. $\square$

Пример 9.1.8. Пусть $\mathbb{G}$ – группа и $\mathbb{H}$ – ее нетривиальная подгруппа. Рассмотрим пространство правых смежных классов $\mathbb{G}/\mathbb{H}$, состоящее из элементов вида $\mathbb{H}a$, где $a \in \mathbb{G}$. Определим действие группы $\mathbb{G}$ в пространстве смежных классов $\mathbb{G}/\mathbb{H}$ соотношением

$$\mathbb{G}/\mathbb{H} \times \mathbb{G} \ni \mathbb{H}a, b \mapsto \mathbb{H}ab \in \mathbb{G}/\mathbb{H}.$$

Это действие транзитивно для любых $\mathbb{G}$ и $\mathbb{H}$. Действие $\mathbb{G}$ на $\mathbb{G}/\mathbb{H}$ эффективно тогда и только тогда, когда $\mathbb{H}$ не содержит нормальных подгрупп. Действительно, если $\mathbb{H}$ – нормальная подгруппа, то для всех $h \in \mathbb{H}$ и $a \in \mathbb{G}$ выполнены равенства:

$$(\mathbb{H}a)h = a\mathbb{H}h = a\mathbb{H} = \mathbb{H}a.$$

Действие группы никогда не является свободным, т.к. $(\mathbb{H}e)h = \mathbb{H}e$ для всех $h \in \mathbb{H}$. $\square$

В теории групп преобразований важную роль играет

Теорема 9.1.1. Пусть $\mathbb{G}$ – группа Ли и $\mathbb{H}$ – ее замкнутая подгруппа. Тогда на фактор пространстве правых смежных классов $\mathbb{G}/\mathbb{H}$ существует единственная структура вещественно аналитического многообразия такая, что действие $\mathbb{G}$ на $\mathbb{G}/\mathbb{H}$ вещественно аналитично. В частности, проекция $\mathbb{G} \rightarrow \mathbb{G}/\mathbb{H}$ вещественно аналитична.

Доказательство. См., например, в [66], глава II, теорема 4.2. $\square$

Определение. Точка $x \in \mathbb{M}$ называется *стационарной* или *неподвижной* относительно действия группы преобразований, если $xa = x$ для всех $a \in \mathbb{G}$. $\square$

Если группа действует на многообразии свободно или транзитивно, то стационарные точки отсутствуют.

Пример 9.1.9. Представление группы $(\mathbb{V}, \mathbb{G})$ не может быть транзитивной или свободной группой преобразований, т.к. нулевой вектор из $\mathbb{V}$ является стационарной точкой. $\square$

Определение. Группой изотропии $\mathbb{G}_x$ точки многообразия $x \in \mathbb{M}$ называется множество элементов группы $\mathbb{G}$, оставляющих точку x неподвижной,

$$\mathbb{G}_x := \{a \in \mathbb{G} : xa = x\}.$$

Нетрудно проверить, что это множество элементов образует группу. Группы изотропии $\mathbb{G}_{x_1}$ и $\mathbb{G}_{x_2}$ двух точек, лежащих на одной орбите, $x_2a = x_1$, являются сопряженными подгруппами в $\mathbb{G}$:

$$\mathbb{G}_{x_2} = a\mathbb{G}_{x_1}a^{-1}.$$

Ядро $\mathbb{K}$ группы преобразований $(\mathbb{M}, \mathbb{G})$ связано с группами изотропии простым соотношением

$$\mathbb{K} = \bigcap_{x \in \mathbb{M}} \mathbb{G}_x,$$

где объединение берется по всем точкам $x \in \mathbb{M}$.

Из теоремы 9.1.1 следует, что пара $(\mathbb{G}/\mathbb{H}, \mathbb{G})$ является группой транзитивных преобразований. Представляет большой интерес обратное утверждение: если пара $(\mathbb{M}, \mathbb{G})$ – транзитивная группа преобразований, то она имеет вид $(\mathbb{G}/\mathbb{H}, \mathbb{G})$ для некоторой подгруппы $\mathbb{H} \subset \mathbb{G}$ и при некоторых дополнительных предположениях, которые мы уточним ниже. Для ясности сначала будет доказана теорема безотносительно топологии и дифференцируемой структуры, заданной на $\mathbb{M}$.

Теорема 9.1.2. Пусть $\mathbb{G}$ – группа, которая действует транзитивно на некотором множестве $\mathbb{M}$. Тогда для каждой точки $x \in \mathbb{M}$ существует биекция правых смежных классов $\mathbb{G}/\mathbb{G}_x$ на множество $\mathbb{M}$, заданное отображением

$$\mathbb{G}/\mathbb{G}_x \ni \mathbb{G}_x a \mapsto xa \in \mathbb{M}, \quad (9.2)$$

где $\mathbb{G}_x$ – группа изотропии точки x .

Доказательство. Отображение (9.2) определено для каждого представителя из правого смежного класса. Покажем сначала, что это отображение не зависит от выбора представителя. Пусть a и b – два представителя какого-либо одного смежного класса. Тогда существует элемент $h \in \mathbb{G}_x$ такой, что $a = hb$ и поэтому $xa = xhb = xb$. То есть отображение (9.2) не зависит от выбора представителя смежного класса и, следовательно, определено корректно.

Выберем по представителю из двух смежных классов $\mathbb{G}_x a$ и $\mathbb{G}_x b$ и подействуем на точку x . В результате получим точки $x_1 = xa$ и $x_2 = xb$. Эти точки совпадают тогда и только тогда, когда группа изотропии содержит элемент $ba^{-1} \in \mathbb{G}_x$. Это означает, что отображение (9.2) является взаимно однозначным.

Поскольку группа преобразований $(\mathbb{M}, \mathbb{G})$ является транзитивной, то отображение (9.2) является взаимно однозначным отображением на, т.е. биекцией. $\square$

Следствие. Если группа преобразований $(\mathbb{M}, \mathbb{G})$ транзитивна и свободна, т.е. $\mathbb{G}_x = e$ для всех точек $x \in \mathbb{M}$, то отображение

$$\mathbb{G} \ni a \mapsto xa \in \mathbb{M}$$

является биекцией. $\square$

Теперь разберемся с дифференцируемой структурой.

Предложение 9.1.4. Пусть $(\mathbb{M}, \mathbb{G})$ – группа транзитивных преобразований. Выберем произвольную точку $x \in \mathbb{M}$. Тогда ее группа изотропии $\mathbb{G}_x$ является замкнутой подгруппой в $\mathbb{G}$.

Доказательство. Поскольку группа преобразований транзитивна, то отображение $\mathbb{G} \ni a \mapsto xa \in \mathbb{M}$, определенное для каждого $x \in \mathbb{M}$, является сюръективным и непрерывным. Точка x является замкнутым подмножеством в $\mathbb{M}$, и, в силу теоремы критерия 1) из теоремы 1.4.3, группа изотропии $\mathbb{G}_x$ замкнута в $\mathbb{G}$. $\square$

Следствие. Группа изотропии $\mathbb{G}_x$ для всех $x \in \mathbb{M}$ является замкнутой подгруппой в $\mathbb{G}$ для произвольной группы Ли преобразований $(\mathbb{M}, \mathbb{G})$.

Доказательство. Каждая точка $x \in \mathbb{M}$ принадлежит орбите $x\mathbb{G}$, которая является подмногообразием в $\mathbb{M}$. На каждой орбите группа $\mathbb{G}$ действует транзитивно, и поэтому пара $(x\mathbb{G}, \mathbb{G})$ – транзитивная группа преобразований. Следовательно, группа изотропии $\mathbb{G}_x$ является замкнутой подгруппой в $\mathbb{G}$. $\square$

Согласно теореме 9.1.1 на пространстве смежных классов $\mathbb{G}/\mathbb{G}_x$ существует единственная дифференцируемая структура. Поскольку отображение (9.2) является биекцией, то эту дифференцируемую структуру можно перенести на $\mathbb{M}$, превратив его тем самым в многообразие. Тогда отображение (9.2) становится диффеоморфизмом.

Таким образом, если группа преобразований $(\mathbb{M}, \mathbb{G})$ транзитивна, то многообразие $\mathbb{M}$ представляет собой фактор пространство $\mathbb{G}/\mathbb{G}_x$. При этом

$$\dim \mathbb{M} = \dim \mathbb{G} - \dim \mathbb{G}_x.$$

Отсюда следует, что размерность многообразия $\mathbb{M}$, на котором действует транзитивная группа преобразований, не может превышать размерность группы.

Определение. Произвольное множество $\mathbb{M}$, на котором группа преобразований $(\mathbb{M}, \mathbb{G})$ действует транзитивно называется *однородным пространством*. Если, вдобавок, действие группы эффективно, то это множество называется *главным однородным пространством* над $\mathbb{G}$. $\square$

Поскольку каждое свободное действие группы преобразований является также эффективным, то группа преобразований $(\mathbb{M}, \mathbb{G})$, действующая транзитивно и свободно является главным однородным пространством. Многие авторы определяют главное однородное пространство именно в таком более узком смысле. В этом случае в силу следствия из теоремы 9.1.2 главное однородное пространство – это есть просто пара $(\mathbb{G}, \mathbb{G})$, где группа Ли $\mathbb{G}$ действует на себе либо правыми, либо левыми сдвигами.

Пример 9.1.10. Пара $(\mathbb{G}/\mathbb{H}, \mathbb{G})$ (пример 9.1.8) является однородным пространством. Если подгруппа $\mathbb{H} \subset \mathbb{G}$ не содержит нормальных подгрупп, то пара $(\mathbb{G}/\mathbb{H}, \mathbb{G})$ – главное однородное пространство. $\square$

Пример 9.1.11. Сфера единичного радиуса $\mathbb{S}^n$ задается в евклидовом пространстве $\mathbb{R}^{n+1}$ алгебраическим уравнением

$$(x^1)^2 + \dots + (x^{n+1})^2 = 1.$$

Линейные действия групп вращений $\mathbb{SO}(n+1)$ и $\mathbb{O}(n+1)$ на $\mathbb{R}^{n+1}$ порождают транзитивные действия этих групп на сфере $\mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1}$. Группа изотропии точки $(1, 0, \dots, 0) \in \mathbb{S}^n$ состоит из блочно диагональных матриц вида

$$\begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix}, \quad S \in \mathbb{SO}(n) \quad \text{или} \quad \mathbb{O}(n).$$

Поэтому из теоремы 9.1.2 следует, что сфера $\mathbb{S}^n$ является однородным пространством,

$$\mathbb{S}^n \approx \frac{\mathbb{SO}(n+1)}{\mathbb{SO}(n)} \approx \frac{\mathbb{O}(n+1)}{\mathbb{O}(n)}. \quad (9.3)$$

Конечно, подгруппа $\mathbb{SO}(n) \subset \mathbb{SO}(n+1)$ является группой изотропии для каждой точки $x \in \mathbb{S}^n$ в силу транзитивности действия группы преобразований.

Только две сферы $\mathbb{S}^1$ и $\mathbb{S}^3$ можно рассматривать как главные однородные пространства

$$\mathbb{S}^1 \approx \mathbb{U}(1) \simeq \mathbb{SO}(2), \quad \mathbb{S}^3 \approx \mathbb{SU}(2),$$

на которых группы преобразований действуют транзитивно и свободно. $\square$

Пример 9.1.12. Если сфера имеет нечетную размерность $2n+1$, то ее можно задать уравнением

$$|z^1|^2 + |z^2|^2 + \dots + |z^{n+1}|^2 = 1$$

в комплексном пространстве $\mathbb{C}^{n+1}$. На ней транзитивно действуют группы $\mathbb{U}(n+1)$ и $\mathbb{SU}(n+1)$. Аналогично вещественному случаю группы изотропии точки $(1, 0, \dots, 0) \in \mathbb{S}^{n+1}$ состоят из блочно диагональных матриц, которые изоморфны группам $\mathbb{U}(n)$ или $\mathbb{SU}(n)$. Поэтому

$$\mathbb{S}^{2n+1} \approx \frac{\mathbb{U}(n+1)}{\mathbb{U}(n)} \approx \frac{\mathbb{SU}(n+1)}{\mathbb{SU}(n)}. \quad \square \quad (9.4)$$

Пример 9.1.13. Вещественное проективное пространство $\mathbb{RP}^n$ состоит из всех прямых в евклидовом пространстве $\mathbb{R}^{n+1}$, проходящих через начало координат. На пространстве $\mathbb{RP}^n$ транзитивно действует группа вращений $\mathbb{O}(n+1)$: матрица вращений

$S \in \mathbb{O}(n+1)$ переводит прямую с направляющим вектором X в прямую с направляющим вектором SX . Группа изотропии прямой с направлением $X = (1, 0, \dots, 0)$ образована матрицами вида

$$\begin{pmatrix} \pm 1 & 0 \\ 0 & S \end{pmatrix}, \quad S \in \mathbb{O}(n).$$

Отсюда следует, что вещественное проективное пространство является однородным:

$$\mathbb{RP}^n \approx \frac{\mathbb{O}(n+1)}{\mathbb{O}(1) \times \mathbb{O}(n)}, \quad (9.5)$$

где $\mathbb{O}(1) \simeq \mathbb{Z}_2 = \{1, -1\}$. □

Пример 9.1.14. Комплексное проективное пространство $\mathbb{CP}^n$ представляет собой множество комплексных прямых (плоскостей), проходящих через начало координат. Эти прямые задаются комплексным направляющим вектором $X \neq 0$ и состоят из всех векторов вида aX , где $a \in \mathbb{C}$. Не ограничивая общности, можно считать, что направляющий вектор имеет единичную длину, $|X| = 1$. На пространстве единичных векторов в $\mathbb{C}^n$ транзитивно действует унитарная группа $\mathbb{U}(n+1)$. Группа изотропии комплексной прямой с направляющим вектором $X = (1, 0, \dots, 0)$ образована всеми матрицами вида

$$\begin{pmatrix} e^{i\varphi} & 0 \\ 0 & S \end{pmatrix}, \quad \varphi \in \mathbb{R}, \quad S \in \mathbb{U}(n).$$

Поэтому проективное пространство $\mathbb{CP}^n$ является однородным:

$$\mathbb{CP}^n \approx \frac{\mathbb{U}(n+1)}{\mathbb{U}(1) \times \mathbb{U}(n)}. \quad (9.6)$$

Когда мы нормируем направляющие векторы, считая их единичными, мы фактически рассматриваем единичную сферу $\mathbb{S}^{2n+1}$, вложенную в $\mathbb{C}^{n+1}$. Единичные векторы X_1 и X_2 задают одну и ту же точку в проективном пространстве $\mathbb{CP}^n$, если и только если они пропорциональны: $X_1 = e^{i\varphi} X_2$. Отсюда вытекает, что группа $\mathbb{U}(1) = \{e^{i\varphi}\}$ действует на сфере $\mathbb{S}^{2n+1}$ умножениями: $X \mapsto e^{i\varphi} X$. Это действие определяет группу изотропии точки из $\mathbb{CP}^n$, и его пространство орбит – гладкое многообразие

$$\mathbb{CP}^n \approx \frac{\mathbb{S}^{2n+1}}{\mathbb{U}(1)}. \quad (9.7)$$

Так как $\mathbb{CP}^1 = \mathbb{S}^2$ и $\mathbb{U}(1) \approx \mathbb{S}^1$, то возникает проекция

$$\mathbb{S}^3 \xrightarrow{\pi} \mathbb{S}^2 = \frac{\mathbb{S}^3}{\mathbb{U}(1)}$$

трехмерной сферы на двумерную сферу, слоем которой является окружность. Таким образом возникает расслоение Хопфа $\mathbb{S}^3(\mathbb{S}^2, \pi, \mathbb{U}(1))$ (пример 12.1.9). □

9.2 Инфинитезимальные преобразования

Пусть задана группа преобразований $(\mathbb{M}, \mathbb{G})$, где $\mathbb{G}$ – группа Ли. Тогда имеет смысл говорить об инфинитезимальных преобразованиях, которые взаимно однозначно определяются векторными полями на $\mathbb{M}$ и $\mathbb{G}$. Ниже мы рассмотрим соотношения между алгебрами Ли векторных полей на $\mathbb{G}$ и $\mathbb{M}$.

Построим отображение $\mu : \mathfrak{g} \rightarrow \mathcal{X}(\mathbb{M})$ алгебры Ли $\mathfrak{g}$ группы Ли $\mathbb{G}$ (множества левоинвариантных векторных полей в $\mathcal{X}(\mathbb{G})$) в бесконечномерную алгебру Ли векторных полей $\mathcal{X}(\mathbb{M})$ на многообразии $\mathbb{M}$. Это можно сделать тремя эквивалентными способами.

1) Каждому элементу алгебры Ли $X \in \mathfrak{g}$ соответствует однопараметрическая подгруппа (см. раздел 8.9)

$$\exp_X : \mathbb{R} \ni t \mapsto a(t) \in \mathbb{G},$$

где $a(t) := \exp_X(t)$ (экспоненциальное отображение) и $a(0) = e$ – единица группы. Тогда через каждую точку $x \in \mathbb{M}$ проходит единственная кривая $xa(t)$. Касательные векторы к этим кривым образуют векторное поле $X^*(x)$ на многообразии $\mathbb{M}$. Таким образом, мы построили отображение μ алгебры Ли $\mathfrak{g}$ в алгебру векторных полей $\mathcal{X}(\mathbb{M})$:

$$\mu : \mathfrak{g} \ni X \mapsto \mu(X) := X^* \in \mathcal{X}(\mathbb{M}).$$

2) Для каждой точки $x \in \mathbb{M}$ определим отображение $\mu_x : \mathbb{G} \ni a \mapsto xa \in \mathbb{M}$. Тогда отображение μ определено соотношением

$$X^*(x) = \mu(X)|_x := (\mu_x)_* X_e, \quad (9.8)$$

где $(\mu_x)_*$ – дифференциал отображения μ_x , действующий на касательный вектор $X_e := X(e)$ к единице группы, который соответствует элементу алгебры Ли $X \in \mathfrak{g}$.

3) Каждое векторное поле взаимно однозначно определяется дифференцированием в алгебре функций. Поэтому рассмотрим гладкую функцию $f \in \mathcal{C}^\infty(\mathbb{M})$ и определим векторное поле соотношением

$$X^*(f) := \left. \frac{d}{dt} f(xa(t)) \right|_{t=0},$$

где $a(t) := \exp_X(t)$.

Поскольку произвольное левоинвариантное векторное поле X на группе Ли $\mathbb{G}$ полно, то полно также векторное поле X^* на $\mathbb{M}$. Это следует из определения, т.к. отображение μ определено для всех $t \in (-\infty, \infty)$.

Предложение 9.2.1. *Отображение μ есть гомоморфизм алгебр Ли. Если $\mathbb{G}$ действует эффективно на $\mathbb{M}$, то μ есть изоморфизм алгебры Ли $\mathfrak{g}$ и ее образа $\mu(\mathfrak{g}) \subset \mathcal{X}(\mathbb{M})$. Если $\mathbb{G}$ действует свободно на $\mathbb{M}$, то для каждого ненулевого элемента алгебры Ли $X \in \mathfrak{g}$, его образ $\mu(X) = X^*$ не обращается в нуль на $\mathbb{M}$.*

Доказательство. Поскольку дифференциал отображения – это линейная операция, то из определения (9.8) следует, что μ есть линейное отображение из $\mathfrak{g}$ в $\mathcal{X}(\mathbb{M})$. Доказательство того факта, что отображение μ сохраняет коммутатор векторных полей:

$$[X^*, Y^*] = \mu[X, Y], \quad \text{где } X^*, Y^* \in \mathcal{X}(\mathbb{M}), X, Y \in \mathfrak{g},$$

технично и приведено, например, в [48], глава I, предложение 4.1.

Докажем второе утверждение. Пусть $\mu(X) = 0$ всюду на $\mathbb{M}$. Это значит, что однопараметрическая группа действует тривиально на $\mathbb{M}$, т.е. $a(t) = e$ для всех t . Верно и обратное утверждение. Пусть теперь $X \in \ker \mu$, т.е. $\mu(X) = 0$. Это значит, что $a(t) = e$ для всех t . Отсюда следует, что ядро $\ker \mu$ тривиально, $\mu(X) = 0$. Поэтому отображение μ – изоморфизм.

Пусть $\mu(X) = 0$ в некоторой точке $x \in \mathbb{M}$. Тогда точка x неподвижна, $xa(t) = x$ для всех t . Обратно. Если точка x неподвижна, то $a(t) = e$ для всех t и, следовательно, $X = 0$. $\square$

Это предложение является основой для рассмотрения главных расслоений со структурной группой Ли в разделе 12.1.

Мы видим, что транзитивные группы преобразований устроены довольно просто. В этом случае многообразие $\mathbb{M}$ – это одна орбита, и фактор пространство $\mathbb{M}/\mathbb{G}$ состоит из одной точки. Если действие группы $\mathbb{G}$ на многообразии $\mathbb{M}$ не является транзитивным, то орбит может быть много. В общем случае пространство орбит устроено сложно, что показывает следующий

Пример 9.2.1. Стационарной точкой в евклидовом пространстве $\mathbb{R}^3$ относительно действия группы трехмерных вращений $\mathrm{SO}(3)$ является начало координат. Группа изотропии начала координат совпадает со всей группой, а группой изотропии любой другой точки является подгруппа двумерных вращений $\mathrm{SO}(2) \subset \mathrm{SO}(3)$ вокруг оси, проходящей через данную точку и начало координат. Пространство орбит представляет собой множество сфер $\mathbb{S}_r^2$, которое можно параметризовать их радиусом $r \in \mathbb{R}_+$, объединенное с началом координат $\{0\} \in \mathbb{R}^3$,

$$\mathbb{R}^3/\mathrm{SO}(3) \approx \mathbb{R}_+ \cup \{0\} = \overline{\mathbb{R}}_+.$$

То есть факторпространство представляет собой положительную полуось $\overline{\mathbb{R}}_+$, включая начало координат, и не является многообразием, т.к. не является открытым подмножеством в $\mathbb{R}$ (оно является многообразием с краем). $\square$

Если группа преобразований $(\mathbb{M}, \mathbb{G})$ не является транзитивной, то многообразие $\mathbb{M}$ представляет собой объединение орбит, на каждой из которых группа $\mathbb{G}$ действует транзитивно. Выберем точку x_i на каждой орбите и обозначим через $\mathbb{G}_i$ ее группу изотропии. Тогда существует диффеоморфизм каждой орбиты на фактор пространство $\mathbb{G}/\mathbb{G}_i$. Другими словами, каждая орбита является однородным пространством и на нем существует естественная дифференцируемая структура (теорема 9.1.1). Согласно предложению 9.1.3 каждая орбита является замкнутым подмногообразием в $\mathbb{M}$.

Рассмотрим частный случай. Пусть L_A , $A = 1, \dots, \dim \mathbb{G}$, – базис алгебры Ли. Если векторные поля $\mu(L_A)$ линейно независимы в каждой точке, то они определяют инволютивное распределение на многообразии $\mathbb{M}$. Тогда, согласно теореме Фробениуса, каждая орбита с естественной дифференцируемой структурой является подмногообразием в $\mathbb{M}$. В общем случае, несмотря на изоморфизм отображения μ , векторные поля $\mu(L_A)$ могут быть линейно зависимы. Это показывает пример 9.2.1, в котором орбиты двумерны (за исключением 0-мерного начала координат), а алгебра Ли $\mathfrak{so}(3)$ – трехмерна. Вопрос о том можно ли ввести на пространстве орбит (фактор пространстве $\mathbb{M}/\mathbb{G}$) дифференцируемую структуру сложен и выходит за рамки данной книги.

В заключение настоящего раздела докажем утверждение, которое будет использовано при построении связностей в главном расслоении.

Предложение 9.2.2. Пусть a есть диффеоморфизм многообразия $\mathbb{M}$,

$$a : \mathbb{M} \ni x \mapsto a(x) \in \mathbb{M}.$$

Если векторное поле $X \in \mathcal{X}(\mathbb{M})$ порождает однопараметрическую группу преобразований $s_t : \mathbb{M} \rightarrow \mathbb{M}$, то векторное поле a_*X , где a_* – дифференциал отображения a , порождает однопараметрическую группу преобразований $a \circ s_t \circ a^{-1}$.

Доказательство. Ясно, что $a \circ s_t \circ a^{-1}$ есть однопараметрическая группа преобразований. Покажем, что эта группа преобразований индуцирует векторное поле a_*X . Пусть $x \in \mathbb{M}$ – произвольная точка и $y = a^{-1}(x)$. Так как отображение s_t индуцирует векторное поле X , то вектор $X_y \in \mathbb{T}_y(\mathbb{M})$ касается кривой $\gamma_1(t) = s_t(y)$ в точке $y = \gamma_1(0)$. Из этого следует, что вектор

$$(a_*X)_x = a_*X_y \in \mathbb{T}_x(\mathbb{M})$$

касается кривой $\gamma_2(t) = a \circ s_t(y) = a \circ s_t \circ a^{-1}(x)$. $\square$

Замечание. В данном предложении преобразование a не обязано принадлежать какой-либо группе. Это может быть произвольный фиксированный диффеоморфизм многообразия $\mathbb{M}$. $\square$

Следствие. Векторное поле $X \in \mathcal{X}(\mathbb{M})$ инвариантно относительно диффеоморфизма $a : \mathbb{M} \rightarrow \mathbb{M}$, т.е. $a_*X = X$, тогда и только тогда, когда диффеоморфизм a перестановочен с однопараметрической группой преобразований s_t , порожденной векторным полем X .

9.3 Инвариантные структуры

Рассмотрим группу преобразований $(\mathbb{M}, \mathbb{G})$. Предположим, что на многообразии $\mathbb{M}$ задана некоторая геометрическая структура, например, тензорное поле. Возникает вопрос при каких условиях эта структура является инвариантной относительно данной группы преобразований? Ответ на этот вопрос очень важен для приложений, поскольку часто мы ищем решения некоторой системы уравнений, обладающие определенной симметрией. Такие структуры называются $\mathbb{G}$ -инвариантными.

Начнем с простейшего случая. Предположим, что на многообразии $\mathbb{M}$ задана функция и действие группы преобразований $(\mathbb{M}, \mathbb{G})$. Определим действие элемента группы преобразований $a \in \mathbb{G}$ на функцию формулой

$$f(x) \mapsto \hat{f}(x) := f(xa^{-1}).$$

Отсюда следует равенство $\hat{f}(xa) = f(x)$. То есть значение новой функции $\hat{f}$ в точке xa равно значению исходной функции в точке x . Это – обычное преобразование функции при диффеоморфизмах.

Определение. Назовем функцию $\mathbb{G}$ -инвариантной, если она удовлетворяет условию

$$a : \quad f(xa) = f(x) \tag{9.9}$$

для всех $a \in \mathbb{G}$. $\square$

Из определения следует, что $\mathbb{G}$ -инвариантные функции постоянны на орбитах группы преобразований.

Пример 9.3.1. Сферически симметричные функции в трехмерном евклидовом пространстве $\mathbb{R}^3$ постоянны на сферах $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ с центром в начале координат. Эти функции в сферической системе координат зависят только от радиуса $f = f(r)$. $\square$

Пример 9.3.2. Если группа преобразований $(\mathbb{M}, \mathbb{G})$ транзитивна, то $\mathbb{G}$ -инвариантные функции на $\mathbb{M}$ постоянны. $\square$

Рассмотрим более сложную ситуацию.

Определение. Пусть на многообразии $\mathbb{M}$, на котором действует группа преобразований $(\mathbb{M}, \mathbb{G})$, задано поле $\varphi = \{\varphi^i(x)\}$, имеющее N компонент, $i = 1, \dots, N$, которые преобразуются по некоторому представлению группы $\mathbb{G}$. Это значит, что группа преобразований действует не только на точки многообразия, но одновременно и на компоненты φ^i по правилу

$$\begin{aligned} x &\mapsto x' = xa, \\ \varphi^i(x) &\mapsto \varphi^i(x') := \varphi^j(x) S_j^i(a), \end{aligned} \quad (9.10)$$

где $S_j^i(a)$ – матрица представления, соответствующая элементу группы $a \in \mathbb{G}$. Поле $\varphi(x)$ называется $\mathbb{G}$ -инвариантным, если для его компонент выполнено условие

$$\varphi^i(xa) = \varphi^j(x) S_j^i(a). \quad \square \quad (9.11)$$

В отличие от правила преобразования поля (9.10) в левой части этого равенства опущен штрих у поля, и поэтому условие $\mathbb{G}$ -инвариантности представляет собой систему уравнений на компоненты $\varphi^i(x)$. В случае, если представление группы преобразований тривиально, $S_j^i(a) = \delta_j^i$ для всех $a \in \mathbb{G}$, мы имеем набор $\mathbb{G}$ -инвариантных скалярных полей (9.9).

Условие $\mathbb{G}$ -инвариантности можно переписать в эквивалентном виде

$$\varphi^i(x) = \varphi^j(xa^{-1}) S_j^i(a).$$

Пример 9.3.3. Рассмотрим сферически инвариантные векторные поля в трехмерном евклидовом пространстве $\mathbb{R}^3$. Нетрудно проверить, что векторное поле с компонентами

$$\varphi^i(x) = \frac{x^i}{r} f(r), \quad i = 1, 2, 3,$$

заданными в декартовой системе координат x^i , где $f(r)$ – произвольная функция от радиуса $r := \sqrt{(x^1)^2 + \dots + (x^3)^2}$, является сферически симметричным. Можно также доказать обратное утверждение, что любое сферически симметричное векторное поле имеет такой вид для некоторой функции $f(r)$. Таким образом, сферически симметричные векторные поля на $\mathbb{R}^3$ параметризуются одной функцией радиуса $f(r)$. $\square$

В разделе 7.1 приведены также сферически симметричные тензоры второго ранга.

Замечание. При исследовании моделей математической физики предположение о симметрии полей позволяет существенно уменьшить число независимых переменных и число аргументов. Это позволяет в ряде случаев найти явно точное решение уравнений Эйлера–Лагранжа. $\square$

Замечание. При рассмотрении групп преобразований мы предполагаем, что преобразования многообразия $\mathbb{M}$ являются глобальными, т.е. элемент группы $a \in \mathbb{G}$ не зависит от точки многообразия $x \in \mathbb{M}$. При таких преобразованиях калибровочные поля преобразуются, как тензоры. Поэтому условие $\mathbb{G}$ -инвариантности (9.11) можно распространить и на калибровочные поля. $\square$

Рассмотрим общую ситуацию. Пусть задана группа преобразований $(\mathbb{M}, \mathbb{G})$. Каждому элементу группы $a \in \mathbb{G}$ соответствует диффеоморфизм $\mathbb{M} \xrightarrow{a} \mathbb{M}$, который мы обозначим той же буквой, что и элемент группы. У этого отображения существует

дифференциал a_* и возврат отображения a^* . Возврат отображения обратим, т.к. a – диффеоморфизм. Пусть в точке $x \in \mathbb{M}$ задан тензор $T(x) \in \mathbb{T}_{s,x}^r(\mathbb{M})$ типа (r, s) . Тогда его можно перенести из точки x в точку xa с помощью дифференциала отображения и возврата отображения, которые действуют, соответственно, на контравариантные и ковариантные индексы, $T(xa) = (a_*)^r (a^{*-1})^s T(x)$.

В координатах это отображение записывается следующим образом. Пусть $x = \{x^\alpha\}$ и $xa = \{y^\alpha\}$, $\alpha = 1, \dots, \dim \mathbb{M}$, – координаты точек. Тогда

$$T_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_r}(xa) = \frac{\partial x^{\gamma_1}}{\partial y^{\beta_1}} \dots \frac{\partial x^{\gamma_s}}{\partial y^{\beta_s}} T_{\gamma_1 \dots \gamma_s}^{\delta_1 \dots \delta_r}(x) \frac{\partial y^{\alpha_1}}{\partial x^{\delta_1}} \dots \frac{\partial y^{\alpha_r}}{\partial x^{\delta_r}},$$

где матрицы Якоби вычисляются в точке x .

Определение. Произвольное тензорное поле $T \in \mathcal{T}_s^r(\mathbb{M})$, заданное на многообразии $\mathbb{M}$, называется $\mathbb{G}$ -инвариантным, если выполнено условие

$$T(xa) := (a_*)^r (a^{*-1})^s T(x) \quad (9.12)$$

для всех $x \in \mathbb{M}$ и $a \in \mathbb{G}$. □

Замечание. Это определение применимо как для счетных групп, так и для групп Ли. □

Поскольку действие группы преобразований транзитивно на каждой орбите, то тензорное поле на всей орбите однозначно задается его значением в некоторой точке x на данной орбите. Это не значит, что тензор в точке x можно задать произвольно, а затем разнести по орбите. Если группа изотропии $\mathbb{G}_x$ нетривиальна, то в соответствии с определением (9.12) тензор в точке x должен быть выбран симметричным относительно преобразований из $\mathbb{G}_x$.

В случае, когда $\mathbb{G}$ – группа Ли, условие (9.12) можно записать в дифференциальной форме. Генераторы группы Ли L_A , $A = 1, \dots, \dim \mathbb{G}$, – это базис левоинвариантных векторных полей на $\mathbb{G}$. Им соответствуют векторные поля L_A^* на многообразии $\mathbb{M}$. Из определения производной Ли следует, что если тензорное поле является $\mathbb{G}$ -инвариантным, то

$$\mathbb{L}_{L_A^*} T = 0$$

для всех A и $x \in \mathbb{M}$. Это – система дифференциальных уравнений в частных производных первого порядка на компоненты $\mathbb{G}$ -инвариантного тензорного поля.

Замечание. В математической физике, как правило, ставится одна из двух задач. 1) По заданному тензорному полю определить его группу симметрии и 2) найти тензорное поле, обладающее определенной симметрией. □

В моделях гравитации важную роль играют однопараметрические группы преобразований, которые оставляют инвариантной метрику, заданную на многообразии $\mathbb{M}$. Этим группам преобразований соответствуют векторные поля, которые называются векторными полями Киллинга. Им будет посвящен отдельный раздел 15.

9.4 Отображения групп преобразований

Пусть задано две группы преобразований (M, G) и (N, G) двух многообразий M и N с одной и той же группой G . В частном случае многообразия могут совпадать, $M = N$, но действия группы G на M можно задавать различными способами.

Пример 9.4.1. Можно рассматривать группу вращений плоскости $(\mathbb{R}^2, SO(2))$ относительно произвольной точки $x \in \mathbb{R}^2$. $\square$

Обозначим действие элемента группы $a \in G$ на M и N , соответственно, через a_M и a_N :

$$\begin{aligned} a_M : M \times G &\ni x, a \mapsto xa \in M, \\ a_N : N \times G &\ni y, a \mapsto ya \in N. \end{aligned}$$

Определение. Отображение $f : M \rightarrow N$ называется *эквивариантным*, если диаграмма

$$\begin{array}{ccc} M \times G & \xrightarrow{f \times \text{id}} & N \times G \\ a_M \downarrow & & \downarrow a_N \\ M & \xrightarrow{f} & N \end{array}$$

где id – тождественное отображение $G \rightarrow G$, коммутативна. То есть

$$f \circ a_M = a_N \circ f \quad \text{или} \quad f(xa) = f(x)a$$

для всех $a \in G$ и $x \in M$. $\square$

Понятие эквивариантного отображения можно обобщить. Рассмотрим две группы преобразований (M, G) и (N, H) . Пусть задан гомоморфизм групп $\rho : G \rightarrow H$ и достаточно гладкое отображение многообразий $f : M \rightarrow N$.

Определение. Отображение $f \times \rho : M \times G \rightarrow N \times H$ называется *эквивариантным*, если диаграмма

$$\begin{array}{ccc} M \times G & \xrightarrow{f \times \rho} & N \times H \\ a_M \downarrow & & \downarrow \rho(a)_N \\ M & \xrightarrow{f} & N \end{array}$$

коммутативна. То есть

$$f \circ a_M = \rho(a)_N \circ (f \times \rho) \quad \text{или} \quad f(xa) = f(x)\rho(a)$$

для всех $a \in G$ и $x \in M$. $\square$

Пример 9.4.2. Важным физическим применением указанной конструкции является понятие спинора. Рассмотрим две группы преобразований: группу вращений $SO(3)$ трехмерного евклидова пространства $\mathbb{R}^3$ и группу $SU(2)$ линейных преобразований двумерного комплексного пространства $\mathbb{C}^2$. Согласно теореме 1.8.1, группа $SU(2)$ два

раза покрывает группу вращений $\mathbb{SO}(3)$. Обозначим соответствующий гомоморфизм $\rho : \mathbb{SU}(2) \rightarrow \mathbb{SO}(3)$. Тогда для спиноров имеем коммутативную диаграмму

$$\begin{array}{ccc} \mathbb{C}^2 \times \mathbb{SU}(2) & \xrightarrow{f \times \rho} & \mathbb{R}^3 \times \mathbb{SO}(3) \\ \downarrow a_{\mathbb{C}^2} & & \downarrow \rho(a)_{\mathbb{R}^3}, \\ \mathbb{C}^2 & \xrightarrow{f} & \mathbb{R}^3 \end{array}$$

где $a_{\mathbb{C}^2} = U$ – унитарная 2×2 матрица с единичным определителем линейных преобразований двумерного комплексного пространства $\mathbb{C}^2$, и $\rho(a)_{\mathbb{R}^3}$ – соответствующая ортогональная 3×3 матрица вращений трехмерного евклидова пространства $\mathbb{R}^3$. Отображение f задается формулой $x^i = (\bar{\psi} \sigma^i \psi)$, где σ^i , $i = 1, 2, 3$, – матрицы Паули, $\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} \in \mathbb{C}^2$ – столбец из двух комплексных чисел (компоненты спинора) и $\bar{\psi} = (\psi_1^*, \psi_2^*)$ – строка из двух комплексно сопряженных чисел (компоненты сопряженного спинора). В физике принято называть матрицы U – спинорным представлением группы вращений $\mathbb{SO}(3)$. Строго говоря, соответствие

$$\mathbb{SO}(3) \ni \rho(a) \mapsto a \in \mathbb{SU}(2)$$

не является отображением и, следовательно, представлением, т.к. каждому вращению $\rho(a) \in \mathbb{SO}(3)$ соответствуют две унитарные матрицы $\pm a \in \mathbb{SU}(2)$, которые отличаются знаком. Поэтому спинорное представление называют двузначным представлением. $\square$

Глава 10

Гомотопии и фундаментальная группа

Все многообразия локально устроены одинаково так же как и евклидовы пространства. Однако их глобальная структура может существенно отличаться. Описание глобальной структуры многообразий является сложной задачей, и для ее решения применяются различные методы. Например, на каждом многообразии можно задать некоторую алгебраическую структуру, которая является топологическим инвариантом, т.е. сохраняется при диффеоморфизмах. В настоящем разделе мы опишем один из простейших топологических инвариантов – фундаментальную группу. Этот инвариант довольно груб и не различает все топологически различные многообразия. Однако он полезен в приложениях, т.к. если фундаментальная группа двух многообразий различна, то можно с уверенностью сказать, что данные многообразия недиффеоморфны.

10.1 Пути

Важным методом алгебраического анализа глобальной структуры многообразий является нахождение фундаментальной группы многообразия, которая строится на основе понятия *пути* или *кривой*. Определение кривой в евклидовом пространстве (1.9) дословно переносится на произвольное многообразие $\mathbb{M}$ размерности n .

Определение. *Путем* γ называется кусочно дифференцируемое отображение единичного отрезка $[0, 1] \in \mathbb{R}$ в $\mathbb{M}$,

$$\gamma : [0, 1] \ni t \mapsto x(t) \in \mathbb{M}, \quad (10.1)$$

где t – вещественный параметр вдоль кривой. Говорят, что путь соединяет *начало* пути $x(0)$ и его *конец* $x(1)$. Если начало и конец пути совпадают, $x(0) = x(1)$, то путь называют *замкнутым*. Замкнутый путь называют также *петлей*. Мы будем также писать $\gamma = x(t)$. $\square$

В координатах отображение (10.1) задается набором n кусочно дифференцируемых функций $x(t) = \{x^\alpha(t)\}$, $\alpha = 1, \dots, n$, одной переменной $t \in [0, 1]$.

Замечание. В общем случае кривая $x(t)$ может иметь точки самопересечения, и тогда отображение γ не является взаимно однозначным отображением отрезка $[0, 1]$ на его образ $\gamma([0, 1])$.

В алгебраической топологии, где в роли $\mathbb{M}$ выступают топологические пространства, от отображения (10.1) требуется только непрерывность. Для алгебраического анализа многообразий удобно рассматривать кусочно дифференцируемые пути, т.е. функции $x^\alpha(t)$ считать кусочно дифференцируемыми. Это удобно при умножении путей, которое будет введено позже. $\square$

Замкнутую кривую в $\mathbb{M}$ можно рассматривать, как непрерывный образ окружности, т.к. при совпадении начала и конца пути можно отождествить концы единичного интервала $t \in [0, 1]$.

Определение. *Простой замкнутой или жордановой кривой называется гомеоморфный образ окружности, т.е. замкнутая кривая без точек самопересечения. В общем случае простая замкнутая кривая может иметь изломы, т.е. от нее требуется только непрерывность. Жорданова кривая на евклидовой плоскости, состоящая из конечного числа прямолинейных отрезков, называется жордановым многоугольником.* $\square$

Теорема 10.1.1 (Жордан). *Пусть $\mathbb{S}$ – простая замкнутая кривая на евклидовой плоскости $\mathbb{R}^2$. Тогда $\mathbb{R}^2 \setminus \mathbb{S}$ несвязно и состоит из двух компонент, общей границей которых является $\mathbb{S}$. В точности одна из этих компонент ограничена.*

Доказательство. Несмотря на интуитивную очевидность утверждения теоремы, ее доказательство сложно. Оно приведено, например, в [19], теорема 12A.12. $\square$

В определении пути (10.1) область определения параметра вдоль пути фиксирована единичным отрезком. Этого можно не делать, считая, что параметр принимает значения на отрезке произвольной длины $[a, b] \subset \mathbb{R}$. В этом случае все пути можно разделить на классы эквивалентности. А именно, два пути $x_1(t)$, $t \in [a, b]$, и $x_2(s)$, $s \in [0, 1]$ называются *эквивалентными*, если они связаны достаточно гладкой монотонной заменой параметра $s = s(t)$:

$$x_1(t) = x_2(s(t)), \quad \frac{ds}{dt} > 0.$$

где $s(a) = 0$ и $s(b) = 1$. В частном случае, возможно преобразование параметра t внутри единичного отрезка, $a = 0$, $b = 1$. В дальнейшем под путем будем понимать класс эквивалентных путей, определенных с точностью до гладкой репараметризации.

Замечание. Иногда понятие пути и кривой различают, понимая под путем отображение (10.1), а под кривой – образ пути $\gamma([0, 1])$ в многообразии $\mathbb{M}$. Как правило, мы этого не делаем и употребляем термины путь и кривая как синонимы. При изучении фундаментальных групп отображение (10.1) будем называть путем, как это принято в алгебраической топологии. $\square$

Пример 10.1.1. Простейшим примером пути является *тождественный* путь e_x в некоторой точке $x \in \mathbb{M}$, определенный равенством

$$e_x : [0, 1] \ni t \mapsto x \in \mathbb{M} \quad (10.2)$$

для всех t , где $x \in \mathbb{M}$ – произвольная фиксированная точка многообразия. $\square$

Пример 10.1.2. В классической механике под многообразием $\mathbb{M}$ часто понимается пространство, а под вещественным параметром t – время. Тогда путь $x(t) = \{x^\mu(t)\}$ – это траектория точечной частицы. $\square$

При рассмотрении евклидова пространства $\mathbb{R}^n$ как топологического пространства в разделе 1.3.2 было введено понятие связности. Введем новое понятие связности, которое удобно, в частности, при рассмотрении фундаментальных групп.

Определение. Топологическое пространство $\mathbb{M}$ называется *линейно связным*, если для любых двух точек $x_1, x_2 \in \mathbb{M}$ найдется соединяющий их путь, целиком лежащий в $\mathbb{M}$. Топологическое пространство $\mathbb{M}$ называется *локально линейно связным*, если для любого $x \in \mathbb{M}$ всякая открытая окрестность точки x содержит линейно связную окрестность этой точки. $\square$

Если какую-либо фиксированную точку $x_0 \in \mathbb{M}$ можно соединить путем с любой другой точкой $x_1 \in \mathbb{M}$, то этого достаточно для линейной связности многообразия, потому что тогда две произвольные точки $x_1, x_2 \in \mathbb{M}$ можно соединить путем, проходящим через x_0 .

Между просто связностью и линейной связностью топологических пространств имеется тесная связь, которая дается следующей теоремой.

Теорема 10.1.2. *Всякое линейно связное топологическое пространство связно.*

Доказательство. Зафиксируем точку $x_0 \in \mathbb{M}$. Тогда ее можно соединить некоторой кривой γ_{x_0} , целиком лежащей в $\mathbb{M}$, с произвольной точкой $x \in \mathbb{M}$, потому что $\mathbb{M}$ – линейно связно. Каждая кривая γ_x является связным подмножеством в $\mathbb{M}$. Тогда все пространство $\mathbb{M}$ является объединением, $\mathbb{M} = \bigcup_x \gamma_x$, и связно, как объединение связных подмножеств, имеющих непустое пересечение (теорема 1.3.13). $\square$

Обратное утверждение теоремы 10.1.2 неверно. Приведем пример связного, но не линейно связного топологического пространства.

Пример 10.1.3 (Блоха и гребенка). Пусть подмножество комплексной плоскости $\mathbb{W} \subset \mathbb{C}$ является объединением двух множеств $\mathbb{W} = \mathbb{U} \cup \mathbb{V}$ (см. рис. 10.1), где

$$\begin{aligned} \mathbb{U} &= \{i\} && \text{– блоха} \\ \mathbb{V} &= [0, 1] \cup \{1/n + iy : n \in \mathbb{N}, 0 \leq y \leq 1\} && \text{– гребенка.} \end{aligned}$$

Будем считать, что топология $\mathbb{W}$ индуцирована вложением в комплексную плоскость

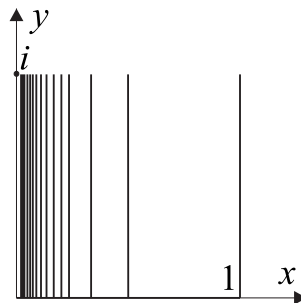


Рис. 10.1: Пример связного, но не линейно связного топологического пространства “блоха и гребенка”.

$\mathbb{C} \approx \mathbb{R}^2$. Прежде всего заметим, что множество $\mathbb{V}$ линейно связно как объединение линейно связных множеств (теорема 1.3.13) и потому связно. Пусть $\mathbb{U}_\epsilon$ – ϵ -окрестность точки i , т.е. пересечение множества $\{z : |z - i| < \epsilon\}$ с $\mathbb{W}$. Подмножество $\mathbb{U}$, состоящее из одной точки, принадлежит замыканию $\overline{\mathbb{V}}$, т.к. для любого $\epsilon > 0$, пересечение $\mathbb{U}_\epsilon \cap \mathbb{V}$ не пусто, поскольку для достаточно больших n точки $1/n + i \in \mathbb{V}$ лежат в ϵ -окрестности точки i (лемма 1.3.2). Отсюда вытекает, что связная компонента $\mathbb{V}$ содержит $\mathbb{U}$ и поэтому совпадает со всем $\mathbb{W}$.

Покажем теперь, что топологическое пространство $\mathbb{W}$ не является линейно связным, т.к. точку $i \in \mathbb{W}$ нельзя соединить путем с любой другой точкой из $\mathbb{W}$. Предположим, что существует путь с началом в точке $\gamma(0) = i$ и концом в некоторой точке $\gamma(1) \in \mathbb{V}$. Не ограничивая общности можно считать, что $\gamma(t) \neq i$ при $t > 0$. Из непрерывности отображения $\gamma(t)$ следует, что для любого положительного ϵ существует такое положительное число δ , что при $t < \delta$ выполняется неравенство $|\gamma(t) - i| < \epsilon$, т.е. путь лежит в ϵ -окрестности $\mathbb{U}_\epsilon$ точки i . Выберем конечное число $\epsilon < 1$ и некоторое значение параметра $0 < t_0 < \delta$. По предположению, $x(t_0) \in \mathbb{V}$. Окрестность $\mathbb{U}_\epsilon$ многосвязна, т.к. состоит из самой точки i и объединения непересекающихся интервалов. Отрезок $[0, t_0]$ связан и, как показано в примере 1.4.10, его невозможно непрерывно отобразить на многосвязную область таким образом, чтобы начало и конец отрезка лежали в разных компонентах. $\square$

В рассмотренном примере топологическое пространство $\mathbb{W}$ не является многообразием по двум причинам. Во-первых, окрестность точки i нельзя отобразить на интервал и, во вторых, множество $\mathbb{V}$ имеет точки самопересечения.

Следующий пример показывает, что не всякое линейно связное топологическое пространство является локально линейно связным.

Пример 10.1.4 (Польская окружность). Рассмотрим подмножество на евклидовой плоскости $\mathbb{U} \subset \mathbb{R}^2$, которое имеет вид $\mathbb{U} = \mathbb{U}_1 \cup \mathbb{U}_2 \cup \mathbb{U}_3$, где

$$\begin{aligned}\mathbb{U}_1 &= \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1, y \geq 0\}, \\ \mathbb{U}_2 &= \{(x, y) \in \mathbb{R}^2 : -1 \leq x \leq 0, y = 0\}, \\ \mathbb{U}_3 &= \{(x, y) \in \mathbb{R}^2 : 0 < x \leq 1, y = \frac{1}{2} \sin \frac{\pi}{x}\},\end{aligned}$$

с индуцированной топологией, см. рис. 10.2. Это пространство линейно связно, но не локально линейно связно, т.к. начало координат $(0, 0)$ не имеет окрестности, содержащей линейно связную окрестность. Это топологическое пространство не является многообразием, поскольку начало координат не имеет окрестности, гомеоморфной прямой $\mathbb{R}$. $\square$

Следующие свойства справедливы для произвольных топологических пространств.

Предложение 10.1.1. *Образ линейно связного многообразия при непрерывном отображении линейно связан.*

Доказательство. См., например, [19], теорема 12.4. $\square$

Следствие. Если $\mathbb{M}_1$ и $\mathbb{M}_2$ – два гомеоморфных топологических пространства, то $\mathbb{M}_1$ линейно связно тогда и только тогда, когда $\mathbb{M}_2$ линейно связно. $\square$

Предложение 10.1.2. *Многообразия $\mathbb{M}_1$ и $\mathbb{M}_2$ линейно связны тогда и только тогда, когда их прямое произведение $\mathbb{M}_1 \times \mathbb{M}_2$ линейно связно.*

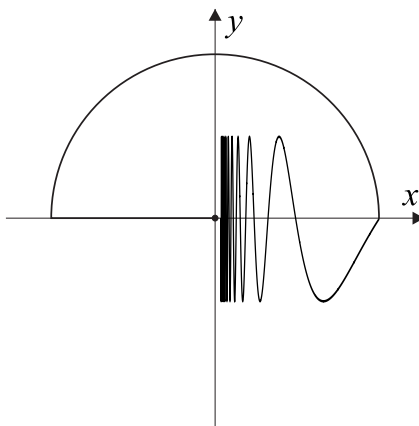


Рис. 10.2: Пример линейно связного, но не локально линейного связного топологического пространства.

Доказательство. См., например, в [19], теорема 12.7. □

Для многообразий понятие связности и линейной связности эквивалентны.

Теорема 10.1.3. *Многообразие M связно тогда и только тогда, когда M – линейно связно.*

Доказательство. Достаточность является утверждением теоремы 10.1.2.

Необходимость докажем от противного. Пусть M связное, но не линейно связное многообразие. Прежде всего заметим, что каждое многообразие можно покрыть счетным числом карт. Каждая карта является линейно связным подмножеством, потому что гомеоморфна $\mathbb{R}^n$. Если две карты пересекаются, то их объединение также линейно связно. Это значит, что многообразие можно представить в виде объединения не более, чем счетного числа непересекающихся подмножеств, $M = \bigcup_i U_i$, где каждое U_i линейно связно и $U_i \cap U_j = \emptyset$ при $i \neq j$. Поскольку M связно, то среди подмножеств U_i существует по крайней мере одно подмножество, пусть это будет U_i , которое содержит предельную точку некоторого другого подмножества U_j . Пусть точка $x_0 \in U_i$ является предельной для подмножества U_j . Это значит, что в U_j существует последовательность точек $\{x_k\} \in U_j$, которая сходится к x_0 . Поэтому, начиная с некоторого N , все точки последовательности x_k , $k > N$, лежат в окрестности $U_0 \ni x_0$. Следовательно, точку x_0 можно соединить кривой с x_k и подмножества U_i и U_j – линейно связны, что противоречит предположению. □

Пример 10.1.5. Евклидово пространство $\mathbb{R}^n$ линейно связно, т.к. любые две точки $x_0, x_1 \in \mathbb{R}^n$ можно соединить путем $x(t) = x_0 + (x_1 - x_0)t$, который представляет собой прямолинейный отрезок. □

Нетрудно также доказать

Предложение 10.1.3. *Каждое многообразие M локально линейно связно.*

В общем случае многосвязное многообразие представляет собой набор компонент связности, каждая из которых является линейно связной.

Для полноты картины приведем достаточное условие линейной связности произвольного топологического пространства.

Предложение 10.1.4. *Если топологическое пространство M связно и локально линейно связно, то оно линейно связно.*

Доказательство. См., например, [19], упражнение 12.10(j). □

10.2 Гомотопия непрерывных отображений

На множестве непрерывных отображений двух топологических пространств $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ и, в частности, многообразий можно ввести отношение эквивалентности – гомотопию. Понятие гомотопии является топологическим, потому что при его определении используется только непрерывность отображений.

Определение. Два непрерывных отображения

$$\begin{aligned} f_0 : \mathbb{M}_1 \ni x &\mapsto f_0(x) \in \mathbb{M}_2, \\ f_1 : \mathbb{M}_1 \ni x &\mapsto f_1(x) \in \mathbb{M}_2, \end{aligned}$$

называются *гомотопными*, если существует непрерывное отображение

$$F : \mathbb{M}_1 \times [0, 1] \ni x, t \mapsto F(x, t) \in \mathbb{M}_2$$

такое, что $F(x, 0) = f_0(x)$ и $F(x, 1) = f_1(x)$. При этом само отображение F называется *гомотопией* между f_0 и f_1 :

$$F : f_0 \sim f_1.$$

Гомотопия $F(x, t)$ называется также *деформацией* отображения f_0 в f_1 . $\square$

Другими словами, гомотопия позволяет непрерывно деформировать одно отображение в другое. Нетрудно проверить, что гомотопия является отношением эквивалентности на множестве всех отображений $\mathbb{M}_1 \rightarrow \mathbb{M}_2$. Отображения, гомотопные отображению f , образуют гомотопический класс отображений, который будем обозначать квадратными скобками $[f]$. Множество всех гомотопических классов отображений $\mathbb{M}_1 \rightarrow \mathbb{M}_2$ будем обозначать через $[\mathbb{M}_1; \mathbb{M}_2]$.

Гомотопию $F(x, t)$ можно представить себе следующим образом. Зафиксируем точку $x \in \mathbb{M}_1$ и будем менять параметр $t \in [0, 1]$. Тогда мы получим траекторию точки x в пространстве $\mathbb{M}_2$. Тогда гомотопия $F(x, t)$ представляет собой семейство всех траекторий, зависящих от параметра $x \in \mathbb{M}_1$.

Пример 10.2.1. Любой путь $\gamma : [0, 1] \ni s \mapsto x(s) \in \mathbb{M}$ гомотопен постоянному пути e_{x_0} с началом и концом в точке $x_0 := x(0)$ посредством гомотопии

$$F : [0, 1] \times [0, 1] \ni s, t \mapsto F(s, t) \in \mathbb{M}, \quad \text{где } F(s, t) := x((1 - t)s).$$

В этой гомотопии конец пути меняется: он скользит вдоль кривой $\gamma(s)$, приближаясь к началу пути при $t \rightarrow 1$. При данном выше определении гомотопии, когда отображения f_0 и f_1 не обязаны совпадать ни в одной точке, все пути на связном многообразии гомотопны между собой. $\square$

Понятие гомотопных отображений можно использовать для определения отношения эквивалентности самих топологических пространств.

Определение. Говорят, что пространства $\mathbb{M}_1$ и $\mathbb{M}_2$ имеют один и тот же *гомотопический тип*, если существуют непрерывные отображения $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ и $g : \mathbb{M}_2 \rightarrow \mathbb{M}_1$ такие, что композиции этих отображений гомотопны тождественным отображениям:

$$g \circ f \sim \text{id}_1 : \mathbb{M}_1 \rightarrow \mathbb{M}_1, \quad f \circ g \sim \text{id}_2 : \mathbb{M}_2 \rightarrow \mathbb{M}_2,$$

где $\text{id}_{1,2}$ – тождественные отображения пространств $\mathbb{M}_1$ и $\mathbb{M}_2$. Отображения f и g называются в этом случае *гомотопическими эквивалентностями*. Говорят также, что пространства $\mathbb{M}_1$ и $\mathbb{M}_2$ в этом случае *гомотопически эквивалентны*. $\square$

Это определение отличается от определения гомеоморфизма тем, что вместо знака равенства стоит знак гомотопической эквивалентности. Поэтому два гомеоморфных пространства имеют одинаковый гомотопический тип. Обратное утверждение неверно.

Пример 10.2.2. n -мерный шар $\mathbb{B}^n \in \mathbb{R}^n$ гомотопически эквивалентен точке $x \in \mathbb{B}^n$, но не гомеоморфен ей. $\square$

Определение. В общем случае топологическое пространство, гомотопически эквивалентное точке, называется *стягиваемым*. $\square$

Другими словами, пространство стягиваемо, если его можно непрерывно деформировать по себе в точку.

Пример 10.2.3. Нестягиваемыми пространствами являются n -мерная сфера $\mathbb{S}^n$ и тор $\mathbb{T}^n$. $\square$

Пример 10.2.4. Евклидово пространство $\mathbb{R}^n$, а также любая выпуклая область $U \subset \mathbb{R}^n$ гомотопически эквивалентна одной точке $x_0 \in U$. Напомним, что область $U \subset \mathbb{R}^n$ называется *выпуклой*, если наряду с произвольными точками $x_1, x_2 \in U$ она содержит также и отрезок $x_1 + (x_2 - x_1)t$, их соединяющий. $\square$

Пример 10.2.5. Евклидово пространство с выколотой точкой, $\mathbb{R}^n \setminus x_0$, гомотопически эквивалентно сфере $\mathbb{S}^{n-1}$, содержащей точку x_0 внутри. В частности, плоскость с выколотой точкой гомотопически эквивалентна окружности. $\square$

Пример 10.2.6. Плоскость с двумя выколотыми точками $x_1, x_2 \in \mathbb{R}^2$ гомотопически эквивалентна *букету окружностей*, изображенному на рис. 10.3. Букет окружностей представляет собой объединение двух окружностей, имеющих одну общую точку. Он является топологическим пространством, но не многообразием. $\square$

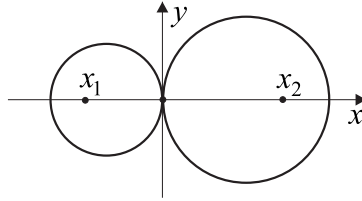


Рис. 10.3: Букет двух окружностей.

Пример 10.2.7. Гомотопически эквивалентными пространствами является тройка: цилиндр, лист Мебиуса и окружность. $\square$

Из этих примеров видно, что гомотопически эквивалентные многообразия могут иметь разную размерность и ориентируемость. Многообразия могут быть также гомотопически эквивалентны топологическим пространствам, которые не являются многообразиями, как показывает пример плоскости с двумя выколотыми точками. Тем не менее некоторые топологические свойства являются общими для гомотопически эквивалентных пространств.

Предложение 10.2.1. Если пространство M_1 связно и имеет тот же гомотопический тип, что и M_2 , то пространство M_2 также связно.

Понятие гомотопии может быть использовано для доказательства нетривиальных утверждений.

Теорема 10.2.1 (О невозможности причесать ежа). *На четномерной сфере $\mathbb{S}^n$ нельзя задать непрерывное векторное поле, которое не обращается в нуль ни в одной точке.*

Доказательство. От противного. Рассмотрим сферу единичного радиуса

$$\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} : x^\alpha x_\alpha = 1, \quad \alpha = 1, \dots, n+1\},$$

вложенную в евклидово пространство $\mathbb{R}^{n+1}$. Допустим, что на сфере задано непрерывное векторное поле X , не имеющее нулей. Тогда векторное поле $T := X/\sqrt{X^2}$ имеет единичную длину и поэтому задает непрерывное отображение сферы на себя $\mathbb{S}^n \rightarrow \mathbb{S}^n$. Касательный вектор к сфере $T = \{T^\alpha(x)\}$ ортогонален радиус-вектору $\{x^\alpha\}$, т.е. $x^\alpha T_\alpha = 0$. При этом мы рассматриваем координаты точки на сфере и касательный вектор как векторы евклидова пространства $\mathbb{R}^{n+1}$. Тогда отображение

$$F(x, t) : \quad x^\alpha, t \mapsto x^\alpha \cos(\pi t) + T^\alpha \sin(\pi t)$$

задает гомотопию

$$F : \quad \mathbb{S}^n \times [0, 1] \rightarrow \mathbb{S}^n.$$

Эта гомотопия переводит произвольную точку на сфере $\{x^\alpha\}$ в ее антиподальную точку с координатами $\{-x^\alpha\}$. Выберем ориентацию на сфере, согласованную с канонической ориентацией евклидова пространства (3.80). Тогда объем сферы равен

$$V = \int_{\mathbb{S}^n} v \neq 0, \tag{10.3}$$

где форма объема на сфере v получена сужением канонической формы объема евклидова пространства v_0 на сферу $\mathbb{S}^n$:

$$v = i_N v_0, \quad v_0 := dx^1 \wedge dx^2 \wedge \dots \wedge dx^{n+1}.$$

Здесь i_N – внутреннее умножение формы объема v_0 на внешнюю нормаль к сфере N . Интеграл (10.3) задает непрерывное отображение $\Lambda_{n+1}(\mathbb{R}^{n+1}) \rightarrow \mathbb{R}$. При отображении $x^\alpha \rightarrow -x^\alpha$ объем сферы преобразуется по правилу

$$V \mapsto (-1)^{n+1} V.$$

Поскольку изменение знака объема при гомотопии невозможно, т.к. она является непрерывным отображением, то для четных n приходим к противоречию. $\square$

Замечание. На нечетномерной сфере непрерывные векторные поля, нигде не обращающиеся в нуль, существуют. А именно, можно доказать [72], что максимальное число линейно независимых нигде не обращающихся в нуль векторных полей на нечетномерной сфере $\mathbb{S}^{n-1}$ равно $2^c + 8d - 1$, где c и d – неотрицательные целые числа, которые определяются следующим образом. Поскольку n четно, то его единственным образом можно представить в виде $n = (2a - 1)2^b$, где a, b – натуральные числа. Тогда $c := b \bmod 4$ и $d := (b - c)/4$. $\square$

Рассмотренный выше пример 10.2.1 гомотопных путей показывает, что определение гомотопии является грубым, т.к. гомотопические классы путей содержат слишком много элементов. Более тонкое деление на классы дает понятие отображений, гомотопных относительно некоторого подмножества $\mathbb{U} \subset \mathbb{M}_1$.

Определение. Два непрерывных отображения топологических пространств

$$f_0, f_1 : \mathbb{M}_1 \rightarrow \mathbb{M}_2$$

называются *гомотопными относительно подмножества* $\mathbb{U} \subset \mathbb{M}_1$, если существует гомотопия

$$F : \mathbb{M}_1 \times [0, 1] \rightarrow \mathbb{M}_2,$$

такая, что $F(x, t)$ не зависит от t при $x \in \mathbb{U}$. Мы пишем $f_0 \sim f_1(\text{rel } \mathbb{U})$. $\square$

В частности, отображения f_0 и f_1 должны совпадать на подмножестве $\mathbb{U}$. Другими словами, при непрерывной деформации отображений допускается их изменение только на множестве $\mathbb{M} \setminus \mathbb{U}$. Соответствующая гомотопия называется *относительной гомотопией* и также определяет отношение эквивалентности на множестве непрерывных отображений топологических пространств, которое является более тонким.

Пример 10.2.8. Можно рассмотреть класс путей, гомотопных относительно начала $x(0)$ и конца $x(1)$ некоторого пути. Этот класс содержит все пути, которые имеют фиксированное начало и конец и которые можно непрерывно деформировать друг в друга по многообразию $\mathbb{M}$. Более того, как показывает пример тора, не все пути с одинаковым началом и концом можно непрерывно деформировать друг в друга, не выходя из многообразия $\mathbb{M}$. Если начало и конец пути отличаются, то пути относительно негомотопны тождественному пути. Это говорит о том, что понятие относительной гомотопии более тонко, чем просто гомотопия. $\square$

Понятие относительной гомотопии для путей будет использовано при определении фундаментальной группы многообразия.

10.3 Фундаментальная группа

Существует два простых, но важных способа получения новых путей из старых. А именно, можно определить произведение путей, когда сначала проходит путь γ_1 , а затем – путь γ_2 , если такое возможно. Кроме того, можно определить *обратный* путь γ^{-1} , который проходит в обратном направлении. Сформулируем это в виде следующего утверждения.

Предложение 10.3.1. 1) Если начало второго пути $\gamma_2 = x_2(t)$ совпадает с концом первого пути $\gamma_1 = x_1(t)$, то можно определить произведение путей $\gamma_2 \circ \gamma_1$ следующей формулой:

$$(\gamma_2 \circ \gamma_1) = x(t) := \begin{cases} x_1(2t) & \text{при } 0 < t < \frac{1}{2}, \\ x_2(2t - 1) & \text{при } \frac{1}{2} < t < 1, \end{cases} \quad (10.4)$$

которое является путем в $\mathbb{M}$.

2) Для любого пути γ можно определить путь γ^{-1} , получаемый как путь γ , проходимый в обратном направлении:

$$\gamma^{-1} := x(1 - t). \quad (10.5)$$

Эти операции похожи, но не являются групповыми операциями на множестве всех путей. Действительно, умножение путей определено только в том случае, если конец первого пути совпадает с началом второго. Кроме того, произведения $\gamma^{-1} \circ \gamma$ и $\gamma \circ \gamma^{-1}$ дают не тождественные, а замкнутые пути в начале и конце в точках $x(0)$ и $x(1)$, соответственно.

Разобьем все множество путей γ на классы эквивалентных путей $[\gamma]$, которые гомотопны относительно начала и конца пути. То есть каждый элемент $[\gamma]$ содержит все пути, которые можно непрерывно деформировать друг в друга при фиксированном начале и конце путей: $\gamma_1 \in [\gamma]$ тогда и только тогда, когда $\gamma_1 \sim \gamma(\text{rel}\{0, 1\})$. В классах эквивалентных путей также можно ввести умножение путей и обратный элемент, выбрав из каждого класса по представителю и воспользовавшись формулами (10.4) и (10.5). Нетрудно показать, что эти операции на множестве относительно гомотопных путей корректно определены, т.к. не зависят от выбора представителя в каждом классе. В этих классах произведения $[\gamma^{-1}] \circ [\gamma] = [e_{x(0)}]$ и $[\gamma] \circ [\gamma^{-1}] = [e_{x(1)}]$ дают классы путей, гомотопных тождественным путям. Однако умножение определено не для всех классов путей.

Ситуацию можно исправить и превратить некоторое подмножество всех путей на $\mathbb{M}$ в группу, если зафиксировать точку многообразия и ограничиться замкнутыми путями, имеющими начало и конец в данной точке.

Зафиксируем произвольную точку многообразия $x_0 \in \mathbb{M}$ и рассмотрим все замкнутые пути, имеющие начало и конец в данной точке. Множество этих путей обозначим через $\Omega(\mathbb{M}, x_0)$. При этом допускается, чтобы пути имели точки самопересечения. Поскольку начала и концы всех путей совпадают, то на этом множестве определена операция умножения (10.4). Для каждого пути определен также обратный путь (10.5). Множество $\Omega(\mathbb{M}, x_0)$ с введенной операцией умножения и обратного элемента группы не образует, потому что произведения двух различных путей на их обратные не совпадают, и, следовательно, невозможно ввести понятие единственного единичного элемента. Эту трудность можно обойти, если разбить множество всех путей $\Omega(\mathbb{M}, x_0)$ на классы путей, гомотопных относительно начала и конца, которые обозначим через $\pi_1(\mathbb{M}, x_0)$. Справедлива следующая

Теорема 10.3.1. *Гомотопические классы путей $\pi_1(\mathbb{M}, x_0)$ образуют группу относительно операции умножения путей, причем обратным элементом является гомотопический класс обратного пути, а единицей – гомотопический класс единичного пути e_{x_0} .*

Доказательство. Сводится к проверке корректности операций и групповых свойств. Детали можно найти, например, в [19], теорема 15.2. $\square$

Таким образом, для определения группы было сделано два шага. Во-первых, чтобы определить операцию умножения на всех элементах, множество всех путей было ограничено только замкнутыми путями, имеющими начало и конец в фиксированной точке многообразия. Во-вторых, для корректного определения единичного элемента мы перешли от умножения отдельных путей к умножению классов относительно гомотопных путей.

Определение. Группа $\pi_1(\mathbb{M}, x_0)$ называется *фундаментальной группой* топологического пространства или многообразия $\mathbb{M}$ в точке x_0 . $\square$

В этом определении используется фиксированная точка x_0 . Для связных многообразий роль фиксированной точки не является существенной.

Теорема 10.3.2. Пусть $\mathbb{M}$ – линейно связное топологическое пространство (многообразие). Тогда для любых точек $x_0, x_1 \in \mathbb{M}$ фундаментальные группы $\pi_1(\mathbb{M}, x_0)$ и $\pi_1(\mathbb{M}, x_1)$ изоморфны.

Доказательство. Пусть δ – путь из x_0 в x_1 и γ – произвольный замкнутый путь в точке x_0 . Тогда $\delta \circ \gamma \circ \delta^{-1}$ – замкнутый путь в точке x_1 . Поэтому между множествами замкнутых путей $\Omega(\mathbb{M}, x_0)$ и $\Omega(\mathbb{M}, x_1)$ устанавливается взаимно однозначное соответствие. Отсюда следует, что между фундаментальными группами существует биекция $f_\delta : \pi_1(\mathbb{M}, x_0) \rightarrow \pi_1(\mathbb{M}, x_1)$. Эта биекция – гомоморфизм групп, т.к.

$$f_\delta([\gamma_2] \circ [\gamma_1]) = [\delta] \circ [\gamma_2] \circ [\gamma_1] \circ \delta^{-1} = [\delta] \circ [\gamma_2] \circ \delta^{-1} \circ \delta \circ [\gamma_1] \circ \delta^{-1} = f_\delta([\gamma_2]) \circ f_\delta([\gamma_1]).$$

Обратное отображение f_δ^{-1} определяется обратным путем δ^{-1} . $\square$

В силу этой теоремы при обозначении фундаментальной группы фиксированную точку x_0 мы часто будем опускать: $\pi_1(\mathbb{M}, x_0) = \pi_1(\mathbb{M})$. Изоморфизм между фундаментальными группами f_δ определяется гомотопическим классом пути $[\delta]$ из x_0 в x_1 . Поскольку нет естественного (предпочтительного) выбора такого класса путей, то естественного изоморфизма между фундаментальными группами в различных точках не существует.

Пример 10.3.1. В евклидовом пространстве $\mathbb{R}^n$ при любом n произвольный замкнутый путь для любой отмеченной точки гомотопен тождественному пути в этой точке. Поэтому фундаментальная группа евклидова пространства тривиальна $\pi_1(\mathbb{R}^n) = e$. $\square$

Определение. Топологическое пространство (многообразие) $\mathbb{M}$ называется *односвязным*, если его фундаментальная группа тривиальна, $\pi_1(\mathbb{M}) = e$. $\square$

Предложение 10.3.2. Если $\mathbb{M}_1$ и $\mathbb{M}_2$ – односвязные многообразия, то их прямое произведение $\mathbb{M}_1 \times \mathbb{M}_2$ односвязно.

Доказательство. Простая проверка. $\square$

Пример 10.3.2. Фундаментальная группа окружности с отмеченной точкой изоморфна группе целых чисел по сложению, $\pi_1(\mathbb{S}^1, x_0) \simeq \mathbb{Z}$. При этом целое число равно количеству полных обходов окружности, причем положительные и отрицательные числа соответствуют обходам против и по часовой стрелке, соответственно. Хотя изоморфизм фундаментальной группы окружности группе целых чисел интуитивно очевиден, доказательство громоздко. Его можно найти, например, в [19]. $\square$

Пример 10.3.3. Фундаментальная группа цилиндра совпадает с фундаментальной группой окружности и изоморфна группе целых чисел по сложению $\mathbb{Z}$. $\square$

В общем случае вычисление фундаментальной группы является сложной задачей.

Предложение 10.3.3. Фундаментальная группа группы Ли $U(n)$ при $n \geq 1$ совпадает с группой целых чисел по сложению $\mathbb{Z}$. Группа специальных унитарных матриц $SU(n)$ при $n \geq 1$ односвязна. Фундаментальная группа группы вращений $SO(n)$ при $n \geq 3$ равна $\mathbb{Z}_2$. Фундаментальная группа группы вращений плоскости $SO(2)$ равна $\mathbb{Z}$.

Доказательство. См., например, [67], лекции 12 и 13. $\square$

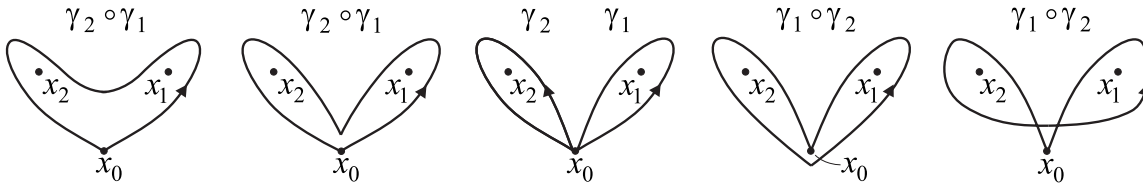


Рис. 10.4: Замкнутые пути γ_1 и γ_2 вокруг точек x_1 и x_2 на евклидовой плоскости (в центре). Непрерывная деформация произведения путей $[\gamma_2] \circ [\gamma_1]$ (влево) и $[\gamma_1] \circ [\gamma_2]$ (вправо).

Отметим, что фундаментальная группа может быть неабелевой.

Пример 10.3.4. Рассмотрим евклидову плоскость $\mathbb{R}^2$ с двумя выколотыми точками x_1 и x_2 и фиксированной точкой x_0 (см. рис.10.4). Пусть γ_1 и γ_2 – замкнутые пути с началом и концом в точке x_0 , которые охватывают, соответственно, точки x_1 и x_2 . На рис.10.4 показаны представители из классов гомотопных путей $[\gamma_2] \circ [\gamma_1]$ и $[\gamma_1] \circ [\gamma_2]$. Ясно, что эти пути не могут быть непрерывно деформированы друг в друга без пересечения выколотых точек и поэтому негомотопны друг другу. Следовательно, фундаментальная группа евклидовой плоскости с двумя выколотыми точками некоммукативна. Плоскость с двумя выколотыми точками гомотопически эквивалентна букету окружностей (восьмерке) (см. пример 10.2.6). Поэтому фундаментальная группа восьмерки также некоммукативна. $\square$

В общем случае фундаментальная группа многообразия представляет собой свободную группу с конечным числом образующих. Напомним

Определение. Группа $\mathbb{G}$ называется *свободной группой* с системой $\mathbb{E} \subset \mathbb{G}$ порождающих (образующих) элементов, если любое отображение множества $\mathbb{E}$ в любую группу $\mathbb{H}$ продолжается до гомоморфизма $\mathbb{G} \rightarrow \mathbb{H}$. Система $\mathbb{E} \subset \mathbb{G}$ называется *системой свободных порождающих*. Ее мощность называется *рангом свободной группы* $\mathbb{G}$. Множество $\mathbb{E}$ называется также *алфавитом*. Элементы $g \in \mathbb{G}$ представляют собой *слова* в алфавите $\mathbb{E}$, т.е. выражения в виде конечного произведения

$$g = e_{i_1}^{\epsilon_1} e_{i_2}^{\epsilon_2} \dots e_{i_k}^{\epsilon_k}, \quad e_{i_j} \in \mathbb{E}, \quad \epsilon_j = \pm 1, \quad \forall j = 1, \dots, k, \quad (10.6)$$

а также пустое слово. Слово называется *несократимым*, если

$$e_{i_j}^{\epsilon_j} \neq e_{i_{j+1}}^{-\epsilon_{j+1}}, \quad \forall j = 1, \dots, k-1.$$

Несократимые слова являются разными элементами свободной группы $\mathbb{G}$, и каждое слово равно единственному несократимому слову. Число k называется *длиной слова* g , если оно несократимо. $\square$

Пример 10.3.5. Рассмотрим расширенную комплексную плоскость $\overline{\mathbb{C}}$ (сферу Римана).

Определение. Область $\mathbb{D} \subset \overline{\mathbb{C}}$ называется *t связной*, если ее граница состоит из t связных компонент. $\square$

Любая t связная область диффеоморфна комплексной плоскости $\mathbb{C}$ с $t-1$ дырками.

Предложение 10.3.4. *Фундаментальная группа произвольной m связной области в расширенной комплексной плоскости $\overline{\mathbb{C}}$ является свободная группа с $m - 1$ образующими.*

Доказательство. Каждой связной компоненте границы соответствует ровно одна компонента дополнения $\overline{\mathbb{C}} \setminus \mathbb{D}$. Обозначим эти компоненты $\mathbb{B}_0, \mathbb{B}_1, \dots, \mathbb{B}_{m-1}$ (см. рис.10.5). Заклучим каждую компоненту $\mathbb{B}_i$ в некоторую окрестность $\mathbb{D}_i$, которая ограничена

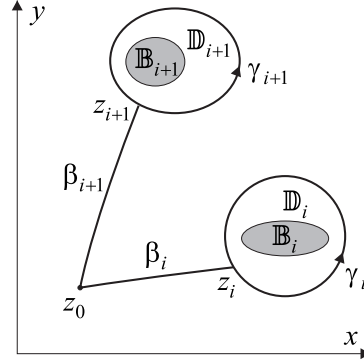


Рис. 10.5: Образующие m связной области в расширенной комплексной плоскости.

простой замкнутой кривой γ_i , целиком лежащей в $\mathbb{D}$. Области $\mathbb{D}_i$ выберем такими, чтобы кривые γ_i не пересекались. На каждой кривой выберем некоторую точку z_i , которую будем считать началом и концом кривой γ_i . Из фиксированной точки z_0 проведем в каждую точку z_i кривую β_i , целиком лежащую в области $\mathbb{D}$. Обозначим

$$e_i := \beta_i^{-1} \circ \gamma_i \circ \beta_i, \quad i = 0, 1, \dots, m - 1.$$

Заметим, что справедливо тождество

$$[e_0] \circ [e_1] \circ \dots \circ [e_{m-1}] = 1.$$

Поэтому гомотопические классы $[e_1], \dots, [e_{m-1}]$ можно выбрать в качестве образующих фундаментальной группы $\pi_1(\mathbb{D}, z_0)$. Можно проверить, что любой путь γ гомотопен конечному произведению:

$$[\gamma] = [\gamma_{i_1}^{\epsilon_1}] \circ [\gamma_{i_2}^{\epsilon_2}] \circ \dots \circ [\gamma_{i_k}^{\epsilon_k}],$$

где все $i_j, j = 1, \dots, k$, принадлежат множеству $\{1, \dots, m - 1\}$. При этом данное представление единственно, если оно несократимо. $\square$ $\square$

Нетрудно доказать, что при непрерывном отображении $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ гомотопически эквивалентные пути $\gamma_1 \sim \gamma_2$ переходят в гомотопически эквивалентные пути $f\gamma_1 \sim f\gamma_2$ и что замкнутые пути переходят в замкнутые. При этом между фундаментальными группами возникает связь.

Теорема 10.3.3. *При непрерывном отображении $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ фундаментальные группы испытывают гомоморфизм*

$$f_* : \pi_1(\mathbb{M}_1, x) \rightarrow \pi_1(\mathbb{M}_2, f(x)),$$

где x – произвольная точка из $\mathbb{M}_1$, который называется индуцированным и одинаков для всех отображений, гомотопных относительно точки x .

Доказательство. См., например, [19], лемма 15.7. $\square$

Следствие. Если отображение двух пространств является гомеоморфизмом, то гомоморфизм фундаментальных групп становится изоморфизмом. $\square$

Понятие фундаментальной группы позволяет осуществить переход от топологии пространств к алгебре и является одним из разделов алгебраической топологии. Этот переход позволяет использовать алгебраические методы при изучении топологических пространств и, следовательно, многообразий. При этом 1) каждому топологическому пространству с отмеченной точкой ставится в соответствие фундаментальная группа, 2) каждому непрерывному отображению пространств ставится в соответствие гомоморфизм групп, 3) тождественному отображению отвечает тождественный гомоморфизм, 4) гомеоморфизму отвечает изоморфизм, 5) композиции непрерывных отображений сопоставляется композиция гомоморфизмов.

Замечание. Отмеченные выше свойства 1) – 5) дают пример функтора из категории топологических пространств с отмеченной точкой и непрерывными отображениями, сохраняющими отмеченную точку, в категорию групп и их гомоморфизмов. $\square$

Рассмотренные примеры окружности и цилиндра показывают, что изоморфизм фундаментальных групп не означает гомеоморфизма пространств. Однако, если фундаментальные группы не изоморфны, то соответствующие пространства заведомо не гомеоморфны.

Теорема 10.3.4. Пусть $f : M_1 \rightarrow M_2$ – гомотопическая эквивалентность двух пространств, тогда индуцированное отображение $f_* : \pi_1(M_1, x) \rightarrow \pi_1(M_2, f(x))$ является изоморфизмом для любой точки $x \in M_1$.

Доказательство. См., например, [19], теорема 15.13. $\square$

Следствие. Стягиваемое пространство имеет тривиальную фундаментальную группу. $\square$

Мы видим, что любое стягиваемое пространство является односвязным. Обратное утверждение неверно.

Пример 10.3.6. Сфера S^n , $n \geq 1$, является односвязным, но не стягиваемым многообразием. $\square$

Пример 10.3.7. Шар B^n , $n \geq 2$, с выколотой точкой является односвязным, но не стягиваемым многообразием. $\square$

Теорема 10.3.5. Пусть M_1 и M_2 – два линейно связных топологических пространства. Тогда фундаментальная группа произведения $M_1 \times M_2$ изоморфна прямому произведению фундаментальных групп, $\pi_1(M_1 \times M_2) \simeq \pi_1(M_1) \times \pi_1(M_2)$.

Доказательство. Доказательство заключается в явном построении изоморфизма. См., например, [19], теорема 15.17. $\square$

Следствие. Произведение $M_1 \times M_2$ двух односвязных пространств является односвязным. $\square$

Пример 10.3.8. Фундаментальная группа тора $T^n \approx S^1 \times \dots \times S^1$ изоморфна прямому произведению групп целых чисел, $\pi_1(T^n) \simeq \underbrace{\mathbb{Z} \times \dots \times \mathbb{Z}}_n$. $\square$

10.4 Фундаментальная группа и ориентируемость

Все связные многообразия можно разделить на два класса: ориентируемые и неориентируемые. Напомним, что многообразие M называется ориентируемым (см. раздел 2), если его можно покрыть системой карт $M = \bigcup_i U_i$, причем якобиан преобразования координат (1.69) во всех пересечениях $U_i \cap U_j$ положителен.

Рассмотрим класс многообразий, на которых можно задать репер $\{e_a\}$, $a = 1, \dots, n$, т.е. n достаточно гладких векторных полей, которые линейно независимы в каждой точке многообразия M , $\dim M = n$. Отсюда, в частности, следует, что векторные поля $e_a(x)$ не могут обращаться в нуль ни в какой точке многообразия M , т.к. в противном случае они были бы линейно зависимы в этой точке. Репер $\{e_a\}$ образует базис для касательных векторных полей. Локально репер всегда существует. Однако его глобальное существование является сильным предположением, и далеко не каждое ориентируемое многообразие допускает существование репера.

Пример 10.4.1. Двумерная сфера S^2 является ориентируемым многообразием, однако она не допускает существование репера, т.к. на ней не существует векторных полей, которые не обращаются в нуль (теорема о невозможности причесать ежа 10.2.1). $\square$

В локальной системе координат репер раскладывается по координатному базису, $e_a = e^\alpha_a \partial_\alpha$, $\det e^\alpha_a \neq 0$.

Введенное ниже определение класса ориентации репера основано на локальном существовании репера и поэтому применимо ко всем многообразиям.

Если от отдельных реперов перейти к классам ориентирующих реперов, то можно дать новое эквивалентное определение ориентируемости многообразий. А именно, пусть в точке $x \in M$ задан *класс ориентации*, т.е. класс реперов $\{e_a\}$, связанных друг с другом линейным преобразованием с положительным определителем. Подробнее, два репера $\{e'_a\}$ и $\{e_a\}$ принадлежат одному классу ориентации, если они связаны преобразованием $e'_a = S_a^b e_b$, причем $\det S(x) > 0$. Если $\det S < 0$, то будем говорить, что в точке x реперы принадлежат классам противоположной ориентации. Таким образом, в каждой точке многообразия имеются ровно два класса ориентации. Поскольку репер можно непрерывно смещать из точки x в близкие точки многообразия, то имеет смысл говорить о непрерывной зависимости классов ориентации от точки многообразия.

Определение. Если в каждой точке многообразия M существует класс ориентации реперов, который непрерывно зависит от точки многообразия, то многообразие M называется *ориентируемым*. $\square$

Класс ориентации принимает только два значения, например, $+1$ или -1 . Поэтому непрерывность означает, что во всех точках ориентируемого многообразия можно выбрать класс ориентации $+1$.

Докажем эквивалентность нового определения и определения ориентируемости многообразий, которое было дано в разделе 2. Пусть многообразие ориентируемо. Тогда на нем существует координатное покрытие $M = \bigcup_i U_i$ такое, что якобиан преобразования координат положителен, $\det \partial_\alpha x^{\alpha'} > 0$, во всех областях пересечения карт. В каждой координатной окрестности выберем координатный репер $\{\partial_\alpha\}$. Множество этих реперов определяет класс ориентации, который непрерывно зависит от точки многообразия. Действительно, если точка принадлежит различным картам, то

координатные реперы принадлежат одному классу ориентации. Обратно. Пусть существует класс ориентации реперов, непрерывно зависящий от точки многообразия. Выберем координатное покрытие многообразия, $\mathbb{M} = \bigcup_i \mathbb{U}_i$. Тогда в двух пересекающихся картах $\mathbb{U}_i$ и $\mathbb{U}_j$ можно выбрать координаты x^α и $x^{\alpha'}$ таким образом, что

$$e_a = e^\alpha_a \partial_\alpha = e^{\alpha'}_a \partial_{\alpha'}, \quad \det e^\alpha_a > 0, \quad \det e^{\alpha'}_a > 0,$$

где $\{e_a\}$ – некоторый представитель заданного класса ориентации. Действительно, если в какой-то системе координат $\det e^\alpha_a < 0$, то, переставив две произвольные координаты, получим $\det e^\alpha_a > 0$. Отсюда следует положительность якобиана преобразования координат в области пересечения карт,

$$J := \det \partial_\alpha x^{\alpha'} = \det (e^\alpha_a e^{\alpha'}_a) > 0. \quad (10.7)$$

Это справедливо для любого представителя из класса ориентации.

На произвольном многообразии, в том числе и неориентируемом, ориентацию можно переносить вдоль путей. Пусть на многообразии $\mathbb{M}$ задан кусочно дифференцируемый путь $\gamma = x(t)$, $t \in [0, 1]$, и класс ориентации реперов $\{e_a(x(t))\}$ во всех точках пути, который непрерывно зависит от точки пути. Тогда класс ориентации $\{e_a(1)\}$ в конечной точке $x(1)$ называется *переносом класса ориентации* репера $\{e_a(0)\}$ из точки $x(0)$ в точку $x(1)$ вдоль пути γ .

Операция переноса ориентации вдоль путей обладает следующими свойствами.

1) Из любой точки $x \in \mathbb{M}$ ориентацию можно однозначно перенести во все близлежащие точки многообразия вдоль путей, целиком лежащих в пределах координатной окрестности точки x .

2) Для любого кусочно дифференцируемого пути перенос ориентации существует и не зависит от выбора репера вдоль пути. Существование очевидно, т.к. достаточно задать произвольным образом компоненты n линейно независимых векторов $\{e_a(t)\}$, как дифференцируемые функции одного переменного. Независимость от выбора репера доказывается просто. Пусть $\{e_a(t)\}$ и $\{e'_a(t)\}$ – два репера вдоль одной кривой $x(t)$, имеющие одинаковую ориентацию при $t = 0$. Матрица перехода от одного репера к другому, $e'_a = S_a^b e_b$, невырождена, $\det S(t) \neq 0$ для всех $t \in [0, 1]$. Поэтому, если в начальный момент $\det S(0) > 0$, то из непрерывности следует, что $\det S(t) > 0$ для всех $t \in [0, 1]$, поскольку матрица S не может быть вырождена.

3) Если два пути $x_1(t)$ и $x_2(t)$ гомотопны относительно начала и конца, то переносы класса ориентации вдоль этих путей совпадают. Действительно, для гомотопных путей существует непрерывное отображение $F(t, s) : [0, 1] \times [0, 1] \rightarrow \mathbb{M}$ такое, что $F(t, 0) = x_1(t)$ и $F(t, 1) = x_2(t)$. Пусть $\{e_a(t)\}$ – репер вдоль $x_1(t)$. Перенесем каким-либо непрерывным образом этот репер из каждой точки кривой γ_1 вдоль кривых $F(t, s)$ по второму аргументу s . Перенос ориентации в конечные точки не зависит от выбора репера $\{e_a(t)\}$ вдоль γ_1 в силу свойства 2). В результате получим некоторый репер вдоль второй кривой γ_2 . Поскольку в начальной и конечной точке реперы по построению совпадают, то совпадают и переносы классов ориентации вдоль гомотопных путей.

Приведенные выше определения удобны для доказательства неориентируемости некоторых многообразий в силу следующей теоремы.

Теорема 10.4.1. *Связное многообразие ориентируемо тогда и только тогда, когда перенос класса ориентации репера вдоль любого замкнутого пути сохраняет ориентацию.*

Доказательство. Пусть многообразие $\mathbb{M}$ ориентируемо. Выберем координатное покрытие такое, чтобы якобиан преобразования координат был положителен, $\det \partial_\alpha x^{\alpha'} > 0$, во всех областях пересечения карт. Выберем произвольную точку $x_0 \in \mathbb{M}$ и зафиксируем какой-либо репер, который принадлежал бы тому же классу ориентации, что и координатный базис, т.е. $\det(e^\alpha_a(x_0)) > 0$. Эту ориентацию можно перенести в любую точку многообразия, т.к. оно линейно связно. Результат переноса ориентации не зависит от пути, потому что ориентацию можно переносить внутри произвольной карты, а при переходе от карты к карте принадлежность репера определенному классу ориентации сохраняется. В частности, принадлежность репера данному классу ориентации сохраняется при переносе ориентации вдоль произвольного замкнутого пути.

Обратно. Выберем произвольным образом репер в какой-либо точке $x_0 \in \mathbb{M}$ и разнесем эту ориентацию во все другие точки $x \in \mathbb{M}$. Результат будет однозначен, т.к. перенос класса ориентации вдоль произвольного замкнутого пути сохраняет ориентацию. Таким образом в каждой точке $x \in \mathbb{M}$ определен класс ориентации $\{e_a(x)\}$. Пусть U_i – координатное покрытие многообразия $\mathbb{M}$. На каждой координатной окрестности можно выбрать систему координат такую, что $\det e^\alpha_a > 0$ для некоторого представителя из класса ориентации. Следовательно, из (10.7) следует, что якобиан преобразования координат во всех пересекающихся областях будет положителен. $\square$

Из этой теоремы вытекает, что если можно указать замкнутый путь, вдоль которого не существует переноса класса ориентации репера, то многообразие является неориентируемым. В частности, неориентируемое многообразие не может быть односвязным. Действительно, любой замкнутый путь на односвязном многообразии можно непрерывно стянуть в точку. Поскольку замкнутые пути в достаточно малой окрестности произвольной точки сохраняют ориентацию, то из непрерывности следует, что на односвязном многообразии любой замкнутый путь сохраняет ориентацию. То есть каждая односвязная компонента многообразия всегда ориентируема. Неодносвязные многообразия могут быть как ориентируемыми, так и неориентируемыми.

Пример 10.4.2. Цилиндр и лист Мебиуса имеют изоморфные фундаментальные группы $\pi \simeq \mathbb{Z}$, однако являются, соответственно, ориентируемой и неориентируемой поверхностями. $\square$

Пусть на многообразии $\mathbb{M}$ задан ориентирующий репер $\{e_a\}$ глобально. Рассмотрим этот репер в качестве базиса касательных векторных полей $X = X^a e_a \in \mathcal{X}(\mathbb{M})$. Зададим на $\mathbb{M}$ тривиальную линейную $\text{GL}(n, \mathbb{R})$ связность: $\omega_{\alpha a}^b = 0$ в каждой карте. Это значит, что при параллельном переносе векторов их компоненты относительно выбранного репера не меняются. Тогда рассматриваемое многообразие будет пространством абсолютного параллелизма, т.к. тензор кривизны для тривиальной связности равен нулю, $R_{\alpha\beta a}^b = 0$. По этой причине, если на многообразии можно задать репер глобально, оно называется *параллелизуемым*. В то же время тензор кручения для тривиальной связности в общем случае будет отличен от нуля. При этом тензор кручения на односвязных многообразиях для заданной нулевой связности однозначно определяется репером. Верно также и обратное утверждение: в односвязном пространстве абсолютного параллелизма, в котором тензор кривизны всюду равен нулю, $R_{\alpha\beta a}^b = 0$, всегда можно задать репер глобально. Для этого достаточно выбрать n линейно независимых векторов в касательном пространстве произвольной точки, а

затем разнести их по всему многообразию, используя то свойство, что результат не зависит от пути параллельного переноса.

Очевидно, что любое параллелизуемое многообразие является ориентируемым.

Пример 10.4.3. Любая группа Ли является параллелизуемым многообразием, если групповое умножение слева (справа) принять за определение параллельного переноса (см. раздел 8.6). Поэтому все группы Ли представляют собой ориентируемые многообразия. При этом левоинвариантные (правоинвариантные) векторные поля представляют собой глобально определенный репер. $\square$

Обратное утверждение, что все ориентируемые многообразия являются параллелизуемыми, неверно, как показывает пример сферы 10.4.1.

Приведем критерий параллелизуемости многообразий.

Теорема 10.4.2. Многообразие M является параллелизуемым тогда и только тогда, когда касательное расслоение тривиально, т.е. диффеоморфно прямому произведению $T(M) \approx M \times \mathbb{R}^n$.

Доказательство. Проверяется тот факт, что глобально определенный репер устанавливает диффеоморфизм расслоений. $\square$

Ориентируемость многообразий можно также определить с помощью дифференциальных форм, рассмотренных в разделе 3. Рассмотрим связное многообразие M , $\dim M = n$. Пространство дифференциальных форм максимальной степени $\Lambda_n(M)$ в каждой точке многообразия одномерно. Пусть $A, B \in \Lambda_n(M)$ и $A \neq 0$, $B \neq 0$ во всех точках $x \in M$. Отношение $A = f(x)B$, где $f \in C(M)$ и $f > 0$, задает отношение эквивалентности в пространстве n -форм, не обращающихся в нуль ни в одной точке многообразия. Это отношение определяет ровно два класса эквивалентности. Назовем *ориентацией* многообразия M соответствующий класс эквивалентности n -форм, если такой класс существует. Это определение ориентируемости эквивалентно приведенному выше в силу следующей теоремы.

Теорема 10.4.3. Связное многообразие M , $\dim M = n$, ориентируемо тогда и только тогда, когда оно допускает непрерывную нигде не обращающуюся в нуль n -форму.

Доказательство. Достаточность. Пусть на связном многообразии M ориентация задана с помощью n -формы A , которая не равна нулю ни в одной точке. Пусть $\{U_i\}$ – локально конечное покрытие многообразия M . Пусть x^α , $\alpha = 1, \dots, n$ – система координат на некоторой окрестности U_i . Тогда n -форма $dx^1 \wedge \dots \wedge dx^n$ не равна нулю ни в какой точке U_i и поэтому отличается от A на некоторый отличный от нуля множитель $f(x)$:

$$dx^1 \wedge \dots \wedge dx^n = fA.$$

Если $f > 0$, то мы говорим, что система координат x^α согласована с ориентацией на M . В противном случае, при $f < 0$, достаточно переставить две координаты для изменения знака n -формы. Обозначим координаты на области U_j , которая пересекается с U_i , через $x^{\alpha'}$. Эти координаты всегда можно выбрать так, чтобы они были согласованы с ориентацией A . Тогда n -форма $dx^{1'} \wedge \dots \wedge dx^{n'}$ также отличается от формы A на некоторый положительный множитель $g(x) > 0$,

$$dx^{1'} \wedge \dots \wedge dx^{n'} = gA.$$

Поскольку

$$dx^{1'} \wedge \dots \wedge dx^{n'} = \det \left(\frac{\partial x^{\alpha'}}{\partial x^{\alpha}} \right) dx^1 \wedge \dots \wedge dx^n,$$

то якобиан соответствующего преобразования координат $x^{\alpha} \rightarrow x^{\alpha'}$ положителен,

$$\det \left(\frac{\partial x^{\alpha'}}{\partial x^{\alpha}} \right) = gf^{-1} > 0.$$

Полученное выражение для якобиана преобразования координат корректно, что легко проверяется для пересечения трех карт.

Необходимость. Пусть f_i – разбиение единицы, подчиненное заданному счетному локально конечному покрытию $\{\mathbb{U}_i\}$ (теорема 2.2.3). Выберем координаты на каждой области $\mathbb{U}_i$ таким образом, чтобы все якобианы преобразований координат в областях пересечения были положительны. На каждой области $\mathbb{U}_i$ можно построить отличную от нуля n -форму $\omega_i = dx^1 \wedge \dots \wedge dx^n$. Тогда n -форма $\sum_i f_i \omega_i$ определена на всем многообразии и отлична от нуля во всех точках. $\square$

Замечание. В доказательстве теоремы использовано разбиение единицы, которое всегда существует, т.к. в определение многообразия мы включили счетность базы топологии. Если это требование исключено из определения, как это часто делается, то в условии теоремы необходимо дополнительно потребовать, например, паракомпактность многообразия. $\square$

В роли n -формы, которая не обращается в нуль, может выступать форма объема (3.79). Если многообразие не допускает существования непрерывной невырожденной n -формы, то оно будет неориентируемым. Если на связном многообразии задана невырожденная n -форма, то мы говорим, что на многообразии задана ориентация. Если многообразие несвязно и состоит из нескольких компонент, то на каждой компоненте ориентацию можно задавать независимо, если каждая компонента связности ориентируема.

Отметим, что каждый репер e_a , заданный глобально, определяет невырожденную n -форму. Действительно, обозначим дуальный базис кокасательного пространства через e^a . Тогда n -форма

$$v := e^1 \wedge \dots \wedge e^n$$

определена глобально и невырождена. Ее можно выбрать в качестве формы объема. Обратное утверждение неверно: не каждая форма объема определяет репер.

Пример 10.4.4. На двумерной сфере $\mathbb{S}^2$ существует форма объема, однако репер определить глобально нельзя, т.к. на ней не существует векторных полей, не обращающихся в нуль. $\square$

Обсудим более подробно связь неориентируемости многообразия с фундаментальной группой. Перенос ориентации с помощью репера показывает, что каждый гомотопический класс замкнутого пути с началом и концом в точке x_0 (т.е. элемент фундаментальной группы $\pi_1(\mathbb{M}, x_0)$) либо сохраняет, либо меняет ориентацию репера на противоположную. Это означает, что существует гомоморфизм σ фундаментальной группы в группу $\mathbb{Z}_2$,

$$\sigma : \pi_1(\mathbb{M}, x_0) \rightarrow \mathbb{Z}_2 = \{\pm 1\}.$$

Другими словами, каждому замкнутому пути ставится в соответствие $+1$ или -1 , в зависимости от того, сохраняется ориентация репера при переносе или меняется на

противоположную. Если многообразие ориентируемо, то гомоморфизм σ тривиален. Для неориентируемых многообразий ввиду наличия путей, обращающих ориентацию, гомоморфизм σ нетривиален. Как следствие получаем, что фундаментальная группа неориентируемых многообразий не может быть тривиальной.

Пример 10.4.5. Фундаментальная группа листа Мебиуса совпадает с группой целых чисел, $\pi_1(\mathbb{M}) = \mathbb{Z}$, т.к. он стягивается к центральной окружности. При гомоморфизме σ все четные и нечетные числа отображаются, соответственно, в $+1$ и -1 . $\square$

Пример 10.4.6. Фундаментальная группа проективного пространства $\mathbb{RP}^n$ нетривиальна: $\pi_1(\mathbb{RP}^n) = \mathbb{Z}_2$, и гомоморфизм σ является изоморфизмом для четных n . Для нечетных n гомоморфизм σ тривиален, т.к. проективные пространства нечетной размерности ориентируемы. Проективное пространство можно параметризовать точками шара, у которого отождествлены диаметрально противоположные точки граничной сферы (см. пример 2.1.8). Тогда диаметр шара будет замкнутым путем (образом окружности), который не стягивается в точку. В то же время, если при отображении окружности в проективное пространство диаметр проходится два раза, то такой образ окружности можно стянуть в точку, как показано на рис. 10.6. Аналогично стягивается в точку любой образ окружности, который проходит диа-

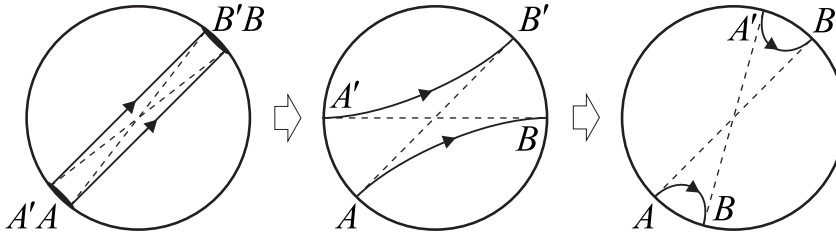


Рис. 10.6: Непрерывная деформация образа окружности в проективном пространстве $\mathbb{RP}^2$, который проходит по диаметру AB два раза, в точку. Пунктиром показано отождествление противоположных точек граничной окружности.

метр произвольное четное число раз. Это поясняет, почему фундаментальная группа проективного пространства равна $\mathbb{Z}_2$. $\square$

Предложение 10.4.1. Любое связное и односвязное многообразие ориентируемо.

Доказательство. См., например, [73], следствие 2.7.6. $\square$

Глава 11

Накрытия

Важным классом непрерывных отображений топологических пространств являются накрытия, которые представляют собой локальный гомеоморфизм. Все, сказанное в настоящей главе о топологических пространствах, относится также к многообразиям.

11.1 Определения и примеры

Определение. Непрерывное отображение топологических пространств $p : \tilde{M} \rightarrow M$ называется *накрытием*, если выполнены следующие условия:

- 1) p сюръективно;
- 2) для любого $x \in M$ найдется открытая окрестность U_x точки x такая, что $p^{-1}(U_x) = \bigcup_{j \in \mathcal{J}} U_j$ для некоторого семейства открытых подмножеств $U_j \subset \tilde{M}$, удовлетворяющих условиям $U_j \cap U_k = \emptyset$ при $j \neq k$ и сужение отображения $p|_{U_j} : U_j \rightarrow U_x$ – гомеоморфизм для всех $j \in \mathcal{J}$.

Мы говорим также, что окрестность U_x *просто накрыта* отображением p . Топологическое пространство M называется *базой* накрытия, а $\tilde{M}$ – *накрывающим пространством*. Каждое множество U_j называется *листом* накрытия области U_x . Если множество $\mathcal{J}$ содержит n элементов, то говорят об n -листном накрытии. Если топологическое пространство $\tilde{M}$ является односвязным, т.е. фундаментальная группа тривиальна, $\pi(\tilde{M}) = e$, то накрытие называется *универсальным*. $\square$

Предложение 11.1.1. Если база M накрытия $p : \tilde{M} \rightarrow M$ линейно связна, то мощность множества $\mathcal{J}$ не зависит от точки $x \in M$.

Доказательство. См., например, [74], глава 2, §3, следствие 8. $\square$

Это утверждение говорит о том, что определение n -листного накрытия корректно, т.к. не зависит от точки $x \in M$.

Замечание. Если $\tilde{M}$ и M являются не просто топологическими пространствами, а многообразиями, то вместо непрерывности отображения p мы требуем его дифференцируемость достаточное число раз, а гомеоморфизм заменяем на диффеоморфизм областей U_j и U_x для всех $j \in \mathcal{J}$. Для многообразий $\dim U_j = \dim U_x$ для всех $j \in \mathcal{J}$, т.к. сужение отображения p на U_j – диффеоморфизм. Кроме того, каждое подмножество U_j , по определению, открыто в $\tilde{M}$. Это возможно только если $\dim \tilde{M} = \dim U_j$. Таким образом, если $p : \tilde{M} \rightarrow M$ накрытие многообразий, то необходимо, чтобы накрывающее многообразие и база имели одинаковую размерность, $\dim \tilde{M} = \dim M$. Для многообразий отображение p является погружением. $\square$

Пример 11.1.1. Всякий гомеоморфизм является однолистным накрытием. В этом случае у каждой области $\mathbb{U}_x$ имеется только один лист. $\square$

Пример 11.1.2. Если $\tilde{\mathbb{M}} = \mathbb{M} \times \mathbb{X}$ – топологическое произведение топологического пространства $\mathbb{M}$ на счетное множество $\mathbb{X}$, то естественная проекция $p : \mathbb{M} \times \mathbb{X} \rightarrow \mathbb{M}$ является накрытием. Число листов равно числу элементов множества $\mathbb{X}$. $\square$

Пример 11.1.3. Пусть задана аналитическая функция $F(z)$ в некоторой области $\mathbb{D} \subset \mathbb{C}$. Если в каждой точке $z \in \mathbb{D}$ функция $F(z)$ имеет N значений, то риманова поверхность функции $F(z)$ является N -листным накрытием области $\mathbb{D}$. $\square$

Пример 11.1.4. Отображение вещественной прямой в окружность,

$$p : \mathbb{R} \ni t \mapsto e^{2\pi it} \in \mathbb{S}^1,$$

является накрытием с бесконечным числом листов. Это накрытие является универсальным. $\square$

Пример 11.1.5. Пусть окружность $\mathbb{S}^1$ задана на комплексной плоскости $\mathbb{C}$ уравнением $|z| = 1$. Тогда отображение окружности в окружность,

$$p : \mathbb{S}^1 \ni z \mapsto z^n \in \mathbb{S}^1, \quad \forall n \in \mathbb{N},$$

является накрытием с n листами. Это же отображение является n -листным накрытием комплексной плоскости с выколотым началом координат $p : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C} \setminus \{0\}$. $\square$

Пример 11.1.6. Пусть $\mathbb{G}$ – группа Ли и $\mathbb{H}$ – ее дискретная подгруппа. Поскольку подгруппа $\mathbb{H}$ замкнута, то по теореме 9.1.1 на фактор пространстве $\mathbb{G}/\mathbb{H}$ существует дифференцируемая структура многообразия. Тогда проекция $p : \mathbb{G} \rightarrow \mathbb{G}/\mathbb{H}$, где $\mathbb{G}/\mathbb{H}$ – пространство левых или правых смежных классов, является накрытием. Число листов равно числу элементов группы $\mathbb{H}$. $\square$

Предложение 11.1.2. Пусть $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$ расслоение, определенное в разделе 2.4, типичным слоем $\mathbb{F}$ которого является 0-мерное многообразие. Тогда проекция $\pi : \mathbb{E} \rightarrow \mathbb{M}$ является накрытием. Наоборот, любое накрытие является расслоением $\mathbb{M}(\mathbb{M}, p, \mathcal{J})$ с 0-мерным типичным слоем $\mathcal{J}$. $\square$

Доказательство. Определение накрытия и расслоения в случае 0-мерного типичного слоя совпадают, если записать $\pi^{-1}(\mathbb{U}_x) = \chi^{-1}(\mathbb{U}_x \times \mathbb{F}) = \bigcup_{j \in \mathcal{J}} \mathbb{U}_j$, где точки 0-мерного многообразия $j \in \mathcal{J} = \mathbb{F}$ нумеруют листы накрытия области $\mathbb{U}_x$. Второе условие в определении расслоения при 0-мерном слое выполняется автоматически, т.к. отображение $\pi \circ \chi^{-1}$ дифференцируемо, а это возможно только при проекции на первый сомножитель, $\pi \circ \chi^{-1} : \mathbb{U}_x \times \mathbb{F} \ni (y, f) \mapsto y \in \mathbb{M}$. $\square$

Замечание. Если типичным слоем расслоения $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$ является многообразие более высокой размерности, $\dim \mathbb{F} \geq 1$, то расслоение не является накрытием, потому что, например, $\dim \pi^{-1}(\mathbb{U}_x) = \dim \mathbb{M} + \dim \mathbb{F} > \dim \mathbb{M}$. $\square$

Отображение накрытия позволяет установить связь между ориентируемыми и неориентируемыми многообразиями.

Предложение 11.1.3. Для любого неориентируемого многообразия $\mathbb{M}$ существует двулистное ориентируемое накрытие $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$.

Доказательство. См., например, [73], предложение 2.7.4. $\square$

Пример 11.1.7. Четномерная сфера $\mathbb{S}^n$, $n = 2, 4, \dots$, ориентируема и является двулистным универсальным накрывающим пространством для проективной плоскости $\mathbb{RP}^n$, которая неориентируема. Накрытие $p : \mathbb{S}^n \rightarrow \mathbb{RP}^n$ осуществляется путем отождествления диаметрально противоположных точек сферы.

Пример 11.1.8. Лист Мебиуса можно получить из цилиндра путем отождествления центрально симметричных точек, как показано на рис. 11.1. Это накрытие двулистно, но не является универсальным. Универсальные накрывающие цилиндра и листа Мебиуса совпадают – это плоскость $\mathbb{R}^2$. $\square$

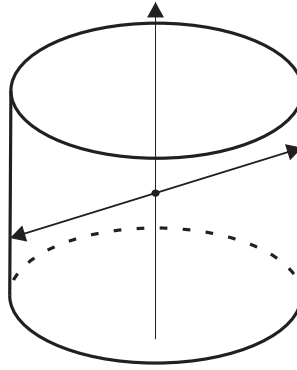


Рис. 11.1: Накрытие листа Мебиуса цилиндром. Отождествив центрально симметричные точки цилиндра, мы получим лист Мебиуса в виде полосы с отождествленными краевыми точками.

Следующее утверждение облегчает изучение накрытий за счет ограничения класса рассматриваемых баз.

Теорема 11.1.1. Если пространство $\mathbb{M}$ – локально связно, то непрерывное отображение $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ тогда и только тогда является накрытием, когда для каждой компоненты связности $U \subset \mathbb{M}$ сужение отображения

$$p|_{p^{-1}(U)} : p^{-1}(U) \rightarrow U$$

является накрытием.

Доказательство. См., например, [74], глава 2, §1, теорема 11. $\square$

Поскольку все многообразия являются локально связными, то в дифференциальной геометрии мы можем ограничиться рассмотрением накрытий над связными базами $\mathbb{M}$. Поэтому в дальнейшем мы будем рассматривать только линейно связные и, следовательно, связные базы накрытий.

Из определения накрытия следует, что прообраз $p^{-1}(U_x)$ является объединением непересекающихся открытых подмножеств U_j , каждое из которых гомеоморфно U_x . Если U_x связно, то каждое из подмножеств U_j также связно. Отсюда следует

Предложение 11.1.4. Пусть U – открытое связное подмножество пространства $\mathbb{M}$, просто накрытое отображением $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$. Тогда p гомеоморфно отображает каждую компоненту связности множества $p^{-1}(U)$ на U .

Следствие. Рассмотрим два накрытия $p_1 : \tilde{M}_1 \rightarrow M$, $p_2 : \tilde{M}_2 \rightarrow M$ с одинаковыми базами и коммутативную диаграмму

$$\begin{array}{ccc} \tilde{M}_1 & \xrightarrow{p} & \tilde{M}_2 \\ & \searrow p_1 & \downarrow p_2 \\ & & M \end{array}$$

где база M локально связна. Если p сюръективно, то оно также является накрытием. $\square$

Теорема 11.1.2. Пусть $p : \tilde{M} \rightarrow M$ – накрытие. Тогда

- 1) p – открытое отображение;
- 2) M имеет фактор топологию относительно p .

Доказательство. См., например, [19], теорема 17.3. $\square$

Если на базе накрытия M задан путь $\gamma : [0, 1] \ni t \mapsto x(t) \in M$, то возникает вопрос, что представляет собой прообраз $p^{-1}(\gamma)$ при накрытии $p : \tilde{M} \rightarrow M$? Для ответа введем новое понятие.

Пусть задано накрытие $p : \tilde{M} \rightarrow M$ и непрерывное отображение некоторого пространства N в базу накрытия, $f : N \rightarrow M$. (Для путей в роли N выступает единичный отрезок $[0, 1]$.)

Определение. Поднятием непрерывного отображения f называется непрерывное отображение $\tilde{f} : N \rightarrow \tilde{M}$ такое, что $p \circ \tilde{f} = f$, т.е. диаграмма

$$\begin{array}{ccc} N & \xrightarrow{\tilde{f}} & \tilde{M} \\ & \searrow p \circ \tilde{f} & \downarrow p \\ & & M \end{array}$$

коммутативна. $\square$

Следующий результат показывает, что, если поднятие отображения f существует, то оно, по существу, единственно.

Предложение 11.1.5. Пусть $p : \tilde{M} \rightarrow M$ – накрытие и $\tilde{f}, \tilde{f}' : N \rightarrow \tilde{M}$ – два поднятия отображения $f : N \rightarrow M$. Если N связно и поднятия совпадают $\tilde{f}(y) = \tilde{f}'(y)$ хотя бы в одной точке $y \in N$, то поднятия совпадают во всех точках, $\tilde{f} = \tilde{f}'$.

Доказательство. См., например, [19], лемма 17.4. $\square$

Теорема 11.1.3 (О накрывающей гомотопии для путей). Пусть $p : \tilde{M} \rightarrow M$ – накрытие. Тогда

- 1) для пути $\gamma : [0, 1] \rightarrow \mathbb{M}$ и точки $\tilde{x} \in \tilde{\mathbb{M}}$ такой, что $p(\tilde{x}) = \gamma(0)$ существует единственный путь $\tilde{\gamma} : [0, 1] \rightarrow \tilde{\mathbb{M}}$, который является поднятием, $p \circ \tilde{\gamma} = \gamma$, пути γ и для которого $\tilde{\gamma}(0) = \tilde{x}$;
- 2) для непрерывного отображения (гомотопии двух путей в базе) $F : [0, 1] \times [0, 1] \rightarrow \mathbb{M}$ и точки $\tilde{x} \in \tilde{\mathbb{M}}$ такой, что $p(\tilde{x}) = F(0, 0)$ существует единственное непрерывное отображение (гомотопия двух путей в накрывающем пространстве) $\tilde{F} : [0, 1] \times [0, 1] \rightarrow \tilde{\mathbb{M}}$, которое является поднятием, $p \circ \tilde{F} = F$, отображения F и для которого $\tilde{F}(0, 0) = \tilde{x}$.

Доказательство. Теорема доказывается путем явного построения отображений. Делать можно найти, например, в [19], теорема 17.6. $\square$

Следствие (Теорема о монодромии). Пусть $\gamma_1 \sim \gamma_2(\text{rel}\{0, 1\})$ – два пути в базе $\mathbb{M}$, гомотопные относительно начала и конца, и начальные точки их поднятий совпадают, $\tilde{\gamma}_1(0) = \tilde{\gamma}_2(0)$. Тогда конечные точки поднятий также совпадают, $\tilde{\gamma}_1(1) = \tilde{\gamma}_2(1)$. $\square$

Напомним, что единица фундаментальной группы $\pi_1(\mathbb{M}, x_0)$ – это класс замкнутых путей с началом и концом в точке x_0 , которые гомотопны тождественному пути e_{x_0} . Из теоремы о монодромии следует, что поднятия всех этих путей также замкнуты в накрывающем пространстве $\tilde{\mathbb{M}}$. Следующая теорема говорит о том, что поднятие замкнутых путей, которые не стягиваются в точку x_0 будет незамкнутым.

Теорема 11.1.4. Пусть $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ – универсальное накрытие, т.е. накрывающее пространство $\tilde{\mathbb{M}}$ – односвязно. Тогда существует взаимно однозначное соответствие между элементами фундаментальной группы $\pi(\mathbb{M}, p(\tilde{x}))$ и множеством прообразов $p^{-1}(p(\tilde{x}))$, где $\tilde{x}$ – произвольная точка $\tilde{\mathbb{M}}$.

Доказательство. Пусть $[\gamma] \in \pi(\mathbb{M}, p(\tilde{x}))$ – класс гомотопных замкнутых путей на базе и $\tilde{\gamma}$ – поднятие какого-либо представителя из данного класса. Определим отображение

$$\varphi : \pi(\mathbb{M}, p(\tilde{x})) \ni [\gamma] \mapsto \varphi([\gamma]) = \tilde{\gamma}(1) \in p^{-1}(p(\tilde{x})).$$

Это отображение определено корректно согласно теореме о монодромии.

Теперь определим обратное отображение $\varphi^{-1} : p^{-1}(p(\tilde{x})) \rightarrow \pi(\mathbb{M}, p(\tilde{x}))$. Выбираем точку $\tilde{x}_1 \in p^{-1}(p(\tilde{x}))$ и некоторый путь $\tilde{\gamma}$ из $\tilde{x}$ в $\tilde{x}_1$. Поскольку $\tilde{\mathbb{M}}$ односвязно, то любые такие пути гомотопны. Поэтому класс замкнутых путей $[p \circ \tilde{\gamma}]$ – корректно определенный элемент фундаментальной группы $\pi(\mathbb{M}, p(\tilde{x}))$. Легко проверить, что $\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = e$, так что φ и φ^{-1} – биекции. $\square$

Эта теорема дает возможность в некоторых случаях вычислить фундаментальную группу. Для этого надо найти универсальное накрытие $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ и групповую структуру на прообразе $p^{-1}(x_0)$, где $x_0 \in \mathbb{M}$, так, чтобы биекция $\varphi : \pi_1(\mathbb{M}, x_0) \rightarrow p^{-1}(x_0)$ стала гомоморфизмом групп. Вообще говоря, сделать это сложно.

11.2 Фундаментальная группа пространства орбит

В настоящем разделе мы рассмотрим группы преобразований $(\tilde{\mathbb{M}}, \mathbb{G})$ и приведем достаточные условия на группу преобразований, которые обеспечивают то, что проекция пространства $\tilde{\mathbb{M}}$ в пространство орбит $\mathbb{M} := \tilde{\mathbb{M}}/\mathbb{G}$ является накрытием. Обсудим также связь группы преобразований $(\tilde{\mathbb{M}}, \mathbb{G})$ с фундаментальными группами базы $\mathbb{M} := \tilde{\mathbb{M}}/\mathbb{G}$ и накрывающего пространства $\tilde{\mathbb{M}}$. Кроме того, обсудим достаточные условия существования в пространстве орбит дифференцируемой структуры.

Замечание. Многие определения и теоремы настоящего раздела справедливы не только для многообразий, но и для произвольных топологических пространств. Поэтому мы будем употреблять термин топологическое пространство, имея в виду, что соответствующее утверждение верно и для многообразий. Там, где важна специфика дифференцируемой структуры, мы будем писать многообразие. $\square$

Определение. Группа $\mathbb{G}$ действует на топологическом пространстве $\tilde{\mathbb{M}}$ *собственно разрывно*, если выполнены следующие условия:

- 1) для любых двух точек $\tilde{x}_1$ и $\tilde{x}_2$ из $\tilde{\mathbb{M}}$, не лежащих на одной орбите, существуют окрестности $\tilde{\mathbb{U}}_1 \supset \tilde{x}_1$ и $\tilde{\mathbb{U}}_2 \supset \tilde{x}_2$ такие, что $\tilde{\mathbb{U}}_1 \mathbb{G} \cap \tilde{\mathbb{U}}_2 = \emptyset$,
- 2) группа изотропии $\mathbb{G}_{\tilde{x}}$ любой точки $\tilde{x} \in \tilde{\mathbb{M}}$ конечна,
- 3) для каждой точки $\tilde{x}$ существует окрестность $\tilde{\mathbb{U}}$, устойчивая относительно группы изотропии $\mathbb{G}_{\tilde{x}}$, т.е. $\tilde{\mathbb{U}} \mathbb{G}_{\tilde{x}} = \tilde{\mathbb{U}}$, и такая, что для всех $a \in \mathbb{G}$, не лежащих в $\mathbb{G}_{\tilde{x}}$, пересечение $\tilde{\mathbb{U}}a \cap \tilde{\mathbb{U}} = \emptyset$.

Ниже мы увидим, что первое условие в данном определении эквивалентно хаусдорфовости фактор пространства $\mathbb{M}$ с фактортопологией.

Предложение 11.2.1. *Если группа $\mathbb{G}$ действует на многообразии $\tilde{\mathbb{M}}$ собственно разрывно, то она конечна или счетна.*

Доказательство. В силу условия 3) орбита произвольной точки является дискретным подмножеством многообразия. Поскольку множество таких точек на многообразии счетно, то, вместе с условием 2) это влечет счетность элементов группы $\mathbb{G}$. В частности, оно может быть конечно. $\square$

Для доказательства следующей теоремы понадобится лемма, которая представляет и самостоятельный интерес.

Лемма 11.2.1. *Пусть $(\tilde{\mathbb{M}}, \mathbb{G})$ – группа преобразований топологического пространства $\tilde{\mathbb{M}}$ и фактор пространство $\mathbb{M} = \tilde{\mathbb{M}}/\mathbb{G}$ снабжено фактор топологией. Тогда каноническая проекция $p: \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ является открытым отображением.*

Доказательство. Пусть $\tilde{\mathbb{U}}$ – открытое подмножество $\tilde{\mathbb{M}}$. Тогда подмножество $p(\tilde{\mathbb{U}})$ открыто в $\mathbb{M}/\mathbb{G}$, т.к. оно снабжено фактор топологией. Прообраз этого подмножества можно представить в виде объединения

$$p^{-1}(p(\tilde{\mathbb{U}})) = \bigcup_{a \in \mathbb{G}} \tilde{\mathbb{U}}a. \quad (11.1)$$

Поскольку отображение $\tilde{\mathbb{U}} \rightarrow \tilde{\mathbb{U}}a$ – гомеоморфизм для всех $a \in \mathbb{G}$, то $p^{-1}(p(\tilde{\mathbb{U}}))$ открыто в $\tilde{\mathbb{M}}$ как объединение открытых подмножеств. $\square$

Теорема 11.2.1. *Пусть группа $\mathbb{G}$ действует в топологическом пространстве $\tilde{\mathbb{M}}$ собственно разрывно и свободно, а пространство орбит $\mathbb{M}/\mathbb{G}$ снабжено фактор топологией. Тогда каноническая проекция $p: \tilde{\mathbb{M}} \rightarrow \mathbb{M}/\mathbb{G}$ является накрытием.*

Доказательство. По построению отображение p сюръективно и непрерывно. По лемме 11.2.1 это отображение открыто. Пусть $\tilde{\mathbb{U}}$ – окрестность точки $\tilde{x} \in \tilde{\mathbb{M}}$, удовлетворяющая условию 3) в определении собственно разрывной группы преобразований. Так как p – открытое отображение, то $p(\tilde{\mathbb{U}})$ – открытая окрестность орбиты $p(\tilde{x}) = \tilde{x}\mathbb{G}$ и ее прообраз имеет вид (11.1). Поскольку действие группы свободно, то все области $\tilde{\mathbb{U}}a$, $a \in \mathbb{G}$, открыты в $\tilde{\mathbb{M}}$ и не пересекаются. Сужение отображения на каждую подобласть $p|_{\tilde{\mathbb{U}}a}: \tilde{\mathbb{U}}a \rightarrow p(\tilde{\mathbb{U}})$ является непрерывным открытым биективным отображением и, следовательно, гомеоморфизмом. $\square$

Теперь обсудим связь между группой преобразований $(\tilde{\mathbb{M}}, \mathbb{G})$, действующей собственноразрывно и свободно, и фундаментальной группой пространства орбит $\mathbb{M} = \tilde{\mathbb{M}}/\mathbb{G}$, на котором определена фактор топология. Пусть $\tilde{x} \in \tilde{\mathbb{M}}$ и $x = p(\tilde{x}) \in \mathbb{M}/\mathbb{G}$. Заметим, что

$$p^{-1}(x) = \{\tilde{x}a \in \tilde{\mathbb{M}} : a \in \mathbb{G}\},$$

т.е. точки орбиты находятся во взаимно однозначном соответствии с элементами группы. Если класс относительно гомотопных путей принадлежит фундаментальной группе пространства орбит, $[\gamma] \in \pi_1(\mathbb{M}/\mathbb{G}, x)$, то существует единственное поднятие $\tilde{\gamma}$ пути γ с началом в точке $\tilde{x} \in \tilde{\mathbb{M}}$ (теорема 11.1.3). Конец поднятого пути $\tilde{x}(1) \in p^{-1}(x)$, и, поскольку группа преобразований действует свободно, то существует единственный элемент $a_{(\gamma)} \in \mathbb{G}$ такой, что $\tilde{x}(1) = \tilde{x}a_{(\gamma)}$. Следовательно, определено отображение

$$\varphi : \pi_1(\tilde{\mathbb{M}}/\mathbb{G}, x) \ni [\gamma] \mapsto a_{(\gamma)} \in \mathbb{G}. \quad (11.2)$$

Теорема 11.2.2. *Отображение (11.2) является гомоморфизмом групп.*

Доказательство. Сводится к проверке равенства $\varphi([\gamma_1][\gamma_2]) = \varphi([\gamma_1])\varphi([\gamma_2])$. Детали можно найти, например, в [19], теорема 19.1. $\square$

Если фундаментальная группа накрывающего пространства нетривиальна, то ядро гомоморфизма также нетривиально.

Предложение 11.2.2. *Ядро гомоморфизма (11.2) совпадает с индуцированной подгруппой $p_*\pi_1(\tilde{\mathbb{M}}, \tilde{x}) \subset \pi_1(\tilde{\mathbb{M}}/\mathbb{G}, x)$.*

Доказательство. См., например, [19], лемма 19.2. $\square$

В частности, $p_*\pi_1(\tilde{\mathbb{M}}, \tilde{x})$ – нормальная подгруппа в $\pi_1(\tilde{\mathbb{M}}/\mathbb{G}, x)$, и поэтому определена фактор группа $\pi_1(\tilde{\mathbb{M}}/\mathbb{G}, x)/p_*\pi_1(\tilde{\mathbb{M}}, \tilde{x})$.

Теорема 11.2.3. *Группы $\pi_1(\tilde{\mathbb{M}}/\mathbb{G}, x)/p_*\pi_1(\tilde{\mathbb{M}}, \tilde{x})$ и $\mathbb{G}$ изоморфны.*

Доказательство. См., например, [19], теорема 19.3. $\square$

Следствие. Если накрывающее топологическое пространство $\tilde{\mathbb{M}}$ односвязно, то $\pi_1(\tilde{\mathbb{M}}/\mathbb{G}, x) \simeq \mathbb{G}$. $\square$

Это следствие позволяет в ряде случаев найти фундаментальную группу многообразия.

Пример 11.2.1. Нетрудно проверить, что окружность $\mathbb{S}^1$ гомеоморфна пространству орбит $\mathbb{R}/\mathbb{Z}$, где группа $\mathbb{Z}$ действует на $\mathbb{R}$ сдвигами на постоянное число. Действие группы $\mathbb{Z}$ на $\mathbb{R}$ свободно и собственноразрывно. Поскольку вещественная прямая $\mathbb{R}$ односвязна, то отображение $p : \mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$ – универсальное накрытие. Следовательно, $\pi_1(\mathbb{S}^1) \simeq \mathbb{Z}$. $\square$

Пример 11.2.2. Тор является факторпространством, $\mathbb{T}^n \approx \mathbb{R}^n/\mathbb{Z}^n$. Поскольку отображение $p : \mathbb{R}^n \rightarrow \mathbb{R}^n/\mathbb{Z}^n$ – универсальное накрытие, то $\pi_1(\mathbb{T}^n) \simeq \mathbb{Z}^n$. $\square$

Следующая теорема дает достаточное условие существования дифференцируемой структуры на пространстве орбит $\mathbb{M}/\mathbb{G}$.

Теорема 11.2.4. Пусть группа преобразований $(\tilde{M}, G)$ действует на многообразии $\tilde{M}$ собственноразрывно и свободно, тогда факторпространство $\tilde{M}/G$ с фактортопологией имеет структуру дифференцируемого многообразия такую, что проекция $p: \tilde{M} \rightarrow \tilde{M}/G$ дифференцируема.

Доказательство. Условие 1) в определении собственноразрывной группы преобразований эквивалентно хаусдорфовости факторпространства $\tilde{M}/G$. Действительно, окрестностью орбиты $x = \tilde{x}G$ является множество орбит $U = \tilde{U}G$, где $\tilde{U}$ – окрестность точки $\tilde{x}$, а условие 1) можно переписать в виде $\tilde{U}_1G \cap \tilde{U}_2G = \emptyset$, т.к. орбиты либо не имеют общих точек, либо совпадают. Это и есть условие хаусдорфовости пространства орбит с фактортопологией.

В условиях теоремы группа G действует свободно и условие 2) в определении собственноразрывной группы преобразований выполняется автоматически, т.к. группа изотропии любой точки состоит только из одного элемента – единицы.

Пусть $U = \tilde{U}G \subset \tilde{M}/G$ – достаточно малая окрестность орбиты $x \in \tilde{M}/G$. Условие 3) вместе с условием 2) значит, что каждая точка x в пространстве орбит имеет окрестность U такую, что прообраз $\tilde{U} = p^{-1}(U)$ состоит не более, чем из счетного числа компонент, $\tilde{U} = \bigcup_i \tilde{U}_i$, и проекция p каждой связной компоненты $\tilde{U}_i$ на U есть гомеоморфизм, т.к. множество орбит можно параметризовать точками из какой-либо окрестности U_i . Зафиксируем связную компоненту $\tilde{U}_i$ в $p^{-1}(U)$. Тогда p – гомеоморфизм $\tilde{U}_i$ на U и существует обратное непрерывное отображение $p^{-1}: U \rightarrow \tilde{U}_i$. Выбрав U достаточно малым, можно считать, что имеется допустимая карта $(\tilde{U}_i, \varphi)$, где $\varphi: \tilde{U}_i \rightarrow \mathbb{R}^n$, многообразия $\tilde{M}$. Теперь можно ввести дифференцируемую структуру в пространстве орбит, рассматривая (U, ψ) , где $\psi = \varphi \circ p^{-1}$, как допустимую карту. $\square$

Последняя теорема является достаточным, но не необходимым условием того, что пространство орбит является многообразием.

Пример 11.2.3. Рассмотрим вращения евклидовой плоскости $\mathbb{R}^2$ вокруг начала координат на фиксированный угол $\alpha = 2\pi/n$, где n – одно из натуральных чисел $2, 3, \dots$. При этом мы отождествляем поворот на 2π с единичным элементом группы. Эта группа G абелева и состоит из n элементов. Действие группы является эффективным. Начало координат является неподвижной точкой и одной из орбит группы. Ее группа изотропии совпадает с G . На остальной части плоскости $\mathbb{R}^2 \setminus \{0\}$ группа G действует свободно. Нетрудно проверить, что группа G действует на плоскость собственноразрывно. Пространство орбит $\mathbb{R}^2/G$ представляет собой конус с углом дефицита $2\pi/n - 2\pi$ (знак минус означает, что угол вырезается, а не вставляется). Конус является гладким двумерным многообразием, т.е. пространство орбит допускает гладкую структуру, несмотря на то, что действие группы не является свободным. Вершина конуса – это “дефект” не многообразия, а вложения. В этом примере каноническая проекция $p: \mathbb{R}^2 \rightarrow \mathbb{R}^2/G$ не является накрытием, т.к. прообраз $p^{-1}(U_0)$ окрестности $U_0 \subset \mathbb{R}^2/G$, содержащей начало координат, состоит из одного связного листа, который не гомеоморфен самой окрестности (нет взаимной однозначности). Фундаментальные группы плоскости $\mathbb{R}^2$ и конуса $\mathbb{R}^2/G$ тривиальны и совпадают. $\square$

Если пара $(\tilde{M}, G)$ – группа преобразований, причем группа G действует на многообразии $\tilde{M}$ свободно и собственной разрывно, то одним из способов изучения факторпространства $\tilde{M}/G$ является построение фундаментальной области.

Определение. Подмножество $\mathbb{D} \subset \tilde{\mathbb{M}}$ называется *фундаментальной областью* многообразия $\tilde{\mathbb{M}}$ для группы преобразований $(\tilde{\mathbb{M}}, \mathbb{G})$, действующей свободно и собственноразрывно, если выполнены следующие условия:

- 1) $\mathbb{D}$ является замкнутым подмножеством в $\tilde{\mathbb{M}}$;
- 2) орбита $\mathbb{D}\mathbb{G}$ совпадает со всем многообразием $\tilde{\mathbb{M}}$;
- 3) покрытие $\tilde{\mathbb{M}}$ множествами $\mathbb{D}a$, $a \in \mathbb{G}$, таково, что с достаточно малой окрестностью произвольной точки $\tilde{\mathbb{M}}$ пересекается лишь конечное число множеств вида $\mathbb{D}a$;
- 4) образ множества всех внутренних точек фундаментальной области, $(\text{int } \mathbb{D})a$, при действии любого преобразования $a \in \mathbb{G}$, отличного от единичного, не пересекается с множеством внутренних точек фундаментальной области, $(\text{int } \mathbb{D})a \cap \text{int } \mathbb{D} = \emptyset$, $a \neq e$. $\square$

Фундаментальная область $\mathbb{D}$ всегда является многообразием с краем той же размерности, что и само $\tilde{\mathbb{M}}$.

Пример 11.2.4. Группа трансляций $\mathbb{G}$ евклидовой плоскости $\mathbb{R}^2$ на всевозможные векторы с целочисленными компонентами

$$\mathbb{R}^2 \ni (x, y) \mapsto (x + m, y + n) \in \mathbb{R}^2, \quad \forall m, n \in \mathbb{Z},$$

действует гладко, свободно и собственноразрывно. Фактор пространство $\mathbb{R}^2/\mathbb{G}$ представляет собой тор $\mathbb{T}^2$. В качестве фундаментальной области можно выбрать единичный квадрат

$$\mathbb{D} = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

Тор можно представить в виде квадрата на плоскости, у которого отождествлены противоположные стороны. $\square$

Из определения фундаментальной области следует, что отображение $\tilde{\mathbb{M}} \rightarrow \text{int } \mathbb{D}$ всегда представляет собой накрытие. Если фундаментальная область известна, то пространство орбит $\tilde{\mathbb{M}}/\mathbb{G}$ получается из фундаментальной области путем склеивания граничных точек.

11.3 Группа скольжений и существование накрытий

В предыдущем разделе было показано, что отображение многообразия в пространство орбит для группы преобразований, действующей свободно и собственноразрывно, является накрытием. Теперь мы рассмотрим обратный вопрос о том, в каком случае заданное накрытие можно представить, в виде отображения накрываемого пространства в пространство орбит относительно действия некоторой группы преобразований и какова эта группа.

Введем новое понятие, которое дает возможность описать произвол, существующий при построении накрываемого пространства, если база задана.

Определение. Группой скольжений накрытия $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ называется группа всех гомеоморфизмов $h : \tilde{\mathbb{M}} \rightarrow \tilde{\mathbb{M}}$, при которых $p \circ h = p$, т.е. диаграмма

$$\begin{array}{ccc} \tilde{\mathbb{M}} & \xrightarrow{h} & \tilde{\mathbb{M}} \\ & \searrow p & \downarrow p \\ & & \mathbb{M} \end{array}$$

коммутативна. Эта группа обозначается $\mathbb{G}(\tilde{M}, p, M)$. $\square$

Теорема 11.3.1. *Если топологическое пространство $\tilde{M}$ связно и локально линейно связно, то действие группы скольжений $\mathbb{G}(\tilde{M}, p, M)$ на $\tilde{M}$ свободно и собственнo разрывно.*

Доказательство. См., например, [19], теорема 21.8. $\square$

В частности, действие группы скольжений на связном многообразии $\tilde{M}$ является свободным и собственнo разрывным. В этом случае группа скольжений $\mathbb{G}(\tilde{M}, p, M)$ конечна или счетна. Поэтому пара $(\tilde{M}, \mathbb{G}(\tilde{M}, p, M))$ является группой преобразований.

Теорема 11.3.2. *Пусть $p : \tilde{M} \rightarrow M$ – накрытие и накрывающее топологическое пространство $\tilde{M}$ связно и локально линейно связно. Если индуцированная группа $p_*\pi_1(\tilde{M}, \tilde{x})$ является нормальной подгруппой $\pi_1(M, x)$, где $x = p(\tilde{x})$, то база M гомотоморфна пространству орбит $\tilde{M}/\mathbb{G}(\tilde{M}, p, M)$.*

Доказательство. См., например, [19], теорема 21.9. $\square$

Следствие. Пусть пространство $\tilde{M}$ связно и локально линейно связно. Если $p_*\pi_1(\tilde{M}, \tilde{x})$ – нормальная подгруппа в $\pi_1(M, x)$, то $\pi_1(M, x)/p_*\pi_1(\tilde{M}, \tilde{x}) \simeq \mathbb{G}(\tilde{M}, p, M)$. $\square$

Доказательство. Прямое следствие теорем 11.3.2 и 11.2.3. $\square$

Следствие. Если накрывающее пространство $\tilde{M}$ односвязно и локально линейно связно, то $\pi_1(M, x) \simeq \mathbb{G}(\tilde{M}, p, M)$. $\square$

По своей сути две предыдущие теоремы и следствия означают, что произвольное накрытие можно представить, как отображение многообразия $\tilde{M}$ в пространство орбит $M = \tilde{M}/\mathbb{G}(\tilde{M}, p, M)$. При этом роль группы преобразований играет группа скольжений $\mathbb{G}(\tilde{M}, p, M)$. Если накрывающее пространство $\tilde{M}$ – многообразие, то действие группы скольжений сводится к перестановке листов накрытия.

Теперь обсудим вопрос о существовании универсального накрытия для заданной базы M . Универсальное накрытие существует, если на топологическое пространство M наложен ряд условий. Чтобы не вводить новых понятий для их формулировки, мы ограничимся многообразиями, для которых эти условия выполняются.

Теорема 11.3.3. *Произвольное связное n -мерное многообразие M имеет универсальное накрытие $p : \tilde{M} \rightarrow M$, где накрывающее пространство $\tilde{M}$ – также n -мерное многообразие (связное и односвязное). Универсальное накрывающее пространство определено с точностью до действия группы скольжений $\mathbb{G}(\tilde{M}, p, M)$, которая действует на $\tilde{M}$ свободно и собственнo разрывно.*

Доказательство. См., например, [19], глава 22. $\square$

Эта теорема очень важна, т.к. позволяет разделить задачу классификации многообразий на два этапа: сначала описать все односвязные многообразия, а затем найти все группы преобразований, действующие на них свободно и собственнo разрывно. Данная задача решена для поверхностей. Для многообразий размерности три и выше вопрос остается открытым.

Глава 12

Главные и ассоциированные расслоения

Теория расслоений или расслоенных пространств играет важнейшую роль в современной математической физике. Достаточно отметить, что в основе общей теории относительности лежит расслоение реперов, которое является главным расслоением со структурной группой $\mathrm{GL}(n, \mathbb{R})$. В основе калибровочных моделей элементарных частиц лежит главное расслоение с полупростой компактной структурной группой, которой обычно является некоторая подгруппа унитарной группы $\mathbb{U}(n)$. В настоящем разделе мы дадим определения и рассмотрим основные свойства главных и ассоциированных расслоений.

12.1 Главные расслоения

В дифференциальной геометрии важную роль играют главные расслоения. Фактически, они лежат в основе многих геометрических конструкций.

Определение. Главным расслоением называется четверка $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, где $\mathbb{P}$ и $\mathbb{M}$ – многообразия, $\mathbb{G}$ – группа Ли, $\pi : \mathbb{P} \rightarrow \mathbb{M}$ – отображение, которые удовлетворяют следующим условиям:

- 1) определено свободное дифференцируемое действие группы $\mathbb{G}$ на $\mathbb{P}$ справа:

$$\mathbb{P} \times \mathbb{G} \ni (p, a) \mapsto pa \in \mathbb{P}; \quad (12.1)$$

- 2) $\mathbb{M}$ есть факторпространство для $\mathbb{P}$ по отношению эквивалентности, индуцированному действием группы $\mathbb{G}$, и каноническая проекция $\pi : \mathbb{P} \rightarrow \mathbb{M} = \mathbb{P}/\mathbb{G}$ дифференцируема;
- 3) каждая точка $x \in \mathbb{M}$ имеет окрестность $\mathbb{U}_x$ такую, что существует диффеоморфизм

$$\chi : \pi^{-1}(\mathbb{U}_x) \ni p \mapsto \chi(p) = (\pi(p), \varphi(p)) \in \mathbb{U}_x \times \mathbb{G} \quad (12.2)$$

такой, что отображение $\varphi : \pi^{-1}(\mathbb{U}_x) \rightarrow \mathbb{G}$ удовлетворяет условию $\varphi(pa) = (\varphi(p))a$ для всех $p \in \pi^{-1}(\mathbb{U}_x)$ и $a \in \mathbb{G}$ (локальная тривиальность).

Многообразие $\mathbb{P}$ называется *пространством расслоения*, $\mathbb{M}$ – *базой расслоения*, $\mathbb{G}$ – *структурной группой* и π – *проекцией*. $\square$

Поскольку отображение (12.2) является диффеоморфизмом, то

$$\dim \mathbb{P} = \dim \mathbb{M} + \dim \mathbb{G}.$$

Иногда в качестве структурной группы мы будем рассматривать не группу Ли, а группу, состоящую из конечного или счетного набора элементов. Для удобства такие группы мы будем считать 0-мерными группами Ли. В этом случае $\dim \mathbb{P} = \dim \mathbb{M}$.

Замечание. Каждое главное расслоение является расслоением в смысле определения, данного в разделе 2.4. В дополнение к общему определению расслоения мы зафиксировали типичный слой, предположив, что им является группа Ли $\mathbb{G}$, и добавили действие этой группы на пространстве расслоения $\mathbb{P}$ так, чтобы оно было согласовано с проекцией. Дифференцируемая структура на $\mathbb{P}$ согласована с дифференцируемыми структурами на $\mathbb{M}$ и $\mathbb{G}$, поскольку отображение χ , по предположению, является диффеоморфизмом. На самом деле, можно было бы потребовать только непрерывность отображения χ , а затем с его помощью перенести дифференцируемую структуру с $\mathbb{M}$ и $\mathbb{G}$ на пространство расслоения $\mathbb{P}$. $\square$

Замечание. Пара $(\mathbb{P}, \mathbb{G})$ является группой преобразований, определенной в разделе 9. Однако, не всякая группа преобразований есть главное расслоение. Напомним, что в общем случае пространство орбит $\mathbb{M}/\mathbb{G}$ группы преобразований $(\mathbb{M}, \mathbb{G})$ может оказаться нехаусдорфовым топологическим пространством и на нем нельзя ввести дифференцируемую структуру. $\square$

Условие локальной тривиальности главного расслоения можно изобразить в виде коммутативной диаграммы

$$\begin{array}{ccc} \mathbb{P} \supset \pi^{-1}(\mathbb{U}_x) & \xrightarrow{\chi} & \mathbb{U}_x \times \mathbb{G} \\ & \searrow \cong & \downarrow \text{pr} = \pi \circ \chi^{-1} \\ & & \mathbb{U}_x \end{array}$$

где pr – естественная проекция прямого произведения $\mathbb{U}_x \times \mathbb{G} \xrightarrow{\text{pr}} \mathbb{U}_x$ на первый сомножитель. Групповое действие на пространстве расслоения описывается эквивалентной диаграммой

$$\begin{array}{ccc} \pi^{-1}(\mathbb{U}_x) \times \mathbb{G} & \xrightarrow{\chi \circ \text{id}} & (\mathbb{U}_x \times \mathbb{G}) \times \mathbb{G} \\ \downarrow a_{\mathbb{P}} & & \downarrow a_{\mathbb{U}_x \times \mathbb{G}} \\ \pi^{-1}(\mathbb{U}_x) & \xrightarrow{\chi} & \mathbb{U}_x \times \mathbb{G} \end{array}$$

где $a_{\mathbb{P}}$ и $a_{\mathbb{U}_x \times \mathbb{G}}$ обозначают действие элемента группы $a \in \mathbb{G}$ на пространстве расслоения и прямом произведении $\mathbb{M} \times \mathbb{G}$.

Определение. Для каждой точки базы $x \in \mathbb{M}$ множество $\pi^{-1}(x)$ есть замкнутое подмногообразие в пространстве расслоения $\mathbb{P}$, которое называется *слоем* над x . *Сечением* или *глобальным сечением* главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ называется дифференцируемое отображение $\sigma : \mathbb{M} \rightarrow \mathbb{P}$ такое, что $\pi \circ \sigma = \text{id}_{\mathbb{M}}$. Дифференцируемое отображение $\sigma : \mathbb{U} \rightarrow \mathbb{P}$, где $\mathbb{U} \subset \mathbb{M}$ – некоторая окрестность базы, называется *локальным сечением* над $\mathbb{U}$, если $\pi \circ \sigma = \text{id}_{\mathbb{U}}$. $\square$

Каждый слой является орбитой $p\mathbb{G}$ какой-либо точки $p \in \pi^{-1}(x)$. То, что каждый слой представляет собой замкнутое подмногообразие в $\mathbb{P}$ является следствием предложения 9.1.3. Поскольку действие группы Ли $\mathbb{G}$ в каждом слое свободно и транзитивно, то между точками слоя и структурной группы имеется взаимно однозначное соответствие. То есть каждый слой $\pi^{-1}(x)$ диффеоморфен $\mathbb{G}$. Этот диффеоморфизм осуществляет функция φ в отображении (12.2) при фиксированном $x \in \mathbb{M}$.

При каждом фиксированном $p \in \mathbb{P}$ отображение (12.1) дифференцируемо. Поэтому каждой точке главного расслоения соответствует диффеоморфизм структурной группы на типичный слой в данной точке

$$p : \mathbb{G} \ni a \mapsto pa \in \pi^{-1}(x), \quad (12.3)$$

где $x = \pi(p)$. В дальнейшем мы иногда будем рассматривать точку p именно в этом смысле, как отображение.

Сечение σ , если оно существует, не может быть сюръективным отображением, т.к. размерность базы в общем случае меньше размерности расслоения. Для накрытий размерность базы совпадает с размерностью главного расслоения. Если накрытие многолистно, то сечение также не является сюръективным отображением. Сечение сюръективно в одном случае, когда структурная группа состоит из единственного элемента – единицы. Сужение проекции π на образ базы $\sigma(\mathbb{M})$ в $\mathbb{P}$ является дифференцируемым отображением, которое обратно к сечению σ . Поэтому пара $(\sigma, \mathbb{M})$ является подмногообразием в $\mathbb{P}$. Если $\dim \mathbb{G} \geq 1$, то это подмногообразие замкнуто в $\mathbb{P}$ (см. раздел 2.10).

Пример 12.1.1. Пусть $\mathbb{G}$ – группа Ли и $\mathbb{M}$ – многообразие. Определим действие группы $\mathbb{G}$ справа на прямом произведении $\mathbb{P} = \mathbb{M} \times \mathbb{G}$:

$$\mathbb{G} \ni b : \mathbb{M} \times \mathbb{G} \ni x, a \mapsto x, ab \in \mathbb{M} \times \mathbb{G}.$$

Тогда четверка $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, где $\pi : \mathbb{M} \times \mathbb{G} \rightarrow \mathbb{M}$ – каноническая проекция, является главным расслоением. Это главное расслоение имеет глобальное сечение $\sigma : \mathbb{M} \rightarrow \mathbb{P}$, $\pi \circ \sigma = \text{id}_{\mathbb{M}}$. Например,

$$\sigma : \mathbb{M} \ni x \mapsto x, e \in \mathbb{M} \times \mathbb{G} \quad \square$$

Определение. Главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ называется *тривиальным*, если оно изоморфно главному расслоению вида $\mathbb{M} \times \mathbb{G} \xrightarrow{\text{pr}_{\mathbb{M}}} \mathbb{M}$. $\square$

Пример 12.1.2. Расслоение реперов $\mathbb{L}(\mathbb{M}) = \mathbb{P}(\mathbb{M}, \pi, \text{GL}(n, \mathbb{R}))$, рассмотренное в разделе 5.4, является главным расслоением со структурной группой $\text{GL}(n, \mathbb{R})$. $\square$

Пример 12.1.3. Пусть $\mathbb{G}$ – группа Ли и $\mathbb{H} \subset \mathbb{G}$ – ее замкнутая подгруппа. Согласно теореме 9.1.1 на пространстве правых смежных классов $\mathbb{H}a \in \mathbb{G}/\mathbb{H}$, где $a \in \mathbb{G}$, можно задать дифференцируемую структуру. Тогда $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$ – главное расслоение с базой $\mathbb{G}/\mathbb{H}$, структурной группой $\mathbb{H}$ и проекцией $\pi : \mathbb{G} \rightarrow \mathbb{G}/\mathbb{H}$. $\square$

Пример 12.1.4. Накрытия $\tilde{\mathbb{M}}(\mathbb{M}, \pi, \mathbb{G}(\tilde{\mathbb{M}}, p, \mathbb{M}))$, где $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ – отображение накрытия, рассмотренные в разделе 11, являются главными расслоениями с 0-мерной структурной группой Ли, которой является группа скольжений $\mathbb{G}(\tilde{\mathbb{M}}, p, \mathbb{M})$. $\square$

Пример 12.1.5. Множество вещественных чисел без нуля образует абелеву группу относительно умножения, которую мы обозначим $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$. Пусть эта группа действует в евклидовом пространстве без начала координат $\mathbb{R}^n \setminus \{0\}$ посредством умножения декартовых координат:

$$\mathbb{R}_0 \ni a : (x^1, \dots, x^n) \mapsto (ax^1, \dots, ax^n).$$

Это действие дифференцируемо, свободно и транзитивно на орбитах. Поэтому $\mathbb{R}^n \setminus \{0\} \xrightarrow{\pi} \mathbb{RP}^{n-1}$ – тривиальное главное расслоение со структурной группой $\mathbb{R}_0$, базой которого является вещественное проективное пространство $\mathbb{RP}^{n-1} = (\mathbb{R}^n \setminus \{0\})/\mathbb{R}_0$. $\square$

Пример 12.1.6. Если в предыдущем примере группу $\mathbb{R}_0$ заменить на группу положительных чисел $\mathbb{R}_+$ по умножению, то получим тривиальное главное расслоение $\mathbb{R}^n \setminus \{0\} \xrightarrow{\pi} \mathbb{S}^{n-1}$ со структурной группой $\mathbb{R}_+$, базой которого является сфера $\mathbb{S}^{n-1} = (\mathbb{R}^n \setminus \{0\})/\mathbb{R}_+$. $\square$

Пример 12.1.7. Пусть $\mathbb{C}_0 := \mathbb{C} \setminus \{0\}$ – группа комплексных чисел по отношению к умножению, которая действует в n -мерном комплексном пространстве без начала координат $\mathbb{C}^n \setminus \{0\}$ умножением,

$$\mathbb{C}_0 \ni a : (z^1, \dots, z^n) \mapsto (az^1, \dots, az^n).$$

Тогда $\mathbb{C}^n \setminus \{0\} \xrightarrow{\pi} \mathbb{CP}^{n-1}$ – тривиальное главное расслоение со структурной группой $\mathbb{C}_0$, базой которого является комплексное проективное пространство $\mathbb{CP}^{n-1} = (\mathbb{C}^n \setminus \{0\})/\mathbb{C}_0$. $\square$

В примере 12.1.1 мы отметили, что произвольное тривиальное расслоение имеет глобальное сечение. Справедливо и обратное утверждение.

Предложение 12.1.1. Если главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ имеет глобальное сечение $\sigma : \mathbb{M} \rightarrow \mathbb{P}$, то оно изоморфно тривиальному, т.е. существует диффеоморфизм

$$f_{\mathbb{P}} : \mathbb{P} \ni p \mapsto (\pi(p), \varphi(p)) \in \mathbb{M} \times \mathbb{G} \quad (12.4)$$

такой, что следующая диаграмма коммутативна:

$$\begin{array}{ccc} \mathbb{P} & \xrightarrow{f_{\mathbb{P}}} & \mathbb{M} \times \mathbb{G} \\ & \searrow \cong & \downarrow \text{pr}_{\mathbb{M}} \\ & & \mathbb{M} \end{array} \quad (12.5)$$

где $\text{pr}_{\mathbb{M}}$ – проекция на первый сомножитель, и справедливо равенство

$$f_{\mathbb{P}}(pa) = (\pi(p), \varphi(p)a) \quad \forall p \in \mathbb{P}, a \in \mathbb{G}. \quad (12.6)$$

Доказательство. Пусть

$$\sigma : \mathbb{M} \ni x \mapsto \sigma(x) \in \mathbb{P}, \quad \pi \circ \sigma = \text{id}_{\mathbb{M}},$$

– глобальное сечение главного расслоения. Для каждой точки слоя $p \in \pi^{-1}(x)$ определим единственный элемент группы $a \in \mathbb{G}$, для которого $p = \sigma(x)a(p)$. Таким образом, определено отображение $p \mapsto f_{\mathbb{P}}(p) = (\pi(p), a(p))$. Легко проверить, что это отображение удовлетворяет требуемым свойствам. $\square$

Следствие. Нетривиальные главные расслоения имеют только локальные сечения. $\square$

В общем случае изоморфизм расслоений будет определен позже в разделе 12.3.

Диаграмму (12.5) вместе с условием (12.6) можно изобразить в виде коммутативной диаграммы

$$\begin{array}{ccc} \mathbb{P} \times \mathbb{G} & \xrightarrow{f_{\mathbb{P}} \circ \text{id}} & (\mathbb{M} \times \mathbb{G}) \times \mathbb{G} \\ \downarrow a_{\mathbb{P}} & & \downarrow a_{\mathbb{M} \times \mathbb{G}} \\ \mathbb{P} & \xrightarrow{f_{\mathbb{P}}} & \mathbb{M} \times \mathbb{G} \end{array},$$

где $a_{\mathbb{P}}$ и $a_{\mathbb{M} \times \mathbb{G}}$ обозначает действие элемента $a \in \mathbb{G}$, соответственно, на $\mathbb{P}$ и $\mathbb{M} \times \mathbb{G}$. Это означает, что отображение (12.4) является эквивариантным (см. раздел 9.4).

Теорема 12.1.1. Если база главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ диффеоморфна евклидову пространству, $\mathbb{M} \approx \mathbb{R}^n$, то главное расслоение изоморфно тривиальному, $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G}) \simeq \mathbb{M} \times \mathbb{G}$.

Доказательство. Рассмотрим главное расслоение $\mathbb{P}(\mathbb{R}^n, \pi, \mathbb{G})$, базой которого является евклидово пространство. Согласно теореме 13.1.2 любое главное расслоение допускает связность. Предположим, что на $\mathbb{P}(\mathbb{R}^n, \pi, \mathbb{G})$ задана какая-либо связность. Рассмотрим в $\mathbb{R}^n$ сферическую систему координат. Тогда каждая отличная от начала координат точка $x \in \mathbb{R}^n$ параметризуется парой (r, n) , где r – расстояние от начала координат и n – единичный вектор, определяющий направление луча, выходящего из начала координат и на котором лежит точка x . Пусть $p_0 \in \pi^{-1}(0)$ – произвольная точка из слоя над началом координат. Согласно предложению 13.3.1 для каждого луча существует его единственный горизонтальный лифт в пространство расслоения $\mathbb{P}$ с началом в точке p_0 . Обозначим через $p(x)$ единственную точку на горизонтальном лифте луча, которая лежит над x , т.е. $\pi(p) = x$. Таким образом, $p(x)$ – это глобальное сечение главного расслоения. Следовательно, по предложению 12.1.1, главное расслоение тривиально $\mathbb{P} = \mathbb{R}^n \times \mathbb{G}$. Если база диффеоморфна $\mathbb{R}^n$, то сферическая система координат просто переносится на $\mathbb{M}$ с помощью диффеоморфизма, и построение глобального сечения повторяется. $\square$

При рассмотрении многообразий в разделе 2.1 были введены функции склейки, с помощью которых осуществляется преобразование координат в двух пересекающихся картах. Обобщением этого понятия на случай главных расслоений являются функции перехода, которые вводятся следующим образом.

Определение. В силу условия 3) в определении главного расслоения на базе можно выбрать такое координатное покрытие, $\mathbb{M} = \bigcup_i \mathbb{U}_i$, что

$$\chi_i : \pi^{-1}(\mathbb{U}_i) \ni p \mapsto (\pi(p), \varphi_i(p)) \in \mathbb{U}_i \times \mathbb{G},$$

причем $\varphi_i(pa) = \varphi_i(p)a$.

Пусть две карты пересекаются, $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$. Если $p \in \pi^{-1}(\mathbb{U}_i \cap \mathbb{U}_j)$, то

$$\varphi_j(pa) \circ \varphi_i(pa)^{-1} = \varphi_j(p) \circ \varphi_i(p)^{-1},$$

где φ_i рассматривается как отображение фиксированного слоя $\pi^{-1}(\pi(p))$ в группу $\mathbb{G}$. Следовательно, отображение $\varphi_j(p) \circ \varphi_i(p)^{-1}$ зависит только от точки базы $x = \pi(p)$. Поэтому определено отображение

$$a_{ji} : \mathbb{U}_i \cap \mathbb{U}_j \ni x = \pi(p) \mapsto a_{ji}(x) := \varphi_j(p) \circ \varphi_i(p)^{-1} \in \mathbb{G}. \quad (12.7)$$

Эти функции на $\mathbb{M}$ со значениями в $\mathbb{G}$ называются *функциями перехода* или *функциями склейки* главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, соответствующими координатному покрытию $\mathbb{M} = \bigcup_i \mathbb{U}_i$. Набор функций склейки называется *склеивающим коциклом* главного расслоения $\mathbb{P}$. $\square$

Функции перехода показывают насколько в различных картах “сдвинуты” образы слоя в структурной группе над фиксированной точкой $x \in \mathbb{M}$.

Из определения (12.7) следует, что функции склейки обладают следующим свойством

$$a_{ij}(x) = (a_{ji}(x))^{-1} \quad \forall x \in \mathbb{U}_i \cap \mathbb{U}_j. \quad (12.8)$$

Кроме того, нетрудно проверить, что для трех пересекающихся карт выполнено равенство

$$a_{ij}(x)a_{jk}(x)a_{ki}(x) = e, \quad \forall x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k, \quad (12.9)$$

где e – единица структурной группы. Поэтому определение функций перехода корректно.

Пример 12.1.8. Пусть $\mathbb{M}$, $\dim \mathbb{M} = n$, – многообразие. Рассмотрим расслоение реперов $\mathbb{L}(\mathbb{M})$ (см. раздел 5.4). Пусть $\mathbb{U}_i$ и $\mathbb{U}_j$ – две пересекающиеся карты с координатами x^α и $x^{\alpha'}$, $\alpha, \alpha' = 1, \dots, n$. Репер в этих картах имеет компоненты $e^\alpha_a \in \mathbb{GL}(n, \mathbb{R})$ и $e^{\alpha'}_a \in \mathbb{GL}(n, \mathbb{R})$, $a = 1, \dots, n$, которые связаны между собой преобразованием

$$e^{\alpha'}_a = e^\alpha_a \partial_\alpha x^{\alpha'}.$$

Таким образом, функциями перехода для расслоения реперов являются матрицы Якоби преобразования координат. Эти матрицы, как легко проверить, удовлетворяют условиям (12.8) и (12.9). $\square$

Таким образом, для каждого главного расслоения можно однозначно построить семейство функций перехода, соответствующих заданному координатному покрытию базы, и эти функции перехода удовлетворяют равенствам (12.8), (12.9). Справедливо также обратное утверждение.

Теорема 12.1.2. Пусть $\mathbb{M}$ – многообразие с координатным покрытием $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и $\mathbb{G}$ – группа Ли. Если заданы отображения $a_{ji} : \mathbb{U}_i \cap \mathbb{U}_j \rightarrow \mathbb{G}$ для всех непустых пересечений $\mathbb{U}_i \cap \mathbb{U}_j$ такие, что выполнены условия (12.8) и (12.9) во всех областях пересечения карт, то существует единственное с точностью до изоморфизма главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с функциями перехода a_{ji} .

Доказательство. Рассмотрим несвязное объединение $\mathbb{Q} := \sqcup_i (\mathbb{U}_i \times \mathbb{G})$. Введем на этом множестве отношение эквивалентности:

$$\begin{aligned} (x, a) &\sim (x, aa_{ij}(x)), & \forall (x, a) \in \mathbb{U}_i \times \mathbb{G}, \quad (x, aa_{ij}(x)) \in \mathbb{U}_j \times \mathbb{G}, \\ (x, a) &\sim (x, a), & \forall (x, a) \in \mathbb{U}_i \times \mathbb{G}. \end{aligned}$$

Из свойств (12.8), (12.9) следует, что это действительно отношение эквивалентности. Обозначим фактор пространство $\mathbb{Q}/\sim$ через $\mathbb{P}$ и введем на нем естественную дифференцируемую структуру. Пусть $p := \langle x, a \rangle$ – точка $\mathbb{P}$, т.е. класс эквивалентности пары (x, a) . Определим действие группы $\mathbb{G}$ на $\mathbb{P}$ формулой

$$\mathbb{P} \times \mathbb{G} \ni \langle x, a \rangle, b \mapsto \langle x, ab \rangle \in \mathbb{P}.$$

Определим также проекцию

$$\pi : \mathbb{P} \ni \langle x, a \rangle \mapsto \pi(\langle x, a \rangle) := x \in \mathbb{M}.$$

Нетрудно проверить, что все свойства главного расслоения для четверки $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ выполнены.

Теперь докажем единственность построенного главного расслоения с точностью до изоморфизма. Пусть множество функций склеек $\{a_{ji}(x)\}$ построено для некоторого главного расслоения $\mathbb{P}'(\mathbb{M}, \pi', \mathbb{G})$ с фиксированным покрытием базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$. Построим гомеоморфизм $f_{\mathbb{P}} : \mathbb{P} \rightarrow \mathbb{P}'$, где $f_{\mathbb{P}}$ совпадает с χ_i^{-1} на каждом $\mathbb{U}_i \times \mathbb{G}$. Для корректности этого определения нужно убедиться, что отображения χ_i^{-1} и χ_j^{-1} совпадают на общей области определения. Действительно, если $x \in \mathbb{U}_i \cap \mathbb{U}_j$ и $(x, a) \in \mathbb{U}_i \times \mathbb{G}$, то по построению пары (x, a) и (x, aa_{ji}) определяют одну и ту же точку $\langle x, a \rangle \in \mathbb{P}$. Из определения функций склейки $a_{ji}(x)$ следуют равенства:

$$\chi_i^{-1}(x, a) = \chi_j^{-1} \circ \chi_j \circ \chi_i^{-1}(x, a) = \chi_j^{-1}(x, aa_{ji}).$$

Эквивариантность отображения $f_{\mathbb{P}}$ и коммутативность диаграммы

$$\begin{array}{ccc} \mathbb{P} & \xrightarrow{f_{\mathbb{P}}} & \mathbb{P}' \\ & \searrow \cong & \downarrow \pi' \\ & & \mathbb{M} \end{array} \quad (12.10)$$

следуют из построения. Это и означает, что главные расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и $\mathbb{P}'(\mathbb{M}, \pi', \mathbb{G})$ изоморфны. (Подробнее изоморфизм расслоений рассматривается в разделе 12.3. Там будет показано, что построенный изоморфизм $f_{\mathbb{P}}$ относится к классу вертикальных автоморфизмов.) $\square$

Склеивающий коцикл определяется расслоением неоднозначно. Он зависит от выбора координатного покрытия базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и тривиализаций $\{\chi_i\}$.

Определение. Два коцикла $\{a_{ji}\}$ и $\{a'_{ji}\}$, соответствующие заданному координатному покрытию базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$, называются *эквивалентными*, если существуют отображения $b_i : \mathbb{U}_i \rightarrow \mathbb{G}$ такие, что выполнены равенства

$$a'_{ji} = b_j a_{ji} b_i^{-1}, \quad \forall x \in \mathbb{U}_i \cap \mathbb{U}_j. \quad \square$$

Теорема 12.1.3. Два главных расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и $\mathbb{P}'(\mathbb{M}, \pi', \mathbb{G})$ изоморфны тогда и только тогда, когда их склеивающие коциклы $\{a_{ji}\}$ и $\{a'_{ji}\}$, соответствующие некоторому координатному покрытию базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$, эквивалентны.

Доказательство. См., например, [75], лекция 1, теорема 1. $\square$

Пример 12.1.9 (Расслоение Хопфа). Реализуем трехмерную сферу $\mathbb{S}^3$ в двумерном комплексном пространстве $\mathbb{C}^2$ с помощью вложения

$$|z^1|^2 + |z^2|^2 = 1, \quad (z^1, z^2) \in \mathbb{C}^2. \quad (12.11)$$

Каждое уравнение $az^1 + bz^2 = 0$, где $(a, b) \in \mathbb{C}^2$ и по крайней мере одно из комплексных чисел a, b отлично от нуля, задает *комплексную прямую*, комплексной размерности один, проходящую через начало координат. Обратно. Комплексная прямая определяет пару комплексных чисел $(a, b) \in \mathbb{C}^2$ с точностью до отношения эквивалентности $(a, b) \sim (a', b')$, если и только если $|a'| = \lambda|a|$, $|b'| = \lambda|b|$, $\lambda \in \mathbb{R}_+$, или $a' = a e^{it}$, $b' = b e^{it}$, $t \in [0, 2\pi]$. Множество всех комплексных прямых, проходящих через начало координат, является одномерным комплексным проективным пространством $\mathbb{CP}^1$ и называется *комплексной проективной прямой*. Каждая комплексная прямая, проходящая через начало координат, пересекает трехмерную сферу $\mathbb{S}^3$ по большой окружности $\mathbb{S}^1$, называемой *окружностью Хопфа*, которая задается одной из двух систем уравнений:

$$\begin{aligned} |z^1|^2 &= \frac{1}{1 + |k|^2}, & z^2 &= k z^1, & k &:= -\frac{a}{b}, & b &\neq 0, \\ |z^2|^2 &= \frac{1}{1 + |k|^2}, & z^1 &= k z^2, & k &:= -\frac{b}{a}, & a &\neq 0. \end{aligned}$$

Поскольку ровно одна окружность Хопфа проходит через каждую точку $\mathbb{S}^3$, эти окружности заполняют всю трехмерную сферу. При этом окружности Хопфа взаимно однозначно соответствуют комплексным прямым в $\mathbb{C}^2$, проходящим через начало координат. Две точки (z^1, z^2) и (z'^1, z'^2) лежат на одной окружности Хопфа тогда и только тогда, когда их координаты отличаются на фазовый множитель: $z'^1 = z^1 e^{it}$, $z'^2 = z^2 e^{it}$. Выбрав подходящим образом параметр t , всегда можно добиться, например, чтобы $\operatorname{im} z^2 = 0$. Тогда каждой прямой, проходящей через начало координат, будет соответствовать одна и только одна точка на окружности Хопфа. Множество таких точек образует двумерную сферу $\mathbb{S}^2$, которая задается в $\mathbb{C}^2$ уравнением

$$(x^1)^2 + (y^1)^2 + (x^2)^2 = 1, \quad y^2 = 0, \quad (12.12)$$

где $z^1 = x^1 + iy^1$, $z^2 = x^2 + iy^2$. Таким образом, комплексная проективная прямая диффеоморфна двумерной сфере, $\mathbb{CP}^1 \approx \mathbb{S}^2$. Соответствующая проекция $\mathbb{S}^3 \xrightarrow{\pi} \mathbb{S}^2$ определяет главное расслоение с базой $\mathbb{S}^2$ и типичным слоем $\mathbb{G} = \mathbb{U}(1) \approx \mathbb{S}^1$, которое называется *расслоением Хопфа*. Структурная группа действует на пространстве расслоения $\mathbb{S}^3$ умножением комплексных координат на e^{it} , $t \in [0, 2\pi]$.

Отображение $g_t : \mathbb{S}^3 \rightarrow \mathbb{S}^3$, $t \in \mathbb{R}$, задаваемое умножением комплексных координат в $\mathbb{C}^2$ на e^{it} , является изометрией для метрики, индуцированной вложением (12.11). Отсюда следует, что у трехмерной сферы $\mathbb{S}^3$ существуют изометрии, не имеющие выделенной оси: в каждой точке это движение выглядит точно так же, как и в любой другой. Это показывает, что трехмерная сфера $\mathbb{S}^3$ “более круглая”, чем ее двумерный аналог $\mathbb{S}^2$. Однопараметрическая группа преобразований g_t называется *поток Хопфа*.

Очевидно, что метрика проективной прямой $\mathbb{CP}^1 \approx \mathbb{S}^2$, индуцированная вложением (12.12), является стандартной метрикой сферы (см. раздел 29.1).

На расслоение Хопфа можно взглянуть с другой точки зрения. Вспомним, что трехмерная сфера может быть оснащена групповой структурой, $\mathbb{S}^3 \approx \mathbb{SU}(2)$ (см.

раздел 1.8). Двумерные унитарные матрицы с единичным определителем могут быть параметризованы двумя комплексными числами

$$U = \begin{pmatrix} z^1 & z^2 \\ -\bar{z}^2 & \bar{z}^1 \end{pmatrix} \in \mathbb{SU}(2),$$

где z^1 и z^2 лежат на сфере (12.11). Группа $\mathbb{SU}(2)$ содержит подгруппу $\mathbb{U}(1)$, которую можно реализовать в виде диагональных матриц

$$\begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix} \in \mathbb{U}(1).$$

Эта подгруппа не является нормальной. Рассмотрим множество правых смежных классов $\mathbb{SU}(2)/\mathbb{U}(1)$. Два элемента группы $U, U' \in \mathbb{SU}(2)$ принадлежат одному смежному классу тогда и только тогда, когда $z'^1 = z^1 e^{it}$ и $z'^2 = z^2 e^{it}$. Поэтому расслоение Хопфа можно представить в эквивалентном виде $\mathbb{SU}(2) \xrightarrow{\pi} \mathbb{M} := \mathbb{SU}(2)/\mathbb{U}(1)$.

Действия группы $\mathbb{SU}(2)$ слева и справа на сферу $\mathbb{S}^3$, которую мы отождествим с $\mathbb{SU}(2)$,

$$U \mapsto a^{-1}Ub, \quad a, b \in \mathbb{SU}(2)$$

коммутируют и являются изометриями. Тем самым мы имеем гомоморфизм $\mathbb{SU}(2) \times \mathbb{SU}(2) \rightarrow \mathbb{SO}(4)$, где $\mathbb{SO}(4)$ является группой движений евклидова пространства $\mathbb{R}^4 \approx \mathbb{C}^2$ и вложенной трехмерной сферы. При этом действие группы $\mathbb{SU}(2)$ слева и справа является транзитивным, а группой изотропии являются преобразования вида $\pm(1, 1) \in \mathbb{SU}(2) \times \mathbb{SU}(2)$. Тем самым мы имеем сюръективный гомоморфизм с ядром $\mathbb{Z}_2$ или изоморфизм

$$\mathbb{SO}(4) \simeq \frac{\mathbb{SU}(2) \times \mathbb{SU}(2)}{\mathbb{Z}_2}. \quad (12.13)$$

Теперь дадим координатное описание расслоения Хопфа. Покроем базу $\mathbb{S}^2 = \overline{\mathbb{C}}$, которую мы отождествим с расширенной комплексной плоскостью, двумя картами следующим образом. Введем координату “окрестности нуля” $z := z^2/z^1$ при $z^1 \neq 0$ и “окрестности бесконечности” $w := z^1/z^2$, при $z^2 \neq 0$. Тогда база $\mathbb{S}^2 = \overline{\mathbb{C}}$ покрывается двумя картами:

$$\mathbb{U}_0 := \{z : z \in \mathbb{C}\}, \quad \text{и} \quad \mathbb{U}_\infty := \{w : w \in \mathbb{C}\}.$$

Тривиализация расслоения Хопфа задается двумя отображениями:

$$\begin{aligned} \chi_0 : \mathbb{S}^3 \ni (z^1, z^2) &\mapsto (z, e^{i \arg z^1}) \in \mathbb{C} \times \mathbb{S}^1, \\ \chi_\infty : \mathbb{S}^3 \ni (z^1, z^2) &\mapsto (w, e^{i \arg z^2}) \in \mathbb{C} \times \mathbb{S}^1. \end{aligned}$$

В области пересечения карт $\mathbb{U}_0 \cap \mathbb{U}_\infty$ координаты связаны преобразованием $z \mapsto w = 1/z$. Следовательно, расслоение Хопфа голоморфно. В области пересечения карт задано отображение:

$$\chi_\infty \circ \chi_0^{-1} : \mathbb{C} \times \mathbb{S}^1 \ni (z, e^{i \arg z^1}) \mapsto (z, e^{i \arg z^2}) \in (\mathbb{C} \setminus \{0\}) \times \mathbb{S}^1,$$

которое мы записали в координате z . Поскольку

$$e^{i \arg z^2} = e^{i \arg z z^1} = e^{i \arg z^1} e^{i \arg z},$$

то функция склейки имеет вид $a_{\infty 0} = e^{i \arg z}$. Таким образом, расслоение Хопфа имеет следующее координатное описание:

$$\begin{aligned} \mathbb{U}_0 &= \mathbb{C}, & \mathbb{U}_\infty &= \overline{\mathbb{C}} \setminus \{0\}, \\ a_{\infty 0} : \mathbb{C} \setminus \{0\} \ni z &\mapsto e^{i \arg z} \in \mathbb{U}(1) \approx \mathbb{S}^1. \end{aligned} \quad \square$$

В заключение настоящего раздела поясним соотношение между связностью (топологической) пространства главного расслоения $\mathbb{P}$ и связностями базы $\mathbb{M}$ и структурной группы $\mathbb{G}$.

Предложение 12.1.2. *Если база $\mathbb{M}$ и структурная группа $\mathbb{G}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ связны, то пространство расслоения $\mathbb{P}$ также связно.*

Доказательство. Пусть $\mathbb{M} = \bigcup_i \mathbb{U}_i$ покрытие базы открытыми множествами такое, что для каждого i существует тривиализация $\chi_i : \pi^{-1}(\mathbb{U}_i) \rightarrow \mathbb{U}_i \times \mathbb{G}$. Поскольку $\mathbb{U}_i$ и $\mathbb{G}$ связны, то $\pi^{-1}(\mathbb{U}_i)$ также связно. Поэтому $\mathbb{P}$ тоже связно, как объединение пересекающихся связных множеств. $\square$

Это предложение дает достаточное, но не необходимое условие связности пространства главного расслоения.

Пример 12.1.10. Рассмотрим главное расслоение $\mathbb{S}^1(\mathbb{S}^1, \pi, \mathbb{Z}^2)$, базой которого является окружность, а структурной группой – $\mathbb{Z}_2 = \{\pm 1\}$. Структурная группа не является связной, однако пространство расслоения связно. Это – двулистное накрытие окружности окружностью, рассмотренное в примере 11.1.5. $\square$

12.2 Ассоциированные расслоения

Дать определение расслоения, ассоциированного с заданным главным расслоением, в нескольких предложениях довольно затруднительно. Ассоциированные расслоения, хотя и просты по своему содержанию, требуют некоторой конструкции, которую мы сейчас опишем.

Определение. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение, и $(\mathbb{F}, \mathbb{G})$ – группа преобразований, т.е. задано дифференцируемое отображение

$$\mathbb{F} \times \mathbb{G} \ni (v, a) \mapsto va \in \mathbb{F}. \quad (12.14)$$

Сейчас мы построим расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$, которое ассоциировано с главным расслоением $\mathbb{P}$, и типичным слоем которого является многообразие $\mathbb{F}$. Во-первых, определим действие группы $\mathbb{G}$ на прямом произведении многообразий $\mathbb{P} \times \mathbb{F}$ по формуле

$$\mathbb{G} \ni a : \mathbb{P} \times \mathbb{F} \ni (p, v) \mapsto (pa, va) \in \mathbb{P} \times \mathbb{F}.$$

Факторпространство для $\mathbb{P} \times \mathbb{F}$ относительно такого действия группы обозначается $\mathbb{E} = \mathbb{P} \times_{\mathbb{G}} \mathbb{F}$. Немного позже мы введем на $\mathbb{E}$ дифференцируемую структуру, а пока $\mathbb{E}$ – всего лишь множество. Проекция π в главном расслоении $\mathbb{P}$ определяет отображение, которое мы обозначим той же буквой

$$\pi : \mathbb{P} \times \mathbb{F} \ni (p, v) \mapsto x = \pi(p) \in \mathbb{M}.$$

Это отображение индуцирует отображение факторпространства,

$$\pi_{\mathbb{E}} : \mathbb{E} \ni u \mapsto x = \pi_{\mathbb{E}}(u) := \pi(p) \in \mathbb{M},$$

где $u = (p\mathbb{G}, v\mathbb{G})$ – орбита точки $(p, v) \in \mathbb{P} \times \mathbb{F}$, которое называется проекцией $\mathbb{E}$ на $\mathbb{M}$. Для каждой точки базы $x \in \mathbb{M}$ множество $\pi_{\mathbb{E}}^{-1}(x)$ называется слоем в $\mathbb{E}$ над x . Поскольку главное расслоение $\mathbb{P}$ локально тривиально, то каждая точка $x \in \mathbb{M}$ имеет

окрестность $U \subset M$ такую, что прообраз $\pi^{-1}(U)$ диффеоморфен прямому произведению $U \times G$. Отождествляя $\pi^{-1}(U)$ с $U \times G$, мы видим, что действие структурной группы G на произведении $\pi^{-1}(U) \times F$ задается отображением

$$G \ni b : U \times G \times F \ni (x, a, v) \mapsto (x, ab, vb) \in U \times G \times F.$$

При этом орбита произвольной точки $(p, v) \in P \times F$ имеет вид

$$(x, aG, vG) = (x, G, vG), \quad \forall a \in G, v \in F.$$

На каждой орбите можно выбрать по одному представителю, соответствующему, например, единице группы $e \in G$. Тогда множество орбит прямого произведения $U \times G$ будет параметризовано парой элементов (x, e, v) , $x \in M$, $v \in F$. Отсюда следует, что диффеоморфизм $\pi^{-1}(U) \approx U \times G$ индуцирует изоморфизм $\pi_E^{-1}(U) \approx U \times F$ (локальная тривиальность). Поэтому на множестве E можно ввести дифференцируемую структуру. А именно, мы потребуем, чтобы множество $\pi_E^{-1}(U)$ было открытым подмногообразием в E , диффеоморфным прямому произведению $U \times F$ относительно изоморфизма $\pi_E^{-1}(U) \approx U \times F$. Таким образом, построено отображение $\chi_E : \pi_E^{-1}(U) \rightarrow U \times F$, которое является диффеоморфизмом. При этом проекция π_E будет дифференцируемым отображением $\pi_E : E \rightarrow M$. Назовем объект $E(M, \pi_E, F, G, P)$ *расслоением* с базой M , проекцией π_E , типичным слоем F и структурной группой G , *ассоциированным* с главным расслоением $P(M, \pi, G)$. $\square$

Поскольку отображение χ_E – диффеоморфизм, то

$$\dim E = \dim M + \dim F.$$

Легко проверить, что расслоение, ассоциированное с главным расслоением, является расслоением в смысле определения, данного в разделе 2.4. В отличие от данного ранее определения на ассоциированном расслоении определено действие структурной группы G справа. Действительно, согласно построению у каждой точки базы $x \in M$ существует окрестность U такая, что прообраз $\pi_E^{-1}(U)$ диффеоморфен прямому произведению $U \times F$. Отождествляя $\pi_E^{-1}(U)$ с $U \times F$, действие структурной группы справа задано равенством

$$G \ni a : U \times F \ni (x, v) \mapsto (x, va) \in U \times F,$$

где произведение va было определено в самом начале формулой (12.14). Таким образом, пара (E, G) представляет собой группу преобразований. Структурная группа действует внутри каждого слоя и не перемешивает слои между собой.

В терминах представителей действие структурной группы на E описывается следующим образом. Если $(p, v) \in P \times F$ – представитель точки $u \in E$, то представителем точки $ua \in E$ является пара $(p, va) \in P \times F$.

Замечание. В моделях математической физики наиболее распространенным примером типичного слоя F является векторное пространство, в котором задано представление структурной группы G . В калибровочных моделях Янга–Миллса поля материи, которым соответствуют скалярные или спинорные поля в пространстве-времени, имеют дополнительный изотопический индекс. Этот индекс нумерует компоненты вектора в некотором векторном (изотопическом) пространстве (типичном слое F), в котором задано представление калибровочной группы (структурной группы G). $\square$

В дальнейшем часть аргументов у ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ мы, для краткости, будем опускать, оставляя лишь те, которые наиболее существенны в данный момент.

Предложение 12.2.1. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение и $(\mathbb{F}, \mathbb{G})$ – группа преобразований. Пусть $\mathbb{E}$ – расслоение, ассоциированное с $\mathbb{P}$. Для каждого $p \in \mathbb{P}$ и каждого $v \in \mathbb{F}$ обозначим через $p(v)$ образ элемента $(p, v) \in \mathbb{P} \times \mathbb{F}$ при действии естественной проекции $\mathbb{P} \times \mathbb{F} \rightarrow \mathbb{E}$. Тогда каждая точка $p \in \mathbb{P}$ определяет отображение типичного слоя $\mathbb{F}$ в слой $\mathbb{F}_x = \pi_{\mathbb{E}}^{-1}(x)$,

$$p: \mathbb{F} \ni v \mapsto p(v) \in \mathbb{F}_x, \quad (12.15)$$

где $x = \pi(p)$, и $(pa^{-1})(v) = p(va)$ для всех $p \in \mathbb{P}$, $a \in \mathbb{G}$ и $v \in \mathbb{F}$, т.е. диаграмма

$$\begin{array}{ccc} \mathbb{F} & \xrightarrow{a} & \mathbb{F} \\ & \searrow pa^{-1} & \downarrow p \\ & & \mathbb{F}_x \end{array}$$

коммутативна.

Доказательство. Следует прямо из определения, если учесть, что точки (pa, va) и (p, v) из произведения $\mathbb{P} \times \mathbb{F}$ проектируются в одну и ту же точку $\mathbb{E}$. $\square$

Под изоморфизмом слоя $\mathbb{F}_x := \pi_{\mathbb{E}}^{-1}(x)$ на другой слой $\mathbb{F}_y := \pi_{\mathbb{E}}^{-1}(y)$, где $x, y \in \mathbb{M}$, мы понимаем диффеоморфизм, который представим в виде $q \circ p^{-1}$, где

$$\begin{aligned} \mathbb{F}_x \ni p: \mathbb{F} &\rightarrow \mathbb{F}_x, \\ \mathbb{F}_y \ni q: \mathbb{F} &\rightarrow \mathbb{F}_y. \end{aligned}$$

В частности, автоморфизм слоя $\mathbb{F}_x$ – это отображение вида $q \circ p^{-1}$, где $p, q \in \mathbb{F}_x$. В этом случае $q = pa$ для некоторого $a \in \mathbb{G}$, так что любой автоморфизм слоя $\mathbb{F}_x$ представим в виде $p \circ a \circ p^{-1}$, где p – произвольная фиксированная точка в том же слое $\mathbb{F}_x$. Поэтому группа автоморфизмов каждого слоя $\mathbb{F}_x$ изоморфна структурной группе $\mathbb{G}$.

Опишем несколько примеров ассоциированных расслоений.

Пример 12.2.1. Рассмотрим группу преобразований $(\mathbb{G}, \mathbb{G})$, когда $\mathbb{G}$ действует на себе справа. Тогда ассоциированное расслоение изоморфно исходному главному расслоению, $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}, \mathbb{G}, \mathbb{P}) \simeq \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. В общем случае они могут не совпадать, т.к. существует произвол в параметризации фактор пространства $\mathbb{E} = \mathbb{P} \times_{\mathbb{G}} \mathbb{F}$. Таким образом, произвольное главное расслоение ассоциировано само с собой. Если пару $(\mathbb{G}, \mathbb{G})$ рассматривать, как группу преобразований, действующую слева, то ассоциированное расслоение также будет изоморфно главному расслоению $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, но структурная группа будет действовать на $\mathbb{P}$ слева. $\square$

Пример 12.2.2 (Цилиндр, лист Мебиуса, бутылка Клейна). Рассмотрим главное расслоение $\mathbb{S}^1(\mathbb{S}^1, \pi, \mathbb{Z}_2)$, где пространство расслоения реализовано в виде единичной окружности $\mathbb{P} \approx \mathbb{S}^1$ на евклидовой плоскости $\mathbb{R}^2$ и структурная группа которого

состоит из двух элементов $\mathbb{Z}_2 = \{1, -1\}$ с обычным умножением. Действие структурной группы $\mathbb{Z}_2$ на пространстве расслоения определено отражением относительно центра окружности,

$$\mathbb{Z}_2 \ni \pm 1 : \quad \mathbb{S}^1 \ni \quad x \mapsto \pm x \quad \in \mathbb{S}^1, \quad x = (x^1, x^2) \in \mathbb{R}^2.$$

База расслоения $\pi(\mathbb{S}^1) = \mathbb{S}^1$ также является окружностью. Это ни что иное, как двулистное накрытие окружности окружностью, рассмотренное в примере 11.1.5. Построенное главное расслоение $\mathbb{S}^1(\mathbb{S}^1, \pi, \mathbb{Z}_2)$ нетривиально, т.к. не диффеоморфно прямому произведению $\mathbb{S}^1 \times \mathbb{Z}_2$.

Построим для этого главного расслоения три ассоциированных расслоения (см. рис.12.1).

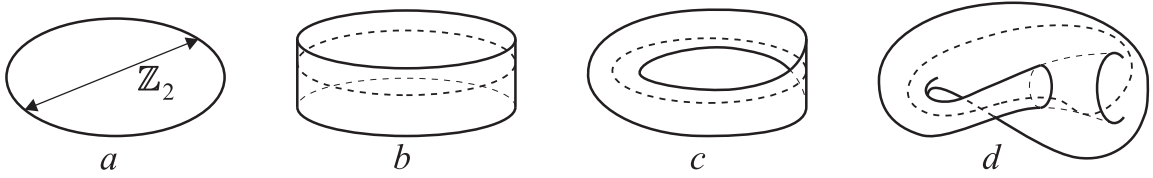


Рис. 12.1: Примеры ассоциированных расслоений. (a) Главное расслоение. (b) Цилиндр. (c) Лист Мёбиуса. (d) Бутылка Клейна.

1). Пусть группа $\mathbb{Z}_2$ действует на отрезок $\mathbb{F} = [-1, 1]$ тривиально,

$$\mathbb{Z}_2 \ni \pm 1 : \quad [-1, 1] \ni \quad v \mapsto v \quad \in [-1, 1].$$

Тогда ассоциированное расслоение представляет собой *цилиндр*, который является прямым произведением $\mathbb{S}^1 \times \mathbb{F}$.

2). Пусть группа $\mathbb{Z}_2$ действует на отрезок $\mathbb{F} = [-1, 1]$ нетривиально,

$$\mathbb{Z}_2 \ni \pm 1 : \quad [-1, 1] \ni \quad v \mapsto \pm v \quad \in [-1, 1].$$

Тогда ассоциированное расслоение представляет собой *лист Мёбиуса*. Лист Мёбиуса является нетривиальным расслоением, т.к. не имеет вида прямого произведения $\mathbb{S}^1 \times \mathbb{F}$.

3). Пусть группа $\mathbb{Z}_2$ действует на окружность $\mathbb{F} = \mathbb{S}^1$, отображая точки окружности относительно диаметра,

$$\mathbb{Z}_2 \ni \pm 1 : \quad \mathbb{S}^1 \ni \quad e^{iv} \mapsto e^{\pm iv} \quad \in \mathbb{S}^1,$$

где окружность $\mathbb{S}^1$ вложена в $\mathbb{C} = \mathbb{R}^2$. Тогда ассоциированное расслоение представляет собой *бутылку Клейна*.

В первых двух примерах типичным слоем является многообразие с краем. Все три ассоциированных расслоения имеют глобальные сечения, несмотря на то, что во втором и третьем случаях ассоциированные расслоения не имеют вида прямого произведения. На рисунках они показаны пунктирными линиями. $\square$

Приведенные примеры показывают, что многие хорошо известные примеры многообразий можно рассматривать, как ассоциированные расслоения. Приведем еще два примера, которые играют важную роль в дифференциальной геометрии.

Пример 12.2.3 (Касательное расслоение). Рассмотрим расслоение реперов $\mathbb{L}(\mathbb{M})$, $\dim \mathbb{M} = n$, определенное в разделе 5.4. Оно состоит из упорядоченных наборов (реперов) $p = \{x, e_a\}$, $a = 1, \dots, n$, где $x \in \mathbb{M}$ и $\{e_a\}$ – упорядоченный набор линейно независимых векторов касательного пространства, т.е. базис в $T_x(\mathbb{M})$. При этом проекция определена отображением $\pi(p) := x \in \mathbb{M}$. Общая линейная группа $\mathbb{GL}(n, \mathbb{R})$ действует в $\mathbb{L}(\mathbb{M})$ справа так. Если $a = \{S_a^b\} \in \mathbb{GL}(n, \mathbb{R})$ и $p = \{x, e_a\}$ – репер, то преобразованный репер $pa = \{x, e'_a\}$ есть, по определению, репер, состоящий из векторов $e'_a := S_a^b e_b$. Мы записываем правое действие структурной группы как матричное умножение слева из-за принятого нами соглашения суммирования “с десяти до четырех”. Отметим, что порядок записи является условным, $S_a^b e_b = e_b S_a^b$, т.к. суммирование проводится по одному нижнему и одному верхнему индексу. Группа $\mathbb{GL}(n, \mathbb{R})$ действует на расслоении реперов $\mathbb{L}(\mathbb{M})$ свободно и сохраняет слои, т.к. $\pi(p) = \pi(q)$, $p, q \in \mathbb{L}(\mathbb{M})$ тогда и только тогда, когда $p = qa$ для некоторого $a \in \mathbb{GL}(n, \mathbb{R})$.

Построим теперь ассоциированное расслоение. Пусть $(\mathbb{R}^n, \mathbb{GL}(n, \mathbb{R}))$ – группа преобразований. Группа $\mathbb{GL}(n, \mathbb{R})$, по определению, действует в $\mathbb{R}^n$ следующим образом. Пусть b_a , $a = 1, \dots, n$, – некоторый базис в $\mathbb{R}^n$. Если $a = \{S_a^b\} \in \mathbb{GL}(n, \mathbb{R})$ – матрица преобразования, то $b_a a := S_a^b b_b \in \mathbb{R}^n$. Покажем, что ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{R}^n, \mathbb{GL}(n, \mathbb{R}), \mathbb{L})$ можно отождествить с касательным расслоением $T(\mathbb{M})$, определенным в разделе 2.6.2. Действительно, базы расслоений совпадают, а слои изоморфны как векторные пространства одинаковой размерности. Покажем, что слои ассоциированного расслоения $\pi_{\mathbb{E}}^{-1}(x)$ для всех $x \in \mathbb{M}$ можно рассматривать, как касательные пространства $T_x(\mathbb{M})$. Пусть $p = \{x, e_a\}$ – репер в точке $x \in \mathbb{M}$. Он же является точкой в главном расслоении $p \in \mathbb{L}(\mathbb{M})$. Согласно предложению 12.2.1, каждому реперу ставится в соответствие невырожденное линейное отображение. Выберем такой репер, что

$$p : \quad \mathbb{R}^n \ni \quad b_a \mapsto e_a \quad \in \pi_{\mathbb{E}}^{-1}(x).$$

С другой стороны, по определению, набор векторов $\{e_a\}$ образует базис касательного пространства $T_x(\mathbb{M})$. Поэтому $\pi_{\mathbb{E}}^{-1}(x) = T_x(\mathbb{M})$.

В компонентах это отождествление выглядит следующим образом. Произвольный касательный вектор $X \in T_x(\mathbb{M})$ в точке $x \in \mathbb{M}$ можно разложить по координатному базису и реперу: $X = X^\alpha \partial_\alpha = X^a e_a$, где $e_a = e_\alpha^a \partial_\alpha$. Компоненты вектора X связаны линейным преобразованием $X^a = X^\alpha e_\alpha^a$, где e_α^a – компоненты обратного репера. Структурная группа $\mathbb{GL}(n, \mathbb{R})$ действует на латинские индексы, не меняя касательного вектора, $X = X^a e_a = X'^a e'_a$, где $X'^a = X^b S_b^{-1a}$ и $e'_a = S_a^b e_b$. При преобразовании координат преобразуются греческие индексы с помощью взаимно обратных матриц Якоби, $X = X^\alpha \partial_\alpha = X^{\alpha'} \partial_{\alpha'}$, где $X^{\alpha'} = X^\alpha \partial_\alpha x^{\alpha'}$ и $\partial_{\alpha'} = \partial_\alpha x^{\alpha'} \partial_\alpha$. В каждой точке $x \in \mathbb{M}$ матрицы Якоби также принадлежат общей линейной группе $\mathbb{GL}(n, \mathbb{R})$. $\square$

Пример 12.2.4 (Тензорные расслоения). Пусть T_s^r – пространство тензоров типа (r, s) над векторным пространством $\mathbb{R}^n$. Общая группа линейных преобразований $\mathbb{GL}(n, \mathbb{R})$ действует в T_s^r обычным образом:

$$\begin{aligned} e_{a_1} \otimes \dots \otimes e_{a_r} \otimes e^{b_1} \otimes \dots \otimes e^{b_s} &\mapsto \\ &\mapsto S_{a_1}^{c_1} e_{c_1} \otimes \dots \otimes S_{a_r}^{c_r} e_{c_r} \otimes e^{d_1} S_{d_1}^{-1b_1} \otimes \dots \otimes e^{d_s} S_{d_s}^{-1b_s}, \end{aligned}$$

где e_a – базис в $\mathbb{R}^n$, e^a – базис в сопряженное пространство $\mathbb{R}^{n*}$ и $\{S_a^b\} \in \mathbb{GL}(n, \mathbb{R})$. $a, b, \dots = 1, \dots, n$. Расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, T_s^r, \mathbb{GL}(n, \mathbb{R}), \mathbb{L})$, ассоциированное с расслоением реперов $\mathbb{L}(\mathbb{M})$, так же как и касательное расслоение, отождествляется с тензорным расслоением $T_s^r(\mathbb{M})$ типа (r, s) , определенным в разделе 2.7. $\square$

В разделе 5.1 дано определение векторного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{GL}(N, \mathbb{R}), \mathbb{P})$, $\dim \mathbb{V} = N$, ассоциированного с главным расслоением $\mathbb{P}(\mathbb{M}, \pi, \mathbb{GL}(N, \mathbb{R}))$. При этом размерность векторного пространства N не обязана совпадать с размерностью базы $n = \dim \mathbb{M}$. Дадим координатное описание векторных расслоений. Рассмотрим координатное покрытие базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$. Каждая область предполагается односвязной, и, следовательно, над каждой областью $\mathbb{U}_i$ расслоение является тривиальным, т.е. существуют диффеоморфизмы (тривиализации)

$$\chi_i : \pi_{\mathbb{E}}^{-1}(\mathbb{U}_i) \rightarrow \mathbb{U}_i \times \mathbb{V}.$$

Отсюда следует, что определено отображение

$$f_{ji} := \chi_j \circ \chi_i^{-1} : \mathbb{V} \rightarrow \mathbb{V},$$

которое называется *функциями склейки*. Если базис векторного пространства $\mathbb{V}$, $\dim \mathbb{V} = N$, фиксирован, то функции склейки принимают значения в группе невырожденных матриц:

$$f_{ji} : \mathbb{U}_i \cap \mathbb{U}_j \rightarrow \mathbb{GL}(N, \mathbb{R}). \quad (12.16)$$

Теорема 12.2.1. Пусть задано векторное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{GL}(N, \mathbb{R}), \mathbb{P})$, координатное покрытие базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и соответствующие тривиализации χ_i . Тогда однозначно определен набор функций склеек (12.16), обладающих свойствами:

$$f_{ij} = f_{ji}^{-1}, \quad \forall x \in \mathbb{U}_i \cap \mathbb{U}_j, \quad (12.17)$$

$$f_{ij} f_{jk} f_{ki} = \text{id}, \quad \forall x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k. \quad (12.18)$$

Обратно. По любому координатному покрытию $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и по любому набору отображений (12.16) со свойствами (12.17), (12.18) можно построить векторное расслоение $\mathbb{E}'(\mathbb{M}, \pi'_{\mathbb{E}}, \mathbb{V}, \mathbb{GL}(N, \mathbb{R}), \mathbb{P})$. Если набор функций склеек был построен по векторному расслоению $\mathbb{E}$, то расслоения $\mathbb{E}'$ и $\mathbb{E}$ изоморфны.

Доказательство. Аналогично доказательству теоремы 12.1.2 для главных расслоений. $\square$

Следствие. Главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G}(N, \mathbb{R}))$ и ассоциированные с ним векторные расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{G}(N, \mathbb{R}), \mathbb{P})$ определяют эквивалентные склеивающие коциклы.

Доказательство. См., например, [75], лекция 2, следствие 1. $\square$

Следующая конструкция показывает, что коциклы можно использовать для построения новых ассоциированных расслоений.

Предложение 12.2.2. Пусть дано векторное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V})$, ассоциированное с главным расслоением $\mathbb{P}(\mathbb{M}, \pi, \mathbb{GL}(N, \mathbb{R}))$, где $N = \dim \mathbb{V}$, с коциклом $\{f_{ji}\}$, соответствующим некоторому координатному покрытию базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$. Тогда существует детерминантное расслоение $|\mathbb{E}|(\mathbb{M}, \pi_{|\mathbb{E}|}, \mathbb{R})$ с одномерным слоем $\mathbb{R}$, определенное коциклом $\{\det f_{ji}\}$, которое ассоциировано с тем же главным расслоением $\mathbb{P}$.

Доказательство. Проверка свойств (12.17), (12.18). Кроме того, отображение $f_{ji} \mapsto \det f_{ji}$ определяет одномерное представление группы $\mathbb{GL}(N, \mathbb{R})$ в $\mathbb{R}$. $\square$

Пример 12.2.5. Пусть $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}) = \mathbb{T}(\mathbb{M})$ – касательное расслоение к многообразию $\mathbb{M}$. Тогда скалярные плотности степени 1, введенные в разделе 2.5, являются сечениями детерминантного расслоения. $\square$

Вернемся к общему случаю. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение, и $\mathbb{H}$ – замкнутая подгруппа в $\mathbb{G}$. Пара $(\mathbb{G}/\mathbb{H}, \mathbb{G})$, где $\mathbb{G}/\mathbb{H}$ – пространство правых смежных классов, является группой преобразований. Поэтому определено ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ со стандартным слоем $\mathbb{G}/\mathbb{H}$. С другой стороны, группа $\mathbb{H}$, будучи подгруппой в структурной группе $\mathbb{G}$, действует на $\mathbb{P}$ справа. Поэтому определено фактор пространство $\mathbb{P}/\mathbb{H}$. Между построенными ассоциированным расслоением и фактор пространством существует связь.

Предложение 12.2.3. *Расслоение $\mathbb{E} = \mathbb{P} \times_{\mathbb{G}} (\mathbb{G}/\mathbb{H})$, ассоциированное с $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, со стандартным слоем $\mathbb{G}/\mathbb{H}$ можно отождествить с $\mathbb{P}/\mathbb{H}$ так. Пусть $(p, \mathbb{H}a) \in \mathbb{P} \times (\mathbb{G}/\mathbb{H})$, где $a \in \mathbb{G}$, – представитель элемента ассоциированного расслоения $\mathbb{E}$. Отождествим его с элементом $pa \in \mathbb{P}$, который является представителем некоторого элемента из фактор пространства $\mathbb{P}/\mathbb{H}$.*

Доказательство. Каждый представитель однозначно определяет элемент соответствующего пространства. Поэтому отождествление представителей приводит к отождествлению точек ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ и фактор пространства $\mathbb{P}/\mathbb{H}$. Легко проверить, что это отождествление не зависит от выбора представителей. $\square$

Следствие. Четверка $\mathbb{P}(\mathbb{E}, \nu, \mathbb{H})$, где $\nu : \mathbb{P} \rightarrow \mathbb{P}/\mathbb{H}$ – естественная проекция, является главным расслоением, базой которого является ассоциированное расслоение $\mathbb{E}$, а структурной группой $\mathbb{H} \subset \mathbb{G}$.

Доказательство. Из предложения 12.2.3 следует, что фактор пространство $\mathbb{P}/\mathbb{H}$ является многообразием и поэтому может являться базой некоторого расслоения. Действие группы $\mathbb{H}$ на $\mathbb{P}$ не двигает точки базы $\mathbb{E}$ по построению. Локальная тривиальность $\mathbb{P}(\mathbb{E}, \nu, \mathbb{H})$ следует из локальной тривиальности расслоений $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ и $\mathbb{G}(\mathbb{G}/\mathbb{H}, \mathbb{H})$. Действительно, пусть $\mathbb{U}$ – окрестность в $\mathbb{M}$ такая, что $\pi^{-1}(\mathbb{U}) \approx \mathbb{U} \times (\mathbb{G}/\mathbb{H})$ и $\mathbb{V}$ – окрестность в $\mathbb{G}/\mathbb{H}$, для которой $\pi_{\mathbb{G}}^{-1}(\mathbb{V}) \approx \mathbb{V} \times \mathbb{H}$, где $\pi_{\mathbb{G}} : \mathbb{G} \rightarrow \mathbb{G}/\mathbb{H}$ – естественная проекция. Пусть $\mathbb{W}$ – открытое подмножество в $\pi_{\mathbb{E}}^{-1}(\mathbb{U}) \subset \mathbb{E}$, которое соответствует прямому произведению $\mathbb{U} \times \mathbb{V}$ при отождествлении $\pi^{-1}(\mathbb{U}) \approx \mathbb{U} \times (\mathbb{G}/\mathbb{H})$. Тогда $\nu^{-1}(\mathbb{W}) \approx \mathbb{W} \times \mathbb{H}$. $\square$

Выпишем связь размерностей многообразий, которые участвовали в приведенной выше конструкции:

$$\begin{aligned} \dim \mathbb{P}(\mathbb{M}, \pi, \mathbb{G}) &= \dim \mathbb{M} + \dim \mathbb{G}, \\ \dim \mathbb{G} &= \dim \mathbb{G}/\mathbb{H} + \dim \mathbb{H}, \\ \dim \mathbb{E} &= \dim \mathbb{M} + \dim \mathbb{G}/\mathbb{H}, \\ \dim \mathbb{P}(\mathbb{E}, \nu, \mathbb{H}) &= \dim \mathbb{M} + \dim \mathbb{G}/\mathbb{H} + \dim \mathbb{H} = \dim \mathbb{P}(\mathbb{M}, \pi, \mathbb{G}). \end{aligned}$$

Теперь обсудим возможность продолжения сечения ассоциированного расслоения, которое задано на некотором подмножестве базы $\mathbb{N} \subset \mathbb{M}$. Для формулировки теоремы нам понадобится

Определение. Отображение $f : \mathbb{N} \rightarrow \mathbb{E}$ подмножества $\mathbb{N}$ многообразия $\mathbb{M}$ в многообразии $\mathbb{E}$ называется *дифференцируемым на $\mathbb{N}$* , если для каждой точки $x \in \mathbb{N}$ существует дифференцируемое отображение $f_x : \mathbb{U}_x \rightarrow \mathbb{E}$, где $\mathbb{U}_x \subset \mathbb{M}$ – некоторая окрестность точки $x \in \mathbb{M}$, такое, что $f_x = f$ на пересечении $\mathbb{U}_x \cap \mathbb{N}$. Если задано ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ и некоторое подмножество базы $\mathbb{N} \subset \mathbb{M}$, то *сечением на $\mathbb{N}$* называется дифференцируемое отображение $\sigma : \mathbb{N} \rightarrow \mathbb{E}$ такое, что $\pi_{\mathbb{E}} \circ \sigma = \text{id}_{\mathbb{N}}$. $\square$

Замечание. В данном определении подмножество $\mathbb{N} \subset \mathbb{M}$ может быть произвольным и в общем случае не является подмногообразием в $\mathbb{M}$. $\square$

Пример 12.2.6. Пусть $f : \mathbb{U} \rightarrow \mathbb{E}$ – дифференцируемое отображение некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ в $\mathbb{E}$. Тогда сужение $f|_{\mathbb{N}}$ на любое подмножество $\mathbb{N} \subset \mathbb{U}$ является дифференцируемым на $\mathbb{N}$. $\square$

Теорема 12.2.2. Пусть $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ – ассоциированное расслоение такое, что типичный слой диффеоморфен евклидову пространству, $\mathbb{F} \approx \mathbb{R}^N$. Пусть $\mathbb{N}$ – замкнутое подмножество (возможно, пустое) в базе $\mathbb{M}$. Тогда любое сечение $\sigma : \mathbb{N} \rightarrow \mathbb{E}$, определенное на $\mathbb{N}$, может быть продолжено до глобального сечения, определенного на всем $\mathbb{M}$. В частности, если подмножество $\mathbb{N}$ пусто, то на ассоциированном расслоении $\mathbb{E}$ с типичным слоем $\mathbb{F} \approx \mathbb{R}^N$ существует глобальное сечение $\sigma : \mathbb{M} \rightarrow \mathbb{E}$.

Доказательство. Явно строится сечение, при этом используется паракомпактность многообразия. Детали можно найти, например, в [48], глава 1, теорема 5.7. $\square$

Следствие. На любом многообразии $\mathbb{M}$ существует векторное поле, например, тождественно равное нулю. Более того, если векторное поле задано на произвольном замкнутом подмножестве $\mathbb{N} \subset \mathbb{M}$, то его всегда можно продолжить до векторного поля на всем многообразии $\mathbb{M}$. Замкнутость подмножества $\mathbb{N}$ существенна. Пример 2.6.5 показывает, что, если векторное поле задано на открытом подмножестве, то продолжение может не существовать. $\square$

Замечание. Данная теорема показывает, что ассоциированные расслоения существенно отличаются от главных. Согласно предложению 12.1.1 главное расслоение $\mathbb{P}$ допускает глобальное сечение тогда и только тогда, когда оно тривиально. Если ассоциированное расслоение имеет вид прямого произведения, $\mathbb{E} = \mathbb{M} \times \mathbb{F}$, то оно, очевидно, также допускает глобальные сечения. Последняя теорема утверждает, что ассоциированное расслоение $\mathbb{E}$ имеет глобальные сечения даже если оно не имеет вида прямого произведения $\mathbb{M} \times \mathbb{F}$. Однако, взамен мы требуем, чтобы типичный слой был диффеоморфен евклидову пространству, $\mathbb{F} \approx \mathbb{R}^N$, некоторой размерности N . $\square$

Пример 12.2.7. Заменяем единичный отрезок $[-1, 1]$ для листа Мёбиуса из примера 12.2.2 на вещественную прямую $\mathbb{R}$. Тогда мы попадаем в условия теоремы 12.2.2, и, следовательно лист Мёбиуса имеет глобальное сечение. $\square$

Определение. Ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ называется *тривиальным*, если соответствующее главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ тривиально. $\square$

Из данного определения следует, что тривиальное ассоциированное расслоение всегда имеет вид прямого произведения, $\mathbb{E} \approx \mathbb{M} \times \mathbb{F}$, т.е. оно тривиально как расслоение. Обратное утверждение неверно.

Пример 12.2.8. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – нетривиальное главное расслоение и $\mathbb{F}$ – произвольное многообразие, на котором структурная группа $\mathbb{G}$ действует тривиально. Тогда ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ тривиально как расслоение, т.е. $\mathbb{E} \approx \mathbb{M} \times \mathbb{F}$, однако оно не является тривиальным ассоциированным расслоением. $\square$

В общем случае справедливо

Предложение 12.2.4. Векторное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathrm{GL}(N, \mathbb{R}), \mathbb{P})$, $\dim \mathbb{V} = N$, имеет вид прямого произведения $\mathbb{E} \simeq \mathbb{M} \times \mathbb{V}$ тогда и только тогда, когда оно имеет N сечений (того же класса гладкости), линейно независимых над каждой точкой базы.

Доказательство. См., например, [75], лекция 2, предложение 1. $\square$

12.3 Отображение расслоений

Рассмотрим два главных расслоения $\mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1)$ и $\mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$. Оба расслоения являются группами преобразований $(\mathbb{P}_1, \mathbb{G}_1)$ и $(\mathbb{P}_2, \mathbb{G}_2)$. Поэтому при определении гомоморфизма расслоений используется понятие эквивариантного отображения, введенного в разделе 9.4.

Определение. Гомоморфизмом расслоений $f : \mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1) \rightarrow \mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$ называется эквивариантное отображение $f = f_{\mathbb{P}} \times f_{\mathbb{G}}$ (пара отображений),

$$\begin{aligned} f_{\mathbb{P}} : \mathbb{P}_1 &\rightarrow \mathbb{P}_2, \\ f_{\mathbb{G}} : \mathbb{G}_1 &\rightarrow \mathbb{G}_2, \end{aligned}$$

где $f_{\mathbb{P}}$ – дифференцируемое отображение пространств расслоений и $f_{\mathbb{G}}$ – гомоморфизм структурных групп, таких, что $f_{\mathbb{P}}(pa) = f_{\mathbb{P}}(p)f_{\mathbb{G}}(a)$ для всех $p \in \mathbb{P}_1$ и $a \in \mathbb{G}_1$, т.е. диаграмма

$$\begin{array}{ccc} \mathbb{P}_1 \times \mathbb{G}_1 & \xrightarrow{f_{\mathbb{P}} \times f_{\mathbb{G}}} & \mathbb{P}_2 \times \mathbb{G}_2 \\ \downarrow a & & \downarrow f_{\mathbb{G}}(a) \\ \mathbb{P}_1 & \xrightarrow{f_{\mathbb{P}}} & \mathbb{P}_2 \end{array}$$

коммутативна.

Расслоения $\mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1)$ и $\mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$ *изоморфны*, если $f_{\mathbb{P}}$ – диффеоморфизм пространств расслоений и $f_{\mathbb{G}}$ – изоморфизм структурных групп. Если $\mathbb{P}_1 = \mathbb{P}_2 = \mathbb{P}$, то изоморфизм расслоений называется *автоморфизмом* расслоения $\mathbb{P}$. $\square$

Из приведенной диаграммы следует, что отображение $f_{\mathbb{P}} : \mathbb{P}_1 \rightarrow \mathbb{P}_2$ отображает каждый слой $p\mathbb{G}_1 \subset \mathbb{P}_1$ расслоения $\mathbb{P}_1$ в слой $f_{\mathbb{P}}(p)\mathbb{G}_2$ из $\mathbb{P}_2$ и поэтому индуцирует дифференцируемое отображение баз

$$f_{\mathbb{M}} : \mathbb{M}_1 \ni \pi_1(p) \mapsto \pi_2(f_{\mathbb{P}}(p)) \in \mathbb{M}_2.$$

Легко проверить, что это отображение не зависит от выбора точки в слое $p \in \pi_1^{-1}(x)$.

Если $f : \mathbb{P}_1 \rightarrow \mathbb{P}_2$ – изоморфизм расслоений, то индуцированное отображение баз $f_{\mathbb{M}}$ является диффеоморфизмом и существует обратный изоморфизм $f^{-1} : \mathbb{P}_2 \rightarrow \mathbb{P}_1$.

Пример 12.3.1 (Вертикальный автоморфизм). Пусть $a(x)$ – произвольная дифференцируемая функция на базе $\mathbb{M}$ со значениями в структурной группе $\mathbb{G}$. Умножим каждую точку главного расслоения $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ на $a(\pi(p))$, что соответствует повороту каждого слоя. В результате получим главное расслоение $\mathbb{P}'(\mathbb{M}, \pi, \mathbb{G})$, которое изоморфно исходному. Этот автоморфизм задается функциями:

$$\begin{aligned} f_{\mathbb{P}} : \mathbb{P} \ni p &\mapsto pa(\pi(p)) \in \mathbb{P}' \\ f_{\mathbb{G}} = \text{id}_{\mathbb{G}} : \mathbb{G} \ni b &\mapsto b \in \mathbb{G}. \end{aligned}$$

При этом отображение баз является тождественным преобразованием, $f_{\mathbb{M}} = \text{id}$. Этот автоморфизм называется *вертикальным*. $\square$

Определение. Гомоморфизм расслоений называется *вложением* или *инъекцией*, если индуцированное отображение баз $f_{\mathbb{M}} : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ есть вложение многообразий и $f_{\mathbb{G}} : \mathbb{G}_1 \rightarrow \mathbb{G}_2$ – мономорфизм (инъективный гомоморфизм). Отождествляя пространство первого расслоения $\mathbb{P}_1$ с его образом $f_{\mathbb{P}}(\mathbb{P}_1)$, $\mathbb{G}_1$ с $f_{\mathbb{G}}(\mathbb{G}_1)$ и $\mathbb{M}_1$ с $f_{\mathbb{M}}(\mathbb{M}_1)$, мы говорим, что $\mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1)$ есть *подрасслоение* для $\mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$. Если, кроме того, $\mathbb{M}_1 = \mathbb{M}_2 = \mathbb{M}$ и индуцированное отображение баз $f_{\mathbb{M}} = \text{id}_{\mathbb{M}}$ есть тождественное отображение, то гомоморфизм расслоений $f : \mathbb{P}_1(\mathbb{M}, \pi_1, \mathbb{G}_1) \rightarrow \mathbb{P}_2(\mathbb{M}, \pi_2, \mathbb{G}_2)$ называется *редукцией* структурной группы $\mathbb{G}_2$ главного расслоения $\mathbb{P}_2(\mathbb{M}, \pi_2, \mathbb{G}_2)$ к подгруппе $\mathbb{G}_1$. Само расслоение $\mathbb{P}_1(\mathbb{M}, \pi_1, \mathbb{G}_1)$ называется *редуцированным расслоением*. Если задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и подгруппа $\mathbb{H}$ в $\mathbb{G}$, то мы говорим, что *структурная группа $\mathbb{G}$ редуцируема к $\mathbb{H}$* , если существует редуцированное главное расслоение $\mathbb{P}'(\mathbb{M}, \pi', \mathbb{H})$. $\square$

Замечание. Поскольку при вложении главных расслоений $f_{\mathbb{M}}$ – вложение и $f_{\mathbb{G}}$ – мономорфизм, то

$$\dim \mathbb{M}_1 \leq \dim \mathbb{M}_2, \quad \text{и} \quad \dim \mathbb{G}_1 \leq \dim \mathbb{G}_2.$$

Кроме того, мы не требуем, чтобы подгруппа $\mathbb{H}$ была замкнутой в $\mathbb{G}$. Эта общность необходима в теории связностей. $\square$

Теорема 12.3.1. *Структурная группа $\mathbb{G}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ редуцируема к подгруппе Ли $\mathbb{H}$ тогда и только тогда, когда существует координатное покрытие базы $\mathbb{M} = \bigcup_i U_i$ такое, что все функции перехода a_{ji} принимают значение в $\mathbb{H}$.*

Доказательство. Доказательство проводится путем построения редуцированного расслоения. См., например, [48], глава I, предложение 5.3. $\square$

Приведем еще один критерий редуцируемости расслоений.

Теорема 12.3.2. *Структурная группа $\mathbb{G}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ редуцируема к подгруппе $\mathbb{H}$ тогда и только тогда, когда ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P}) = \mathbb{P}/\mathbb{H}$ (см. предложение 12.2.3) допускает глобальное сечение $\sigma : \mathbb{M} \rightarrow \mathbb{E}$. Между редуцированными главными расслоениями $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$ и сечениями σ существует естественное взаимно однозначное соответствие.*

Доказательство. См., например, [48], глава I, предложение 5.6. $\square$

Следствие. Структурная группа $\mathbb{G}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ редуцируема к единичному элементу $e \in \mathbb{G}$ тогда и только тогда, когда оно имеет глобальное сечение и, следовательно, тривиально, $\mathbb{P} \simeq \mathbb{M} \times \mathbb{G}$.

Замечание. Если структурная группа состоит только из единичного элемента, то пространство редуцированного главного расслоения $\mathbb{P}'(\mathbb{M}, \pi, e)$ естественным образом отождествляется с базой, $\mathbb{P}' = \mathbb{M}$. Поэтому существование отображения $f_{\mathbb{P}} : \mathbb{P}' = \mathbb{M} \rightarrow \mathbb{P}$ для редуцированного расслоения эквивалентно существованию глобального сечения для $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. $\square$

Теорема о существовании глобальных сечений на ассоциированном расслоении позволяет по новому доказать и взглянуть на существование римановой метрики на произвольном многообразии. Для формулировки результата нам понадобится дополнительное утверждение.

Из разложения Ивасава [76] следует, что любая связная группа Ли диффеоморфна прямому произведению, $\mathbb{G} \approx \mathbb{H} \times \mathbb{R}^n$, где $\mathbb{H}$ – максимальная компактная подгруппа в $\mathbb{G}$ и $\mathbb{R}^n$ – евклидово пространство размерности $n = \dim \mathbb{G} - \dim \mathbb{H}$ ([66], глава VI, теорема 2.2).

Предложение 12.3.1. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение со связной структурной группой Ли $\mathbb{G}$. Тогда структурная группа редуцируема к максимальной компактной подгруппе $\mathbb{H} \subset \mathbb{G}$.

Доказательство. Рассмотрим ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P}) = \mathbb{P}/\mathbb{H}$, где $\mathbb{H}$ – максимальная компактная подгруппа в $\mathbb{G}$. Поскольку типичный слой $\mathbb{G}/\mathbb{H}$ диффеоморфен евклидову пространству (разложение Ивасава), то, согласно теореме 12.2.2, ассоциированное расслоение допускает глобальное сечение. Поэтому из теоремы 12.3.2 вытекает, что структурная группа главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ редуцируема к максимальной компактной подгруппе $\mathbb{H}$. $\square$

Пример 12.3.2 (Существование римановой метрики). Пусть $\mathbb{L}(\mathbb{M})$ – расслоение линейных реперов над многообразием $\mathbb{M}$, $\dim \mathbb{M} = n$. Известно, что максимальной компактной подгруппой в группе общих линейных преобразований $\mathbb{GL}(n, \mathbb{R})$ является группа вращений $\mathbb{O}(n)$ размерности $n(n-1)/2$. В силу разложения Ивасава однородное пространство $\mathbb{GL}(n, \mathbb{R})/\mathbb{O}(n)$ диффеоморфно евклидову пространству $\mathbb{R}^d$ размерности

$$d = n^2 - \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

Из теоремы 12.2.2 следует, что ассоциированное расслоение $\mathbb{E} = \mathbb{L}(\mathbb{M})/\mathbb{O}(n)$ со слоем $\mathbb{GL}(n, \mathbb{R})/\mathbb{O}(n)$ имеет глобальное сечение. Поэтому структурная группа $\mathbb{GL}(n, \mathbb{R})$ может быть редуцирована к группе вращений $\mathbb{O}(n)$ как следствие теоремы 12.3.2.

Пусть в евклидовом пространстве $\mathbb{R}^n$ в декартовых координатах задано скалярное произведение

$$(\tilde{X}, \tilde{Y}) := \tilde{X}^a \tilde{Y}^b \delta_{ab}, \quad \tilde{X}, \tilde{Y} \in \mathbb{R}^n, \quad (12.19)$$

где δ_{ab} – евклидова метрика (1.8). Это скалярное произведение инвариантно относительно вращений евклидова пространства $\mathbb{O}(n)$.

Покажем, что каждая редукция структурной группы $\mathbb{GL}(n, \mathbb{R})$ расслоения реперов $\mathbb{L}(\mathbb{M})$ порождает риманову метрику g на $\mathbb{M}$. Пусть $\mathbb{O}(\mathbb{M}) := \mathbb{O}(\mathbb{M}, \pi, \mathbb{O}(n))$ – редуцированное подрасслоение для расслоения реперов $\mathbb{L}(\mathbb{M})$. Типичный слой редуцированного расслоения состоит из реперов вида $p = \{x, S_a^b e_b\}$, где e_b – некоторый фиксированный базис касательного пространства и $S_a^b \in \mathbb{O}(n)$ – произвольная матрица вращений. Если мы рассматриваем репер $p \in \mathbb{L}(\mathbb{M})$ как линейный изоморфизм

из $\mathbb{R}^n$ на касательное пространство $T_x(\mathbb{M})$, где $x = \pi(p)$, то каждый репер из редуцированного расслоения $p \in \mathcal{O}(\mathbb{M})$ определяет скалярное произведение в касательном пространстве $T_x(\mathbb{M})$ по формуле

$$g(X, Y) = (p^{-1}X, p^{-1}Y), \quad X, Y \in T_x(\mathbb{M}). \quad (12.20)$$

Инвариантность скалярного произведения (12.19) относительно вращений $\mathcal{O}(n)$ влечет за собой независимость $g(X, Y)$ от выбора репера $p \in \mathcal{Q}$. Таким образом, мы доказали, что каждое многообразие допускает риманову метрику.

В компонентах определение (12.20) имеет хорошо знакомый вид

$$g(X, Y) = X^\alpha Y^\beta g_{\alpha\beta} = \tilde{X}^a \tilde{Y}^b \delta_{ab}, \quad g_{\alpha\beta} := e_\alpha^a e_\beta^b \delta_{ab},$$

где $\tilde{X}^a = X^\alpha e_\alpha^a$ и $\tilde{Y}^a = Y^\alpha e_\alpha^a$. В таком виде скалярное произведение уже встречалось в разделе 6.9 (без знаков тильды).

Верно также обратное утверждение. Пусть на $\mathbb{M}$ задана риманова метрика g . Пусть $\mathcal{O}(\mathbb{M})$ – подмножество в расслоении реперов $\mathbb{L}(\mathbb{M})$, состоящее из реперов $p = \{x, e_a\}$, которые ортонормальны относительно g , т.е.

$$(e_a, e_b) = e_a^\alpha e_b^\beta g_{\alpha\beta} = \delta_{ab}.$$

Если репер $p \in \mathbb{L}(\mathbb{M})$ рассматривается как изоморфизм из $\mathbb{R}^n$ в $T_x(\mathbb{M})$, то p принадлежит редуцированному подрасслоению $\mathcal{O}(\mathbb{M})$ тогда и только тогда, когда $(\tilde{X}, \tilde{Y}) = g(p\tilde{X}, p\tilde{Y})$ для всех $\tilde{X}, \tilde{Y} \in \mathbb{R}^n$. Легко проверить, что $\mathcal{O}(\mathbb{M})$ образует редуцированное подрасслоение в $\mathbb{L}(\mathbb{M})$ над базой $\mathbb{M}$ со структурной группой $\mathcal{O}(n)$. Расслоение $\mathcal{O}(\mathbb{M})$ называется *расслоением ортонормальных реперов* над $\mathbb{M}$. Каждый элемент из $\mathcal{O}(\mathbb{M})$ есть *ортонормальный репер*.

Чтобы подчеркнуть принадлежность вектора евклидову пространству, в настоящем примере был использован знак тильды. Поскольку репер устанавливает изоморфизм $\mathbb{R}^n$ и $T_x(\mathbb{M})$, то эти пространства мы отождествим и опустим знак тильды, считая, что X^α и X^a – это компоненты одного касательного вектора $X \in T_x(\mathbb{M})$ относительно координатного базиса и репера. Чтобы их не путать, мы используем буквы греческого и латинского алфавитов. $\square$

В рассмотренном примере мы не только доказали существование римановой метрики на произвольном многообразии, но и следующее утверждение.

Теорема 12.3.3. *Существует взаимно однозначное соответствие между римановыми метриками на многообразии $\mathbb{M}$ и редуциямими структурной группы $\mathbb{GL}(n, \mathbb{R})$ расслоения реперов $\mathbb{L}(\mathbb{M})$ к группе вращений $\mathcal{O}(n)$*

В начале настоящего раздела мы установили, что гомоморфизм расслоений индуцирует дифференцируемое отображение баз расслоений. Теперь мы рассмотрим обратную задачу: в какой степени отображение некоторого многообразия в базу заданного расслоения может быть сопоставлено некоторому гомоморфизму расслоений? Ответ на этот вопрос дает

Теорема 12.3.4. *Пусть дано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и дифференцируемое отображение многообразий $f_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{M}$. Тогда существует единственное с точностью до изоморфизма главное расслоение $\mathbb{Q}(\mathbb{N}, \pi_{\mathbb{Q}}, \mathbb{G})$ с гомоморфизмом $f : \mathbb{Q} \rightarrow \mathbb{P}$, индуцирующим отображение баз $f_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{M}$ и соответствующим тождественному автоморфизму структурной группы $\mathbb{G}$.*

Доказательство. В прямом произведении $\mathbb{N} \times \mathbb{P}$ рассмотрим подмножество $\mathbb{Q}$, состоящее из точек $(y, p) \in \mathbb{N} \times \mathbb{P}$ таких, что $f_{\mathbb{N}}(y) = \pi(p)$. Определим действие структурной группы $\mathbb{G}$ справа на построенном множестве $\mathbb{Q}$:

$$\mathbb{G} \ni a : \quad \mathbb{Q} \ni \quad (y, p) \mapsto (y, p)a := (y, pa) \quad \in \mathbb{Q}.$$

Это отображение не зависит от точки слоя $\pi^{-1}(\pi(p))$. Нетрудно проверить, что $\mathbb{G}$ действует свободно на $\mathbb{Q}$ и что множество $\mathbb{Q}$ представляет собой главное расслоение $\mathbb{Q}(\mathbb{N}, \pi_{\mathbb{Q}}, \mathbb{G})$ с базой $\mathbb{N}$, структурной группой $\mathbb{G}$ и проекцией $\pi_{\mathbb{Q}} : (y, p) \mapsto y$.

Единственность. Пусть $\mathbb{Q}'(\mathbb{N}, \pi'_{\mathbb{Q}}, \mathbb{G})$ – другое главное расслоение с базой $\mathbb{N}$ и структурной группой $\mathbb{G}$ и $f' : \mathbb{Q}' \rightarrow \mathbb{P}$ – гомоморфизм, индуцирующий заданное отображение баз $f_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{M}$ и соответствующий тождественному автоморфизму структурной группы $\mathbb{G}$. Определим отображение $\mathbb{Q}'$ на $\mathbb{Q}$

$$\tilde{f} : \quad \mathbb{Q}' \ni \quad p' \mapsto (\pi'_{\mathbb{Q}}(p'), f'(p')) \quad \in \mathbb{Q}.$$

Тогда отображение $\tilde{f} : \mathbb{Q}' \rightarrow \mathbb{Q}$ – изоморфизм расслоений, индуцирующий тождественное преобразование базы $\mathbb{N}$ и соответствующий тождественному автоморфизму структурной группы $\mathbb{G}$. $\square$

Определение. Расслоение $\mathbb{Q}(\mathbb{N}, \pi_{\mathbb{Q}}, \mathbb{G})$ в утверждении теоремы 12.3.4 называется *расслоением, индуцированным отображением баз $f_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{M}$ из главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, или просто индуцированным расслоением*. Оно обозначается $f_{\mathbb{N}}^{-1}\mathbb{P}$. $\square$

Замечание. Если отображение $f_{\mathbb{N}}$ является вложением, то главное расслоение $\mathbb{Q}(\mathbb{N}, \pi_{\mathbb{Q}}, \mathbb{G})$ есть подрасслоение для $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. $\square$

Пример 12.3.3. Рассмотрим главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и некоторую окрестность базы $\mathbb{U} \subset \mathbb{M}$ тогда $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$ есть подрасслоение в $\mathbb{P}$. Оно же является индуцированным расслоением $f_{\mathbb{U}}^{-1}\mathbb{P}$, где $f : \mathbb{U} \rightarrow \mathbb{P}$ – вложение. $\square$

Глава 13

Связности на главных и ассоциированных расслоениях

Теория связностей на расслоениях играет исключительно важную роль в моделях математической физики, т.к. позволяет определить ковариантную производную. В свою очередь ковариантная производная используется для записи ковариантных уравнений. В настоящем разделе мы определим связность на главном и ассоциированных расслоениях. Покажем, что калибровочные поля Янга–Миллса представляют собой компоненты локальной формы связности. Кроме того, будет рассмотрена группа голономии, которая является одной из важнейших характеристик связности.

13.1 Связность на главном расслоении

Пусть задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с базой $\mathbb{M}$, $\dim \mathbb{M} = n$, и структурной группой Ли $\mathbb{G}$, $\dim \mathbb{G} = n$. Пара $(\mathbb{P}, \mathbb{G})$ представляет собой группу преобразований. При этом группа Ли $\mathbb{G}$ действует на многообразии $\mathbb{P}$ свободно. Каждый слой $\pi^{-1}(x)$, где $x \in \mathbb{M}$, диффеоморфен структурной группе и ее действие на нем транзитивно. Из предложения 9.2.1 следует, что действие структурной группы $\mathbb{G}$ на $\mathbb{P}$ индуцирует гомоморфизм μ алгебры Ли $\mathfrak{g}$ в алгебру Ли векторных полей $\mathcal{X}(\mathbb{P})$. Этот гомоморфизм является мономорфизмом, т.к. действие группы свободно и, следовательно, эффективно.

Определение. Векторное поле $X^* := \mu(X) \in \mathcal{X}(\mathbb{P})$, где $X \in \mathfrak{g}$ – произвольный элемент алгебры Ли структурной группы Ли $\mathbb{G}$, называется *фундаментальным векторным полем, соответствующим $X \in \mathfrak{g}$* . $\square$

Поскольку действие структурной группы на $\mathbb{P}$ отображает каждый слой в себя, то каждое фундаментальное векторное поле X_p^* касается слоя $\pi^{-1}(\pi(p))$ в каждой точке $p \in \mathbb{P}$. Поскольку действие группы свободно, то по предложению 9.2.1 фундаментальные векторные поля нигде не обращаются в нуль. Так как размерность каждого слоя равна размерности алгебры Ли $\mathfrak{g}$, то отображение

$$\mathfrak{g} \ni X \mapsto X_p^* \in \mathbb{T}_p(\mathbb{P})$$

есть линейный мономорфизм алгебры Ли $\mathfrak{g}$ в касательное пространство в точке $p \in \mathbb{P}$ к слою, проходящему через p .

Определение. Образ алгебры Ли $\mathfrak{g}$ в касательном пространстве $\mathbb{T}_p(\mathbb{P})$ называется *вертикальным подпространством* в $\mathbb{T}_p(\mathbb{P})$ и обозначается $\mathbb{V}_p(\mathbb{P})$. $\square$

Пусть $\{L_A\}$, $A = 1, \dots, N = \dim \mathfrak{g}$, – базис алгебры Ли $\mathfrak{g}$ с коммутационными соотношениями $[L_A, L_B] = f_{AB}^C L_C$, где f_{AB}^C – структурные константы группы Ли $\mathbb{G}$. Тогда индуцированные фундаментальные векторные поля L_A^* удовлетворяют тем же коммутационным соотношениям, $[L_A^*, L_B^*] = f_{AB}^C L_C^*$. Множество фундаментальных векторных полей $\{L_A^*\}$ задает дифференцируемое инволютивное распределение вертикальных подпространств $p \mapsto \mathbb{V}_p(\mathbb{P})$ на пространстве главного расслоения $\mathbb{P}$ (см. раздел 2.11). Согласно теореме Фробениуса 2.11.1 через каждую точку $p \in \mathbb{P}$ проходит интегральное подмногообразие этого распределения. Это интегральное подмногообразие есть ни что иное, как слой $\pi^{-1}(\pi(p))$.

Предложение 13.1.1. Пусть X^* – фундаментальное векторное поле, соответствующее элементу алгебры Ли $X = X_0^A L_A \in \mathfrak{g}$, $X_0^A = \text{const}$ для всех A . Тогда для каждого элемента группы $a \in \mathbb{G}$ векторное поле $r_{a*}X^*$, где r_{a*} – дифференциал отображения, индуцированного действием элемента a справа, является фундаментальным векторным полем, которое соответствует левоинвариантному векторному полю $\text{ad}(a^{-1})X = X_0^B S^{-1}_B{}^A(a) L_A \in \mathfrak{g}$, где ad обозначает присоединенное представление $\mathbb{G}$ в $\mathfrak{g}$ и $S_B{}^A(a)$ – матрица присоединенного представления.

Доказательство. Фундаментальное векторное поле X^* индуцируется однопараметрической группой преобразований $r_{b(t)}$, где $b(t) = \exp(tX)$. Векторное поле $r_{a*}X^*$ индуцируется однопараметрической группой преобразований $r_a \circ r_{b(t)} \circ r_{a^{-1}} = r_{a^{-1}b(t)a}$ по предложению 9.2.2. Утверждение предложения следует из того, что преобразования вида $a^{-1}b(t)a$ представляют собой однопараметрическую группу преобразований, порожденную элементом алгебры Ли $\text{ad}(a^{-1})X \in \mathfrak{g}$. $\square$

Пример 13.1.1 (Локальное рассмотрение). Чтобы прояснить абстрактное построение, которое было проведено выше, повторим его в компонентах. В настоящей главе мы будем возвращаться к этому примеру неоднократно. Пусть $U \subset \mathbb{M}$ – окрестность базы, для которой определено отображение (12.2), с координатами x^α , $\alpha = 1, \dots, n$. Используя диффеоморфизм $\chi: \pi^{-1}(U) \rightarrow U \times \tilde{\mathbb{G}}$, мы отождествим подрасслоение $\mathbb{Q} = \pi^{-1}(U)$ с прямым произведением $U \times \tilde{\mathbb{G}}$. Мы отметили структурную группу в прямом произведении знаком тильды, потому что в дальнейшем нам будет необходимо различать точку главного расслоения и точку структурной группы. То есть $p = (x, a)$, где $x \in U$, $a \in \tilde{\mathbb{G}}$, для всех $p \in \mathbb{Q}$. Действие структурной группы $\mathbb{G}$ на $\mathbb{Q}$ имеет вид

$$\mathbb{G} \ni a: \quad \mathbb{Q} \ni \quad p = (x, b) \mapsto pa = (x, ba) \quad \in \mathbb{Q}.$$

Фактически, данный пример относится к произвольному тривиальному главному расслоению, база которого покрывается одной картой, $\mathbb{M} \approx \mathbb{R}^n$.

Любой элемент алгебры Ли $\mathfrak{g}$ (левоинвариантное векторное поле X на группе Ли $\mathbb{G}$) имеет постоянные компоненты X_0^A относительно левоинвариантного базиса, $X = X_0^A L_A$, где L_A – базис алгебры Ли $\mathfrak{g}$ (см. раздел 8.2). Базису алгебры Ли L_A соответствуют фундаментальные векторные поля L_A^* на главном расслоении, которые образуют базис вертикальных подпространств. Выберем базис касательных пространств $\mathbb{T}_p(\mathbb{Q})$ в виде $\{\partial_\alpha, L_A^*\}$. Тогда фундаментальное векторное поле, соответствующее произвольному элементу алгебры Ли $X \in \mathfrak{g}$, будет иметь только вертикальные компоненты, $X^* = X_0^A L_A^*$. Утверждение предложения 13.1.1 сводится к равенству

$$r_{a*}X^* = X_0^A S_A{}^B(a^{-1}) L_B^*, \quad (13.1)$$

где $S_A^B(a^{-1}) = S_A^{-1B}(a)$ – матрица присоединенного представления для обратного элемента $a^{-1} \in \mathbb{G}$. Особенно просто это равенство проверяется вблизи единицы группы, где определена функция композиции. Пусть $X^*(e)$ – значение фундаментального векторного поля в единице группы. Тогда

$$r_{a*}(X^*(e)) = X_0^A R_A^B(a) \partial_B|_a = X_0^A R_A^B(a) L_B^{-1C}(a) L_C^*(a) = X_0^A S_A^B(a^{-1}) L_B^*(a),$$

где матрицы R_A^B , L_A^B и S_A^B были определены в разделах 8.2–8.4.

Произвольное векторное поле $\tilde{X} \in \mathcal{X}(\mathbb{Q})$ имеет вид

$$\tilde{X} = \tilde{X}^\alpha \partial_\alpha + \tilde{X}^A L_A^*.$$

После правого действия группы r_a , оно преобразуется в новое векторное поле

$$r_{a*}\tilde{X} = \tilde{X}^\alpha \partial_\alpha + \tilde{X}^B S_B^{-1A}(a) L_A^*, \quad (13.2)$$

т.к. дифференциал отображения r_{a*} действует только на вертикальные компоненты по правилу (13.1). $\square$

Продолжим общее построение.

Определение. *Связностью* Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ называется распределение

$$\Gamma : \mathbb{P} \ni p \mapsto \mathbb{H}_p(\mathbb{P}) \subset \mathbb{T}_p(\mathbb{P}),$$

которое каждой точке $p \in \mathbb{P}$ ставит в соответствие подпространство $\mathbb{H}_p(\mathbb{P})$ в касательном пространстве $\mathbb{T}_p(\mathbb{P})$ такое, что

- 1) в каждой точке p касательное пространство $\mathbb{T}_p(\mathbb{P})$ разлагается в прямую сумму:

$$\mathbb{T}_p(\mathbb{P}) = \mathbb{V}_p(\mathbb{P}) \oplus \mathbb{H}_p(\mathbb{P});$$

- 2) подпространства $\mathbb{H}_p(\mathbb{P})$ инвариантны относительно правого действия структурной группы:

$$r_{a*}\mathbb{H}_p(\mathbb{P}) = \mathbb{H}_{pa}(\mathbb{P}); \quad (13.3)$$

- 3) $\mathbb{H}_p(\mathbb{P})$ зависит дифференцируемо от p .

Множество касательных векторов $\mathbb{H}_p(\mathbb{P})$ в точке $p \in \mathbb{P}$ называется *горизонтальным подпространством* в $\mathbb{T}_p(\mathbb{P})$. Вектор $\tilde{X} \in \mathbb{T}_p(\mathbb{P})$ называется *вертикальным* или *горизонтальным*, если он лежит, соответственно, в $\mathbb{V}_p(\mathbb{P})$ или $\mathbb{H}_p(\mathbb{P})$. $\square$

Из условия 1) следует, что каждый вектор $\tilde{X} \in \mathbb{T}_p(\mathbb{P})$ может быть единственным образом представлен в виде суммы

$$\tilde{X} = \mathbf{v}\tilde{X} + \mathbf{h}\tilde{X}, \quad \mathbf{v}\tilde{X} \in \mathbb{V}_p(\mathbb{P}), \quad \mathbf{h}\tilde{X} \in \mathbb{H}_p(\mathbb{P}),$$

где $\mathbf{v}$ и $\mathbf{h}$ – проекторы на соответствующие подпространства. Векторы $\mathbf{v}\tilde{X}$ и $\mathbf{h}\tilde{X}$ называются, соответственно, *вертикальной* и *горизонтальной* компонентами касательного вектора $\tilde{X}$. Ясно, что

$$\dim \mathbb{V}_p(\mathbb{P}) = \dim \mathbb{G}, \quad \dim \mathbb{H}_p(\mathbb{P}) = \dim \mathbb{M}, \quad \forall p \in \mathbb{M}.$$

Если главное расслоение тривиально, то у него существует глобальное сечение. В этом случае условие 2) означает, что связность достаточно задать на каком-либо сечении, а затем разнести по всему пространству расслоения $\mathbb{P}$ с помощью группового

действия. В дальнейшем мы сформулируем теорему 13.1.1, определяющую связность на главном расслоении общего вида через семейство локальных форм связности, которые заданы на координатном покрытии базы.

По определению, условие 3) значит, что, если $\tilde{X}$ – дифференцируемое векторное поле на $\mathbb{P}$, то таковы же вертикальная и горизонтальная компоненты $v\tilde{X}$ и $h\tilde{X}$.

Пример 13.1.2. Накрытие $\tilde{\mathbb{M}} \rightarrow \mathbb{M}$ (см. раздел 11) является главным расслоением с 0-мерной структурной группой. В этом случае вертикальные подпространства отсутствуют, а горизонтальное подпространство в точке $p \in \tilde{\mathbb{M}}$ совпадает с касательным пространством $T_p(\mathbb{M})$. Это означает, что связность для накрытий единственна и распределение горизонтальных подпространств совпадает с касательным расслоением $T(\mathbb{M})$. $\square$

Пусть на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана связность Γ . Для того, чтобы конструктивно описать распределение горизонтальных подпространств построим на главном расслоении $\mathbb{P}$ 1-форму связности ω со значениями в алгебре Ли $\mathfrak{g}$, т.е. $\omega = \omega^A L_A$, где $\omega^A \in \Lambda_1(\mathbb{P})$ для всех A . Как было отмечено в начале настоящего раздела, каждому элементу алгебры Ли $X \in \mathfrak{g}$ соответствует единственное фундаментальное векторное поле X^* . При этом отображение $X \mapsto X_p^*$ представляет собой линейный мономорфизм $\mathfrak{g}$ в $T_p(\mathbb{P})$ для всех точек главного расслоения $p \in \mathbb{P}$.

Определение. Для каждого касательного вектора $\tilde{X}_p \in T_p(\mathbb{P})$ определим значение 1-формы $\omega(\tilde{X}_p)$ на векторе $\tilde{X}_p$, как единственный элемент алгебры Ли $X \in \mathfrak{g}$ такой, что фундаментальное векторное поле X_p^* в точке $p \in \mathbb{P}$ совпадает с вертикальной компонентой вектора, $X_p^* = v\tilde{X}_p$. 1-форма ω на $\mathbb{P}$ со значениями в алгебре Ли $\mathfrak{g}$ называется *формой связности* для заданной связности Γ на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. $\square$

По построению, $\omega(\tilde{X}) = 0$ тогда и только тогда, когда векторное поле $\tilde{X}$ горизонтально. В силу следующего предложения 1-форма ω взаимно однозначно определяет распределение горизонтальных подпространств и, следовательно, связность Γ на $\mathbb{P}$.

Предложение 13.1.2. Форма связности ω на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ имеет следующие свойства:

- 1) $\omega(X^*) = X$ для всех $X \in \mathfrak{g}$;
- 2) $r_a^* \omega = \text{ad}(a^{-1})\omega$, т.е.

$$(r_a^* \omega)(\tilde{X}) = \text{ad}(a^{-1})\omega(\tilde{X}) = \omega(\tilde{X})^B S_B^{-1A}(a) L_A$$

для всех $a \in \mathbb{G}$ и каждого векторного поля $\tilde{X}$ на $\mathbb{P}$.

Обратно. Если на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана 1-форма ω со значениями в алгебре Ли $\mathfrak{g}$, удовлетворяющая условиям 1) и 2), то на $\mathbb{P}$ существует единственная связность Γ такая, что ее форма связности есть ω .

Доказательство. Пусть ω – форма связности на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Условие 1) следует немедленно из определения ω . Поскольку каждое векторное поле $\tilde{X} \in \mathcal{X}(\mathbb{P})$ разлагается на вертикальную и горизонтальную составляющую, то достаточно проверить условие 2) для этих двух компонент. Если $\tilde{X}$ горизонтально, то $r_{a*} \tilde{X}$ также горизонтально для всех $a \in \mathbb{G}$, что следует из условия 2) в определении связности. Поэтому $(r_a^* \omega)_p(\tilde{X}) = \omega_{pa}(r_{a*} \tilde{X})$ и $\text{ad}(a^{-1})\omega_p(\tilde{X})$ одновременно обращаются в нуль. Если $\tilde{X}$ вертикально, то его значение в точке $p \in \mathbb{P}$ определяется некоторым фундаментальным векторным полем, $\tilde{X}_p = X_p^*$, для некоторого $X \in \mathfrak{g}$. Тогда из предложения 13.1.1

следует, что векторное поле $r_{a*}X^*$ соответствует элементу алгебры Ли $\text{ad}(a^{-1})X \in \mathfrak{g}$. Поэтому

$$(r_a^*\omega)_p(X^*) = \omega_{pa}(r_{a*}X^*) = \text{ad}(a^{-1})X = \text{ad}(a^{-1})(\omega_p(X^*)).$$

Обратно. Пусть задана 1-форма ω , удовлетворяющая условиям 1) и 2). Определим распределение горизонтальных подпространств

$$\mathbb{H}_p(\mathbb{P}) := \{\tilde{X} \in T_p(\mathbb{P}) : \omega(\tilde{X}) = 0\} \quad \forall p \in \mathbb{P}.$$

Теперь нетрудно проверить, что распределение $\Gamma : p \mapsto \mathbb{H}_p(\mathbb{P})$ является связностью, для которой ω – форма связности. $\square$

Пусть на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задано две связности Γ_1 и Γ_2 с формами связности ω_1 и ω_2 , соответственно. Нетрудно проверить, что 1-форма $(1-t)\omega_1 + t\omega_2$, где $t \in [0, 1]$, также задает некоторую связность на $\mathbb{P}$. Таким образом, любые две связности, заданные на одном главном расслоении, гомотопны.

Пример 13.1.3 (Локальное рассмотрение). Продолжим локальное построение, начатое в примере 13.1.1. У нас есть левоинвариантный базис $\{\partial_\alpha, L_A^*\}$ касательных пространств $T_p(\mathbb{Q})$ к главному подрасслоению $\mathbb{Q}(\mathbb{U}, \pi, \mathbb{G}) \subset \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Обозначим дуальный к нему базис 1-форм на $\mathbb{Q}$ через $\{dx^\alpha, \omega^{*A}\}$. Теперь построим форму связности ω на $\mathbb{Q}$. Произвольную 1-форму на $\mathbb{Q}$ со значениями в алгебре Ли $\mathfrak{g}$ можно разложить по этому базису,

$$\omega(x, a) = (dx^\alpha \omega_\alpha^A + \omega^{*B} \omega_B^A) L_A,$$

где $\omega_\alpha^A(x, a)$ и $\omega_B^A(x, a)$ – некоторые компоненты, зависящие от точки $(x, a) \in \mathbb{Q}$. Свойство 1) предложения 13.1.2 взаимно однозначно определяет компоненту ω_B^A :

$$\omega(X^*) = X \quad \Leftrightarrow \quad \omega_B^A(x, a) = \delta_B^A.$$

Нетрудно проверить, что 1-форма $\omega^{*A} L_A$ удовлетворяет свойству 2) предложения 13.1.2. Действительно,

$$r_a^*(\omega^{*A} L_A)(\tilde{X}) = (\omega^{*A} L_A)(r_{a*}\tilde{X}) = \tilde{X}^B S_B^{-1A}(a) L_A,$$

где мы учли действие дифференциала отображения на векторное поле (13.2). Отсюда следует, что форма $\omega^{*A} L_A$ действительно удовлетворяет условию 2):

$$r_a^*(\omega^{*A} L_A) = \omega^{*B} S_B^{-1A}(a) L_A.$$

Проведя аналогичные вычисления, получаем, что 1-форма со значениями в алгебре Ли $dx^\alpha \omega_\alpha^A L_A$ удовлетворяет свойству 2) предложения 13.1.2 тогда и только тогда, когда ее компоненты имеют вид

$$\omega_\alpha^A(x, b) = \overset{\circ}{A}_\alpha^B(x) S_B^{-1A}(b), \quad (13.4)$$

где $\overset{\circ}{A}_\alpha^B(x)$ – произвольные функции от $x \in \mathbb{U}$. Таким образом, форма связности на $\mathbb{Q}(\mathbb{U}, \pi, \mathbb{G}) = \mathbb{U} \times \mathbb{G}$ имеет вид

$$\omega(x, a) = [dx^\alpha \omega_\alpha^A(x, a) + \omega^{*A}(a)] L_A, \quad (13.5)$$

где компоненты ω_α^A определены равенством (13.4). Мы видим, что форма связности ω на главном расслоении $\mathbb{Q} = \mathbb{U} \times \tilde{\mathbb{G}}$ параметризуется $n \times N$ произвольными функциями $\overset{\circ}{A}_\alpha^B(x)$ на координатной окрестности $\mathbb{U} \subset \mathbb{M}$. Функции $\overset{\circ}{A}_\alpha^B(x)$ определяют компоненты формы связности на *нулевом локальном сечении*

$$\sigma_0 : \mathbb{U} \ni x \mapsto (x, e) \in \mathbb{U} \times \tilde{\mathbb{G}}. \quad (13.6)$$

Вектор $\tilde{X}_p = \tilde{X}_p^\alpha \partial_\alpha + \tilde{X}_p^A L_A^* \in \mathbb{T}_x(\mathbb{Q})$ горизонтален тогда и только тогда, когда выполнено равенство

$$\omega(\tilde{X}_p) = (\tilde{X}_p^\alpha \omega_\alpha^A + \tilde{X}_p^A) L_A = 0,$$

где ω_α^A имеет вид (13.4). Поэтому его вертикальная и горизонтальная составляющие равны

$$\begin{aligned} \mathbf{v}\tilde{X}_p &= (\tilde{X}_p^A + \tilde{X}_p^\alpha \omega_\alpha^A) L_A^*, \\ \mathbf{h}\tilde{X}_p &= \tilde{X}_p^\alpha \partial_\alpha - \tilde{X}_p^A \omega_\alpha^A L_A^*. \end{aligned} \quad (13.7)$$

Перемешивание компонент $\tilde{X}_p^\alpha$ и $\tilde{X}_p^A$ связано с тем, что векторные поля ∂_α в общем случае не являются горизонтальными. Они горизонтальны тогда и только тогда, когда $\omega_\alpha^A = 0 \Leftrightarrow \overset{\circ}{A}_\alpha^A = 0$. $\square$

Замечание. Из вида формы связности (13.5) следует, что она совпадает с канонической 1-формой θ на группе Ли (8.27), если база расслоения состоит из одной точки. В этом смысле форма связности ω на главном расслоении $\mathbb{P}$ является обобщением канонической 1-формы θ на группе Ли $\mathbb{G}$. $\square$

Проекция $\pi : \mathbb{P} \rightarrow \mathbb{M}$ соответствует дифференциал отображения $\pi_* : \mathbb{T}_p(\mathbb{P}) \rightarrow \mathbb{T}_x(\mathbb{M})$, который линейно отображает касательное пространство к главному расслоению в каждой точке $p \in \mathbb{P}$ в касательное пространство к базе в точке $x = \pi(p) \in \mathbb{M}$. Ядром дифференциала проекции π_* является вертикальное подпространство $\mathbb{V}_p(\mathbb{P})$. Действительно, любой вертикальный вектор касается некоторой кривой, целиком лежащей в слое. При проекции вся кривая отображается в одну точку x . Поэтому каждый вертикальный вектор отображается в нулевой вектор из $\mathbb{T}_x(\mathbb{M})$. Поскольку дифференциал проекции – это сюръективный гомоморфизм, то отсюда следует, что отображение горизонтального подпространства $\pi_* : \mathbb{H}_p(\mathbb{P}) \rightarrow \mathbb{T}_x(\mathbb{M})$ является изоморфизмом.

Определение. *Горизонтальным лифтом* или *подъемом* (или просто *лифтом*) векторного поля X на базе $\mathbb{M}$ называется единственное векторное поле $\tilde{X}$ на $\mathbb{P}$, которое горизонтально, $\tilde{X}_p \in \mathbb{H}_p(\mathbb{P})$, и проектируется на X , т.е. $\pi_*(\tilde{X}_p) = X_{\pi(p)}$ для всех $p \in \mathbb{P}$. $\square$

Предложение 13.1.3. *Если на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана связность Γ и на базе $\mathbb{M}$ задано векторное поле X , то существует единственный горизонтальный лифт $\tilde{X}$ векторного поля X . Лифт $\tilde{X}$ инвариантен относительно действия структурной группы, $\tau_{a*} \tilde{X}_p = \tilde{X}_{\tau_a p}$, для всех $a \in \mathbb{G}$. Обратно. Каждое горизонтальное векторное поле $\tilde{X}$ на $\mathbb{P}$, инвариантное относительно действия структурной группы, является горизонтальным лифтом некоторого векторного поля X на базе $\mathbb{M}$.*

Доказательство. См., например, [48], глава II, предложение 1.2. $\square$

Предложение 13.1.4. Пусть $\tilde{X}$ и $\tilde{Y}$ – горизонтальные лифты, соответственно, векторных полей X и Y , заданных на базе $\mathbb{M}$. Тогда:

- 1) $\tilde{X} + \tilde{Y}$ – горизонтальный лифт векторного поля $X + Y$;
- 2) для каждой функции $f \in C^\infty(\mathbb{M})$ произведение $\tilde{f}\tilde{X}$ есть горизонтальный лифт для $fX \in \mathcal{X}(\mathbb{M})$, где функция $f \in C^\infty(\mathbb{P})$ постоянна на слоях и определена равенством $\tilde{f} = f \circ \pi$;
- 3) горизонтальная компонента коммутатора $[\tilde{X}, \tilde{Y}]$ есть горизонтальный лифт коммутатора $[X, Y]$.

Доказательство. Первые два утверждения очевидны. Доказательство третьего просто:

$$\pi_*(h[\tilde{X}, \tilde{Y}]) = \pi_*([\tilde{X}, \tilde{Y}]) = [X, Y]. \quad \square$$

Пусть x^α , $\alpha = 1, \dots, \dim \mathbb{M}$ – система координат на некоторой окрестности базы $\mathbb{U} \subset \mathbb{M}$. Пусть D_α – горизонтальный лифт векторного поля ∂_α в $\pi^{-1}(\mathbb{U})$ для всех α . Тогда векторные поля $\{D_\alpha\}$ образуют локальный базис распределения $\Gamma : p \mapsto \mathbb{H}_p(\mathbb{P})$ в окрестности $\pi^{-1}(\mathbb{U}) \subset \mathbb{P}$.

Пример 13.1.4 (Локальное рассмотрение). Продолжим локальное построение, начатое в примере 13.1.1. Пусть $X(x) = X_x^\alpha \partial_\alpha$ – произвольное векторное поле на базе $\mathbb{M}$. Его горизонтальный лифт в $\mathbb{Q}$ имеет вид

$$\tilde{X}_p = X_x^\alpha \partial_\alpha - X_x^\alpha \omega_\alpha^A L_A^*, \quad x = \pi(p),$$

для всех $p \in \mathbb{Q}$. Вблизи единицы группы второе слагаемое имеет вид

$$X_x^\alpha \overset{\circ}{A}_\alpha^A S_A^{-1B} L_B^* = X_x^\alpha \overset{\circ}{A}_\alpha^A R_A^*,$$

где R_A^* правоинвариантные векторные поля на группе Ли $\tilde{\mathbb{G}}$ и мы воспользовались выражением матрицы присоединенного представления через производные от функции композиции (8.37). Отсюда сразу следует, что горизонтальное векторное поле $\tilde{X}$ инвариантно относительно действия структурной группы справа. Утверждения предложения 13.1.2 становятся тривиальными.

В частности, горизонтальный лифт координатного базиса ∂_α на $\mathbb{M}$ имеет вид

$$D_\alpha = \partial_\alpha - \omega_\alpha^A(p) L_A^* = \partial_\alpha - \overset{\circ}{A}_\alpha^B(x) S_B^{-1A}(a) L_A^* \quad (13.8)$$

для всех точек $p = (x, a) \in \mathbb{P}$. □

Теперь спустимся на базу и определим форму связности на $\mathbb{P}$ через семейство 1-форм на $\mathbb{M}$. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение с формой связности ω . Пусть $\mathbb{U} \subset \mathbb{M}$ – некоторая координатная окрестность базы, на которой задано локальное сечение $\sigma : \mathbb{U} \rightarrow \mathbb{P}$. Тогда форма связности ω определяет на $\mathbb{U}$ 1-форму с помощью возврата отображения σ^* .

Определение. 1-форма $\omega_\sigma := \sigma^* \omega$ со значениями в алгебре Ли $\mathfrak{g}$ структурной группы, $\omega_\sigma = \omega_\sigma^A L_A$, где $\omega_\sigma^A \in \Lambda_1(\mathbb{U})$ называется *локальной формой связности* на $\mathbb{U} \subset \mathbb{M}$. В компонентах она имеет вид

$$\omega_\sigma = dx^\alpha A_\alpha^A L_A, \quad (13.9)$$

где $A_\alpha^A(x)$ – некоторые функции на $\mathbb{U}$. □

Напомним, что, по определению, $\omega_\sigma(X) = \omega(\sigma_*X)$ для всех векторных полей на базе $X \in \mathcal{X}(\mathbb{M})$, где σ_* – дифференциал локального сечения.

Каждая форма связности ω на $\mathbb{P}$ однозначно определяет локальную форму связности ω_σ на $\mathbb{U} \subset \mathbb{M}$. Обратное, конечно, неверно, т.к. локальная форма связности ω_σ определена только на окрестности $\mathbb{U}$ и, вдобавок, зависит от сечения σ .

Пример 13.1.5 (Локальное рассмотрение). Рассмотрим зависимость локальной формы связности от сечения. Пусть на главном расслоении $\mathbb{Q} = \mathbb{U} \times \tilde{\mathbb{G}}$ задано два произвольных сечения σ и σ' . Поскольку две точки одного слоя связаны некоторым групповым преобразованием, то σ' можно выразить через первое сечение, $\sigma'(x) = \sigma(x)a(x)$, где $a(x) : \mathbb{U} \rightarrow \mathbb{G}$ – некоторая функция. Таким образом,

$$\sigma : x \mapsto (x, b(x)) \quad \text{и} \quad \sigma' : x \mapsto (x, b(x)a(x)). \quad (13.10)$$

Рассмотрим случай, когда все элементы структурной группы Ли $\mathbb{G}$ находятся в окрестности единицы группы, соответствующей локальной группе Ли (см. раздел 8.1). То есть каждый элемент группы имеет координаты, $a = \{a^\Lambda\} \in \mathbb{G}$, $\Lambda = 1, \dots, N$, и задана функция композиции $ba = f = \{f^\Lambda(b, a)\}$. Тогда действие дифференциалов сечений на векторное поле $X = X^\alpha \partial_\alpha \in \mathcal{X}(\mathbb{M})$ имеет вид

$$\begin{aligned} \sigma_*X &= X^\alpha \partial_\alpha + X^\alpha \partial_\alpha b^\Lambda \partial_\Lambda = X^\alpha \partial_\alpha + X^\alpha \partial_\alpha b^\Lambda L_\Lambda^{-1\mathbb{B}}(b) L_\mathbb{B}^*(b), \\ \sigma'_*X &= X^\alpha \partial_\alpha + X^\alpha \partial_\alpha f^\Lambda \partial_\Lambda = X^\alpha \partial_\alpha + X^\alpha \partial_\alpha f^\Lambda L_\Lambda^{-1\mathbb{B}}(f) L_\mathbb{B}^*(f), \end{aligned}$$

где использован явный вид формы связности (13.5). Локальные формы связности ω_σ и $\omega_{\sigma'}$ являются 1-формами на $\mathbb{U}$ со значениями в алгебре Ли $\mathfrak{g}$. Поэтому они представимы в виде

$$\begin{aligned} \omega_\sigma &= dx^\alpha A_\alpha^\Lambda(x) L_\Lambda, \\ \omega_{\sigma'} &= dx^\alpha A'_\alpha^\Lambda(x) L_\Lambda, \end{aligned} \quad (13.11)$$

где $A_\alpha^\Lambda(x)$ и $A'_\alpha^\Lambda(x)$ – компоненты этих форм (некоторые дифференцируемые функции от $x \in \mathbb{U}$). С другой стороны, по определению,

$$\begin{aligned} \omega_\sigma(X) &= \omega(\sigma_*X) = X^\alpha \left[\overset{\circ}{A}_\alpha^\mathbb{B} S_\mathbb{B}^{-1\Lambda}(b) + \partial_\alpha b^\mathbb{B} L_\mathbb{B}^{-1\Lambda}(b) \right] L_\Lambda, \\ \omega_{\sigma'}(X) &= \omega(\sigma'_*X) = X^\alpha \left[\overset{\circ}{A}_\alpha^\mathbb{B} S_\mathbb{B}^{-1\Lambda}(f) + \partial_\alpha f^\mathbb{B} L_\mathbb{B}^{-1\Lambda}(f) \right] L_\Lambda. \end{aligned} \quad (13.12)$$

Рассмотрим нулевое сечение $\sigma_0 = (x, e)$, где e – единица группы. Тогда

$$\omega_0 = dx^\alpha \overset{\circ}{A}_\alpha^\Lambda L_\Lambda.$$

То есть функции $\overset{\circ}{A}_\alpha^\Lambda(x)$, параметризующие форму связности в (13.5), представляют собой компоненты локальной формы связности для нулевого сечения. Тогда компоненты локальной формы связности ω_σ связаны с компонентами локальной формы связности для нулевого сечения простым соотношением

$$A_\alpha^\Lambda = \overset{\circ}{A}_\alpha^\mathbb{B} S_\mathbb{B}^{-1\Lambda}(b) + \partial_\alpha b^\mathbb{B} L_\mathbb{B}^{-1\Lambda}(b). \quad (13.13)$$

Сравнивая формулы (13.12), получим связь между компонентами локальных форм связности:

$$A'_\alpha^\Lambda = A_\alpha^\mathbb{B} S_\mathbb{B}^{-1\Lambda}(a) + \partial_\alpha f^\mathbb{B} L_\mathbb{B}^{-1\Lambda}(f) - \partial_\alpha b^\mathbb{B} L_\mathbb{B}^{-1\mathbb{C}}(b) S_\mathbb{C}^{-1\Lambda}(a), \quad (13.14)$$

где функция $a(x)$ связывает данные сечения: $\sigma' = \sigma a = (x, f = ba)$.

Полученная связь компонент двух локальных форм связности (13.14), заданных на одной окрестности базы $\mathbb{U} \subset \mathbb{M}$ неудобна в приложениях, т.к. содержит координатные функции $a^\Lambda(x)$, $b^\Lambda(x)$, f^Λ , которые в явном виде можно задать только в редких случаях. Чтобы обойти эту трудность выберем присоединенное представление алгебры Ли $\mathfrak{ad}(\mathfrak{g})$. То есть вместо локальных 1-форм ω_σ со значениями в алгебре Ли $\mathfrak{g}$ будем рассматривать 1-формы связности со значениями в присоединенном представлении алгебры Ли $\mathfrak{ad}(\mathfrak{g})$. Этому соответствует переход к матрицам

$$A_\alpha^\Lambda \mapsto A_{\alpha\mathbf{B}}^\mathbf{C} := -A_\alpha^\Lambda f_{\Lambda\mathbf{B}}^\mathbf{C},$$

где $f_{\Lambda\mathbf{B}}^\mathbf{C}$ – структурные константы группы Ли $\mathbb{G}$. Компоненты локальной формы связности в присоединенном представлении мы будем обозначать $A_\alpha = \{A_{\alpha\mathbf{B}}^\mathbf{C}\}$, опуская, для краткости, матричные индексы. Продифференцируем матрицу присоединенного представления $S_{\mathbf{A}}^\mathbf{B}(b)$,

$$\partial_\alpha S_{\mathbf{A}}^\mathbf{B} = \partial_\alpha b^\mathbf{C} \partial_{\mathbf{C}} S_{\mathbf{A}}^\mathbf{B} = \partial_\alpha b^\mathbf{C} L_{\mathbf{C}}^{-1\mathbf{D}} f_{\mathbf{D}\mathbf{A}}^\mathbf{E} S_{\mathbf{E}}^\mathbf{B},$$

где мы воспользовались формулой дифференцирования (8.41). Умножив это равенство справа на S^{-1} , получим равенство

$$\partial_\alpha S_{\mathbf{A}}^\mathbf{B} S_{\mathbf{B}}^{-1\mathbf{C}} = \partial_\alpha b^\mathbf{B} L_{\mathbf{B}}^{-1\mathbf{D}} f_{\mathbf{D}\mathbf{A}}^\mathbf{C}.$$

Тогда для компонент локальной формы связности (13.13) со значениями в присоединенном представлении алгебры Ли справедлива формула

$$A_\alpha = S \overset{\circ}{A}_\alpha S^{-1} + \partial_\alpha S S^{-1}, \quad (13.15)$$

где мы опустили матричные индексы. Отсюда следует правило преобразования компонент локальной формы связности при переходе от одного сечения $\sigma(x) : (x, b(x))$ к другому $\sigma'(x) : (x, b(x)a(x))$,

$$A'_\alpha = S A_\alpha S^{-1} + \partial_\alpha S S^{-1}, \quad (13.16)$$

где $S = (S_{\mathbf{A}}^\mathbf{B}(a))$ – матрица присоединенного представления. Это есть ни что иное, как хорошо известная формула калибровочного преобразования полей Янга–Миллса. Эта формула имеет явные преимущества по сравнению с (13.14), т.к. позволяет проводить вычисления с матрицами присоединенного представления, элементы которых зависят от точки базы $x \in \mathbb{M}$.

Вместо присоединенного можно рассматривать произвольное представление алгебры Ли $\mathfrak{g}$. Если $L_{\Lambda i}^j$ – матрицы, соответствующие левоинвариантным векторным полям L_Λ , и $A_{\alpha i}^j := A_\alpha^\Lambda L_{\Lambda i}^j$ – компоненты локальной формы связности в данном представлении, то формулы преобразования компонент при переходе между сечениями (13.16) сохраняются, если под S понимать матрицу соответствующего преобразования, $S = \{S_i^j(a)\}$. $\square$

Функции $a(x)$ можно рассматривать как функции, задающие вертикальный автоморфизм главного расслоения $\mathbb{Q} = \mathbb{U} \times \tilde{\mathbb{G}}$ (см. пример 12.3.1). Тогда формула (13.16) задает преобразование компонент локальной формы связности при вертикальном автоморфизме.

Отображение $\mathbb{U} \ni x \mapsto S_{\mathbf{A}}^\mathbf{B}(x) \in \mathfrak{ad}(\mathfrak{g})$ сопоставляет каждой точке базы матрицу присоединенного представления структурной группы. В формуле (13.16) каждое

слагаемое определяет 1-форму на $\mathbb{U}$ и не зависит явно от функции композиции. Поэтому, несмотря на то, что формулы преобразования компонент локальной формы связности были получены в окрестности единицы группы, они справедливы для произвольных локальных сечений на $\mathbb{U} \subset \mathbb{M}$.

Чтобы установить связь рассмотренной конструкции с понятиями, которые широко используются в математической физике, дадим

Определение. Калибровочным полем или полем Янга–Миллса $A_\alpha^A(x)$ на координатной окрестности $\mathbb{U} \subset \mathbb{M}$ называются компоненты локальной формы связности (13.9), после добавления соответствующих уравнений движения. Калибровочным преобразованием называется переход между двумя локальными сечениями σ и σ' на $\mathbb{U}$ или, что эквивалентно, преобразование компонент локальной формы связности при вертикальном автоморфизме. Структурная группа $\mathbb{G}$ называется калибровочной группой. В электродинамике структурной группой является абелева группа $\mathbb{U}(1)$, а компоненты локальной формы связности $A_\alpha(x)$ на $\mathbb{U}$, после добавления уравнений Максвелла, называются электромагнитным потенциалом. Конечно, когда мы говорим про калибровочные поля, то подразумеваем, что они удовлетворяют некоторой системе уравнений движения (уравнения Янга–Миллса или уравнения Максвелла). $\square$

Замечание. Форма связности ω на $\mathbb{P}$ определена инвариантным образом и ее компоненты $\omega_\alpha^A(p)$ в (13.5) являются тензорными полями на пространстве главного расслоения $\mathbb{P}$. Важное неоднородное слагаемое в калибровочном преобразовании (13.16) для компонент локальной формы связности $A_\alpha^A(x)$ возникает только при переходе к локальным сечениям. $\square$

Как уже было отмечено, локальная форма связности ω_σ не определяет форму связности ω и, следовательно, связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Однако, если задано координатное покрытие базы $\mathbb{M}$ и семейство локальных форм связности на каждой координатной окрестности, то это семейство однозначно определяет связность на $\mathbb{P}$. Опишем соответствующую конструкцию. Пусть задано координатное покрытие базы, $\mathbb{M} = \bigcup_i \mathbb{U}_i$, семейство изоморфизмов

$$\chi_i : \pi^{-1}(\mathbb{U}_i) \rightarrow \mathbb{U}_i \times \mathbb{G}$$

и соответствующие функции перехода

$$a_{ji} : \mathbb{U}_i \cap \mathbb{U}_j \rightarrow \mathbb{G}.$$

Это необходимо для однозначного определения главного расслоения в терминах функций перехода (теорема 12.1.2). На каждой окрестности $\mathbb{U}_i$ выберем сечение $\sigma_i : \mathbb{U}_i \rightarrow \mathbb{P}$, которое соответствует единичному элементу группы $e \in \mathbb{G}$, положив $\sigma_i(x) = \chi_i^{-1}(x, e)$. Пусть θ – (левоинвариантная $\mathfrak{g}$ -значная) каноническая 1-форма на $\mathbb{G}$, которая определена в разделе 8.2.

Для каждого непустого пересечения $\mathbb{U}_i \cap \mathbb{U}_j$ определим $\mathfrak{g}$ -значную 1-форму на $\mathbb{U}_i \cap \mathbb{U}_j$ с помощью возврата отображения a_{ji} ,

$$\theta_{ji} = a_{ji}^* \theta. \quad (13.17)$$

Для каждой окрестности $\mathbb{U}_i$ определим локальную форму связности на $\mathbb{U}_i$ с помощью возврата сечения σ_i ,

$$\omega_i = \sigma_i^* \omega,$$

где ω – форма связности на $\mathbb{P}$. Тогда справедлива

Теорема 13.1.1. *Формы θ_{ji} и ω_i удовлетворяют условиям:*

$$\omega_j = \text{ad}(a_{ji}^{-1})\omega_i + \theta_{ji}, \quad \text{на } \mathbb{U}_i \cap \mathbb{U}_j. \quad (13.18)$$

Обратно. Для каждого семейства локальных форм связности $\{\omega_i\}$, заданных на координатном покрытии $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и удовлетворяющих условиям (13.18) во всех пересечениях карт, существует единственная форма связности ω на $\mathbb{P}$, которая порождает семейство 1-форм $\{\omega_i\}$ вышеописанным образом.

Доказательство. См., например, [48], глава II, предложение 1.4. $\square$

Пример 13.1.6 (Локальное рассмотрение). Продолжим локальное рассмотрение, чтобы прояснить содержание последнего утверждения. Допустим, что мы находимся в окрестности единицы структурной группы. Тогда каноническая форма на группе Ли имеет вид $\theta = da^B L_B^{-1A} L_A$, где L_A – левоинвариантный базис алгебры Ли $\mathfrak{g}$. Пусть в пересекающихся картах $\mathbb{U}_i$ и $\mathbb{U}_j$ заданы координаты x^α и $x^{\alpha'}$, соответственно. Пусть

$$X = X^\alpha \partial_\alpha = X^{\alpha'} \partial_{\alpha'} \in \mathcal{X}(\mathbb{U}_i \cap \mathbb{U}_j)$$

– касательный вектор к базе на пересечении окрестностей. Тогда дифференциал функций перехода отображает этот вектор в касательное пространство группы

$$a_{ji*}X = X^\alpha \partial_\alpha a_{ji}^A \partial_A = X^\alpha \partial_\alpha a_{ji}^B L_B^{-1A} L_A.$$

Такая же формула имеет место в штрихованной системе координат. Поэтому $\mathfrak{g}$ -значная 1-форма (13.17) имеет вид

$$\theta_{ji} = dx^\alpha \partial_\alpha a_{ji}^B L_B^{-1A} (a_{ji}) L_A = dx^{\alpha'} \partial_{\alpha'} a_{ji}^B L_B^{-1A} (a_{ji}) L_A.$$

Пусть

$$\omega_i = dx^\alpha A_\alpha^A L_A \quad \text{и} \quad \omega_j = dx^{\alpha'} A_{\alpha'}^A L_A$$

– локальные формы связности, соответственно, на $\mathbb{U}_i$ и $\mathbb{U}_j$. Тогда формула (13.18) приводит к равенству

$$A_{\alpha'}^A = \partial_{\alpha'} x^\alpha [A_\alpha^B S_B^{-1A}(a_{ji}) + \partial_\alpha a_{ji}^B L_B^{-1A}(a_{ji})]. \quad (13.19)$$

Это преобразование калибровочного поля отличается от описанного ранее (13.13) только множителем $\partial_{\alpha'} x^\alpha$, соответствующим преобразованию координат в пересечении $\mathbb{U}_i \cap \mathbb{U}_j$. $\square$

Таким образом, для того, чтобы описать связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ в локальных терминах, необходимо задать координатное покрытие базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$, функции перехода $a_{ji}(x)$ в каждом непустом пересечении $\mathbb{U}_i \cap \mathbb{U}_j$, которые удовлетворяют условию (12.9) в областях пересечения трех карт, и семейство локальных форм связности (калибровочных полей), заданных на каждой координатной окрестности, такое, что в областях пересечения карт выполнено условие (13.19).

Замечание. В квантовой теории поля обычно рассматривают тривиальные главные расслоения $\mathbb{P} = \mathbb{R}^{1,3} \times \mathbb{G}$, где $\mathbb{R}^{1,3}$ – четырехмерное пространство Минковского и $\mathbb{G}$ – калибровочная группа. В этом частном случае локальная форма связности $dx^\alpha A_\alpha^A L_A$, заданная на всем пространстве-времени $\mathbb{R}^{1,3}$, взаимно однозначно определяет форму связности ω и, следовательно, связность Γ на $\mathbb{P}$. $\square$

В заключение данного раздела обсудим вопрос о существовании связностей. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение и $\mathbb{N} \subset \mathbb{M}$ – некоторое подмножество базы. Мы говорим, что связность определена над $\mathbb{N}$, если в каждой точке $p \in \pi^{-1}(\mathbb{N}) \subset \mathbb{P}$ из прообраза $\mathbb{N}$ определено горизонтальное подпространство $\mathbb{H}_p(\mathbb{P})$ таким образом, что выполнены первые два условия (прямая сумма и инвариантность относительно действия структурной группы) в определении связности и распределение $\mathbb{H}_p(\mathbb{P})$ дифференцируемо зависит от точки p в следующем смысле. Для каждой точки $x \in \mathbb{N}$ существует окрестность $\mathbb{U} \ni x$ и связность на $\pi^{-1}(\mathbb{U})$ такая, что ее сужение на $\pi^{-1}(x)$ совпадает с $\mathbb{H}_p(\mathbb{P})$ для всех точек слоя $p \in \pi^{-1}(x)$.

Теорема 13.1.2. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение и $\mathbb{N}$ – замкнутое подмножество (возможно, пустое) в базе $\mathbb{M}$. Любая связность, определенная над $\mathbb{N}$ может быть продолжена до связности на всем $\mathbb{P}$. В частности, любое главное расслоение $\mathbb{P}$ допускает связность.

Доказательство. Доказательство проводится путем явного построения связности. При этом существенно используется паракомпактность многообразия. См., например, [48], глава II, теорема 2.1. $\square$

Пример 13.1.7. Рассмотрим главное расслоение $\mathbb{P}(\mathbb{M}, \pi, e)$, структурная группа которого состоит из единственного элемента – единицы. Тогда на нем существует единственная связность. В этом случае вертикальные касательные пространства отсутствуют, а распределение горизонтальных пространств совпадает с касательным расслоением $\mathbb{T}(\mathbb{M})$. Форма связности и форма кривизны, которая будет определена в следующем разделе, при этом тождественно равны нулю. $\square$

Продолжение связности не является единственным.

Пример 13.1.8. Рассмотрим расслоение реперов $\mathbb{L}(\mathbb{R}^n)$ над евклидовым пространством $\mathbb{R}^n$. Поскольку $\mathbb{R}^n$ покрывается одной картой, то форма связности $\omega = dx^\alpha \omega_{\alpha a}^b e_b^a$, где e_b^a – базис алгебры Ли $\mathfrak{gl}(n, \mathbb{R})$, взаимно однозначно определяются своими компонентами $\omega_{\alpha a}^b$ на $\mathbb{L}$. Чтобы их задать, достаточно задать компоненты локальной формы связности на каком-либо сечении, а затем разнести их по всему пространству расслоения с помощью действия структурной группы. Если компоненты локальной формы связности $\omega_{\alpha a}^b(x)$ заданы на замкнутом подмножестве $\mathbb{N} \subset \mathbb{R}^n$, то их можно продолжить на все евклидово пространство $\mathbb{R}^n$ многими достаточно гладкими способами. Тем самым мы определим связность на всем $\mathbb{L}(\mathbb{R}^n)$. $\square$

13.2 Форма кривизны и структурное уравнение

Пусть задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ со связностью $\Gamma : \mathbb{P} \ni p \mapsto \mathbb{H}_p \subset \mathbb{T}_p(\mathbb{P})$. Каждая связность Γ на $\mathbb{P}$ взаимно однозначно определяет $\mathfrak{g}$ -значную 1-форму связности $\omega = \omega^A L_A$, где $\omega^A \in \Lambda_1(\mathbb{P})$ для всех $A = 1, \dots, N$, которой мы поставим в соответствие единственную 2-форму кривизны $R = R^A L_A$, где $R^A \in \Lambda_2(\mathbb{P})$ для всех A , со значениями в алгебре Ли $\mathfrak{g}$. Форма кривизны R является важнейшей характеристикой заданной связности Γ . Для ее определения нам понадобятся новые понятия.

Пусть задано представление ρ структурной группы Ли $\mathbb{G}$ в конечномерном векторном пространстве $\mathbb{V}$,

$$\rho : \mathbb{G} \ni a \mapsto \rho(a) \in \text{aut}(\mathbb{V}).$$

При этом $\rho(ab) = \rho(a) \circ \rho(b)$ для всех $a, b \in \mathbb{G}$.

Определение. r -форма $\varphi \in \Lambda_r(\mathbb{P})$ на главном расслоении $\mathbb{P}$ со значениями в векторном пространстве $\mathbb{V}$ называется *псевдотензориальной r -формой типа $(\rho, \mathbb{V})$* , если выполнено условие

$$r_a^* \varphi|_p = \rho(a^{-1}) \varphi|_{pa} \quad \forall a \in \mathbb{G} \quad \text{и} \quad \forall p \in \mathbb{P},$$

где r_a^* – возврат отображения $r_a : p \mapsto pa$. Псевдотензориальная r -форма φ называется *тензориальной r -формой типа $(\rho, \mathbb{V})$* , если она *горизонтальна*, т.е. $\varphi(\tilde{X}_1, \dots, \tilde{X}_r) = 0$, как только один из касательных векторов $\tilde{X}_i \in \mathcal{X}(\mathbb{P})$ вертикален. Под тензориальной 0-формой типа $(\rho, \mathbb{V})$ мы понимаем функцию $\varphi : \mathbb{P} \rightarrow \mathbb{V}$, удовлетворяющую условию $\varphi(p) = \rho(a^{-1}) \varphi(pa)$ или, что эквивалентно, $\varphi(pa) = \rho(a) \varphi(p)$. $\square$

Пример 13.2.1. В силу свойства 2) предложения 13.1.2 форма связности ω является псевдотензориальной 1-формой типа $(\text{ad}, \mathfrak{g})$. Она не является тензориальной, т.к. не обращается в нуль на вертикальных векторных полях. Для краткости, будем говорить, что форма типа $(\text{ad}, \mathfrak{g})$ является формой типа $\text{ad } \mathbb{G}$. $\square$

Пример 13.2.2. Пусть ρ_0 – тривиальное представление группы $\mathbb{G}$ в $\mathbb{V}$, т.е. $\rho_0(a)$ есть тождественное преобразование $\mathbb{V}$ для всех $a \in \mathbb{G}$. Тогда каждая тензориальная r -форма φ типа $(\rho_0, \mathbb{V})$ может быть взаимно однозначно представлена в виде $\varphi = \pi^* \varphi_{\mathbb{M}}$, где $\varphi_{\mathbb{M}}$ некоторая r -форма на базе $\mathbb{M}$ со значениями в $\mathbb{V}$. Для этого достаточно показать, что тензориальная форма φ однозначно определяет форму на базе $\varphi_{\mathbb{M}}$. Поскольку r -форма φ горизонтальна, то ее значения на произвольном наборе касательных векторных полей определяются только горизонтальными компонентами,

$$\varphi(\tilde{X}_1, \dots, \tilde{X}_r) = \varphi(\mathbf{h}\tilde{X}_1, \dots, \mathbf{h}\tilde{X}_r).$$

Так как r -форма φ правоинвариантна, то ее значения на правоинвариантных векторных полях не зависят от точки слоя $p \in \pi^{-1}(x)$. Поэтому r -форма φ взаимно однозначно определяет r -форму $\varphi_{\mathbb{M}}$ на базе следующим равенством

$$\varphi_{\mathbb{M}}(X_1, \dots, X_r) = \varphi(\tilde{X}_1, \dots, \tilde{X}_r),$$

где $\tilde{X}_1, \dots, \tilde{X}_r$ – произвольный набор правоинвариантных векторных полей таких, что $X_i = \pi_* \tilde{X}_i$ для всех $i = 1, \dots, r$. $\square$

Пример 13.2.3. Пусть $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{G}, \mathbb{P})$ – ассоциированное с $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ расслоение, типичным слоем которого является векторное пространство $\mathbb{V}$, в котором задано представление ρ структурной группы $\mathbb{G}$. Тогда тензориальную r -форму φ типа $(\rho, \mathbb{V})$ можно рассматривать как сопоставление каждой точке $x \in \mathbb{M}$ мультилинейного антисимметричного отображения

$$\tilde{\varphi}_x : \underbrace{\mathbb{T}_x(\mathbb{M}) \times \dots \times \mathbb{T}_x}_{r} \rightarrow \pi_{\mathbb{E}}^{-1}(x) \quad (13.20)$$

с помощью равенства

$$\tilde{\varphi}_x(X_1, \dots, X_r) := p(\varphi(\tilde{X}_1, \dots, \tilde{X}_r)), \quad X_i \in \mathbb{T}_x(\mathbb{M}),$$

где p – любая точка слоя $\pi^{-1}(x)$ и $\tilde{X}_i$ – произвольный касательный вектор в p , для которого $\pi_* \tilde{X}_i = X_i$ для каждого i . Подробнее, $\varphi(\tilde{X}_1, \dots, \tilde{X}_r)$ есть тогда элемент стандартного слоя $\mathbb{V}$, а p – линейное отображение из $\mathbb{V}$ на слой $\pi_{\mathbb{E}}^{-1}(x)$. Поэтому

$p(\varphi(\tilde{X}_1, \dots, \tilde{X}_r))$ является элементом слоя $\pi_{\mathbb{E}}^{-1}(x)$. Нетрудно проверить, что этот элемент не зависит от выбора точки p и вектора $\tilde{X}_i$ в слое $\pi^{-1}(x)$.

Обратно. Пусть задано мультилинейное антисимметричное отображение (13.20) для всех $x \in \mathbb{M}$. Тогда можно определить тензориальную r -форму типа $(\rho, \mathbb{V})$ на $\mathbb{P}$ следующим образом

$$\varphi(\tilde{X}_1, \dots, \tilde{X}_r) := p^{-1}(\tilde{\varphi}_x(\pi_*\tilde{X}_1, \dots, \pi_*\tilde{X}_r)), \quad \tilde{X}_i \in \mathbb{T}_p(\mathbb{P}),$$

где $x = \pi(p)$. В частности, тензориальная 0-форма типа $(\rho, \mathbb{V})$, т.е. функция $\varphi : \mathbb{P} \rightarrow \mathbb{V}$, удовлетворяющая условию $\varphi(pa) = \rho(a)\varphi(p)$, может быть отождествлена с сечением ассоциированного расслоения $\sigma_{\mathbb{E}} : \mathbb{M} \rightarrow \mathbb{E}$. Несколько специальных случаев данного примера будет использовано в дальнейшем. $\square$

Продолжим обсуждение связности Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Пусть $\mathbb{V}_p(\mathbb{P})$ и $\mathbb{H}_p(\mathbb{P})$ – вертикальное и горизонтальное подпространства в $\mathbb{T}_p(\mathbb{P})$ и $\mathbf{h} : \mathbb{T}_p(\mathbb{P}) \rightarrow \mathbb{H}_p(\mathbb{P})$ – проекция на горизонтальное подпространство. Введем новое понятие внешней ковариантной производной от псевдотензориальной r -формы, с помощью которой будет определена кривизна.

Предложение 13.2.1. Пусть φ – псевдотензориальная r -форма на $\mathbb{P}$ типа $(\rho, \mathbb{V})$. Тогда:

1) r -форма $\varphi\mathbf{h}$, определенная равенством

$$\varphi\mathbf{h}(\tilde{X}_1, \dots, \tilde{X}_r) := \varphi(\mathbf{h}\tilde{X}_1, \dots, \mathbf{h}\tilde{X}_r), \quad \tilde{X}_i \in \mathbb{T}_p(\mathbb{P}), \quad i = 1, \dots, r,$$

есть тензориальная форма типа $(\rho, \mathbb{V})$;

2) $d\varphi$ есть псевдотензориальная $(r+1)$ -форма типа $(\rho, \mathbb{V})$;

3) $(r+1)$ -форма $D\varphi$, определенная как $D\varphi := (d\varphi)\mathbf{h}$, является тензориальной формой типа $(\rho, \mathbb{V})$.

Доказательство. Распределение горизонтальных векторных полей инвариантно относительно действия структурной группы, поэтому $r_a \circ \mathbf{h} = \mathbf{h} \circ r_a$. Отсюда следует, что $\varphi\mathbf{h}$ является псевдотензориальной r -формой типа $(\rho, \mathbb{V})$. Очевидно, что

$$\varphi\mathbf{h}(\tilde{X}_1, \dots, \tilde{X}_r) = 0,$$

если один из касательных векторов $\tilde{X}_i$ вертикален, и, следовательно, форма $\varphi\mathbf{h}$ горизонтальна. Второе утверждение следует из равенства $r_a^* \circ d = d \circ r_a^*$ для всех $a \in \mathbb{G}$ (3.45). Третье утверждение является прямым следствием двух первых. $\square$

Определение. Форма $D\varphi := (d\varphi)\mathbf{h}$ называется внешней ковариантной производной от псевдотензориальной r -формы φ на $\mathbb{P}$. Оператор D называется внешним ковариантным дифференцированием. $\square$

Пример 13.2.4 (Локальное рассмотрение). Выпишем ковариантную производную от тензориальной 0-формы на $\mathbb{Q} = \mathbb{U} \times \tilde{\mathbb{G}}$ типа $(\rho, \mathbb{V})$ в компонентах. Пусть e_i , $i = 1, \dots, \dim \mathbb{V}$, – базис векторного пространства $\mathbb{V}$ и

$$\rho : \mathbb{G} \ni a \mapsto S_i^j(a) \in \text{aut } \mathbb{V}$$

– представление структурной группы. Пусть $L_{\Lambda j}^i$ – представление генераторов (базиса) L_{Λ} структурной группы. Тогда справедливы формулы (8.67) и (8.68). Поскольку φ – тензориальная 0-форма типа $(\rho, \mathbb{V})$, то она имеет вид

$$\varphi(p) = \varphi^i(p)e_i = \tilde{\varphi}^j(x)S_j^{-1i}(b)e_i, \quad p = (x, b) \in \mathbb{Q},$$

где $\overset{\circ}{\varphi}(x, 0) = \overset{\circ}{\varphi}^i(x)e_i$ – значение этой функции на нулевом сечении $\sigma_0(x) = (x, 0) \in \mathbb{P}$. Внешний дифференциал от компонент φ^i равен

$$\begin{aligned} d\varphi^i &= dx^\alpha \partial_\alpha \overset{\circ}{\varphi}^j S_j^{-1i} + db^\Lambda \overset{\circ}{\varphi}^j \partial_\Lambda S_j^{-1i} = dx^\alpha \partial_\alpha \overset{\circ}{\varphi}^j S_j^{-1i} + \omega^{*\Lambda} \overset{\circ}{\varphi}^j L_\Lambda^* S_j^{-1i} = \\ &= dx^\alpha \partial_\alpha \overset{\circ}{\varphi}^j S_j^{-1i} + \omega^{*\Lambda} \overset{\circ}{\varphi}^j S_j^{-1k} L_{\Lambda k}{}^i. \end{aligned}$$

Поскольку горизонтальная составляющая вектора имеет вид (13.7), то значение внешней ковариантной производной на произвольном векторном поле $\tilde{X} \in \mathbb{X}(\mathbb{P})$ равно

$$D\varphi^i(\tilde{X}) = d\varphi^i(\mathbf{h}\tilde{X}) = \tilde{X}^\alpha \overset{\circ}{\nabla}_\alpha \overset{\circ}{\varphi}^j S_j^{-1i}, \quad (13.21)$$

где введены обозначения

$$\overset{\circ}{\nabla}_\alpha \overset{\circ}{\varphi}^j := \partial_\alpha \overset{\circ}{\varphi}^j + \overset{\circ}{\varphi}^k \overset{\circ}{A}_{\alpha k}{}^j, \quad \overset{\circ}{A}_{\alpha k}{}^j := -\overset{\circ}{A}_\alpha{}^\Lambda L_{\Lambda k}{}^j.$$

Теперь спустимся на базу $\mathbb{U}$. Рассмотрим два произвольных сечения

$$\sigma : x \mapsto (x, b(x)), \quad \text{и} \quad \sigma' : x \mapsto (x, b(x)a(x)),$$

связанных калибровочным преобразованием, которое параметризуется функцией $a(x)$. Обозначим значение компонент $\varphi^i(p)$ на этих сечениях через

$$\varphi^i(x) = \varphi^i|_\sigma, \quad \text{и} \quad \varphi'^i(x) = \varphi^i|_{\sigma'}. \quad (13.22)$$

Эти функции на $\mathbb{U}$ связаны калибровочным преобразованием

$$\varphi'^i = \varphi^j S_j^{-1i}(a), \quad a = a(x). \quad (13.23)$$

Внешний ковариантный дифференциал $D\varphi$ после проектирования на базу с помощью возвратов отображений σ^* и σ'^* принимает вид

$$\begin{aligned} \sigma^*(D\varphi^i) &= dx^\alpha \nabla_\alpha \varphi^i, \\ \sigma'^*(D\varphi^i) &= dx^\alpha \nabla'_\alpha \varphi'^i, \end{aligned} \quad (13.24)$$

где

$$\begin{aligned} \nabla_\alpha \varphi^i &:= \partial_\alpha \varphi^i + \varphi^j A_{\alpha j}{}^i, \\ \nabla'_\alpha \varphi'^i &:= \partial_\alpha \varphi'^i + \varphi'^j A'_{\alpha j}{}^i. \end{aligned} \quad (13.25)$$

Выше мы ввели обозначения: $A_{\alpha j}{}^i := -A_\alpha{}^\Lambda L_{\Lambda j}{}^i$ и $A'_{\alpha j}{}^i := -A'_\alpha{}^\Lambda L_{\Lambda j}{}^i$, где $A_\alpha{}^\Lambda$ и $A'_\alpha{}^\Lambda$ – компоненты локальных форм связности (13.11). Нетрудно проверить, что при калибровочном преобразовании (13.23), (13.15) ковариантная производная ведет себя ковариантно:

$$\nabla'_\alpha \varphi'^i = \nabla_\alpha \varphi^j S_j^{-1i}. \quad (13.26)$$

Этого следовало ожидать, т.к. определение ковариантной производной было дано в инвариантном виде. Таким образом, мы видим, что инвариантное определение внешней ковариантной производной для тензориальной 0-формы типа $(\rho, \mathbb{V})$ совпадает с обычным определением ковариантной производной в теории калибровочных полей. В примере 13.2.3 было показано, что тензориальную 0-форму типа $(\rho, \mathbb{V})$ можно отождествить с сечением ассоциированного расслоения, типичным слоем которого является векторное пространство $\mathbb{V}$. Таким образом, формулы (13.25) определяют ковариантные производные от сечений $\sigma_\mathbb{E} : \mathbb{M} \rightarrow \mathbb{E}$ ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_\mathbb{E}, \mathbb{V}, \mathbb{G}, \mathbb{P})$. $\square$

В примере 13.2.1 было отмечено, что форма связности ω на $\mathbb{P}$ есть псевдотензорная 1-форма типа $\mathfrak{ad} \mathbb{G}$. Используя предложение 13.2.1, дадим

Определение. Внешняя ковариантная производная $R := D\omega$ от формы связности ω является тензорной 2-формой на $\mathbb{P}$ типа $\mathfrak{ad} \mathbb{G}$ и называется *формой кривизны* для формы связности ω . Если L_A базис алгебры Ли, то $R = R^A L_A$, где $R^A \in \Lambda_2(\mathbb{P})$ для всех $A = 1, \dots, N$. $\square$

Теорема 13.2.1 (Структурное уравнение). Пусть ω – форма связности на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и $R := D\omega$ – ее форма кривизны. Тогда

$$d\omega(\tilde{X}, \tilde{Y}) = -\frac{1}{2} [\omega(\tilde{X}), \omega(\tilde{Y})] + R(\tilde{X}, \tilde{Y}) \quad (13.27)$$

для всех $\tilde{X}, \tilde{Y} \in \mathbb{T}_p(\mathbb{P})$ и $p \in \mathbb{P}$.

Доказательство. Каждый вектор в $\mathbb{T}_p(\mathbb{P})$ единственным образом разлагается в сумму вертикального и горизонтального векторов. Так как каждый член в (13.27) билинеен и антисимметричен по $\tilde{X}$ и $\tilde{Y}$, то достаточно проверить равенство (13.27) в трех случаях.

1) $\tilde{X}_p$ и $\tilde{Y}_p$ горизонтальны для всех $p \in \mathbb{P}$. В этом случае $\omega(\tilde{X}) = \omega(\tilde{Y}) = 0$ и равенство (13.27) сводится к определению формы кривизны R , т.к. $D\omega = d\omega$ для горизонтальных векторных полей.

2) $\tilde{X}_p$ и $\tilde{Y}_p$ вертикальны для всех $p \in \mathbb{P}$. В этом случае их значения в точке p соответствуют некоторым фундаментальным векторным полям, т.е. $\tilde{X}_p = X_p^*$ и $\tilde{Y}_p = Y_p^*$ для некоторых $X, Y \in \mathfrak{g}$. Из формулы (3.37) следует равенство

$$\begin{aligned} 2d\omega(X^*, Y^*) &= X^*(\omega(Y^*)) - Y^*(\omega(X^*)) - \omega([X^*, Y^*]) = \\ &= -[X, Y] = -[\omega(X^*), \omega(Y^*)], \end{aligned}$$

т.к. $\omega(X^*) = X$, $\omega(Y^*) = Y$ и $[X^*, Y^*] = [X, Y]^*$. Поскольку для фундаментальных векторных полей $R(X^*, Y^*) = 0$, то формула (13.27) в рассматриваемом случае имеет место.

3) $\tilde{X}_p$ горизонтально, $\tilde{Y}_p$ вертикально для всех $p \in \mathbb{P}$. Продолжим $\tilde{X}_p \in \mathbb{H}_p(\mathbb{P})$ до горизонтального векторного поля $\tilde{X}$ на $\mathbb{P}$. Это всегда возможно в силу следствия теоремы 12.2.2. Пусть $\tilde{Y}_p = Y_p^*$ для некоторого $Y \in \mathfrak{g}$. В рассматриваемом случае правая часть равенства (13.27) равна нулю, поэтому достаточно доказать равенство $d\omega(\tilde{X}, Y^*) = 0$. Из формулы (3.37) следует, что

$$2d\omega(\tilde{X}, Y^*) = \tilde{X}(\omega(Y^*)) - Y^*(\omega(\tilde{X})) - \omega([\tilde{X}, Y^*]) = -\omega([\tilde{X}, Y^*]).$$

Теперь достаточно доказать следующее утверждение

Лемма 13.2.1. Пусть $\tilde{X}$ – горизонтальное векторное поле и Y^* – фундаментальное векторное поле, соответствующее элементу алгебры $Y \in \mathfrak{g}$. Тогда коммутатор $[\tilde{X}, Y^*]$ горизонтален.

Доказательство. Фундаментальное векторное поле Y^* индуцировано действием $r_{a(t)}$, где $a(t)$ – 1-параметрическая подгруппа в $\mathbb{G}$, порожденная элементом алгебры

$Y \in \mathfrak{g}$ (экспоненциальное отображение). Поскольку коммутатор векторных полей совпадает с производной Ли, то из (2.124) следует равенство

$$[\tilde{X}, Y^*] = -L_{Y^*}\tilde{X} = -\lim_{t \rightarrow 0} \frac{\tilde{X} - r_{a(t)*}\tilde{X}}{t}.$$

Если векторное поле $\tilde{X}$ горизонтально, то $r_{a(t)*}\tilde{X}$ тоже горизонтально. Поэтому коммутатор $[\tilde{X}, Y^*]$ горизонтален. $\square$

Следствие. Если $\tilde{X}$ и $\tilde{Y}$ – горизонтальные векторные поля, то

$$\omega([\tilde{X}, \tilde{Y}]) = -2R(\tilde{X}, \tilde{Y}). \quad \square \quad (13.28)$$

Доказательство. Для горизонтальных векторных полей $\omega(\tilde{X}) = \omega(\tilde{Y}) = 0$ и

$$2d\omega(\tilde{X}, \tilde{Y}) = -\omega([\tilde{X}, \tilde{Y}])$$

как следствие (3.37). $\square$

Структурное уравнение (13.27) называют также *структурным уравнением Картана* и часто для простоты записывают в виде

$$d\omega = -\frac{1}{2}[\omega, \omega] + R. \quad (13.29)$$

Приведем еще одну форму записи структурного уравнения. Пусть $L_A, A = 1, \dots, N$, – базис алгебры Ли $\mathfrak{g}$ с коммутационными соотношениями

$$[L_A, L_B] = f_{AB}^C L_C,$$

где f_{AB}^C – структурные константы алгебры. Тогда формы связности и кривизны можно разложить по базису, $\omega = \omega^A L_A$ и $R = R^A L_A$, и структурные уравнения принимают вид

$$d\omega^A = -\frac{1}{2}\omega^B \wedge \omega^C f_{BC}^A + R^A. \quad (13.30)$$

Замечание. Структурное уравнение (13.27) отличается от формулы Маурера–Картана для групп Ли (8.23) только слагаемым с кривизной. Для фундаментальных векторных полей $R(X^*, Y^*) = 0$ и формулы просто совпадают. $\square$

Теперь спустимся на базу. Пусть задано локальное сечение $\sigma : \mathbb{U} \rightarrow \mathbb{P}$ на некоторой окрестности $\mathbb{U} \subset \mathbb{M}$.

Определение. 2-форма на $\mathbb{U}$ со значениями в алгебре Ли $\mathfrak{g}$, определенная возвратом сечения, $R_\sigma := \sigma^* R$, где R – форма кривизны на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, называется *локальной формой кривизны* формы связности ω . $\square$

В компонентах локальная форма кривизны имеет вид

$$R_\sigma = F^A L_A = \frac{1}{2} dx^\alpha \wedge dx^\beta F_{\alpha\beta}^A L_A, \quad (13.31)$$

где $F^A = \frac{1}{2} dx^\alpha \wedge dx^\beta F_{\alpha\beta}^A \in \Lambda_2(\mathbb{U})$ для всех $A = 1, \dots, N$.

Пример 13.2.5 (Локальное рассмотрение). В настоящем примере мы выразим компоненты локальной формы кривизны через компоненты локальной формы связности. Чтобы это сделать, сначала определим компоненты формы кривизны на $\mathbb{Q} = \mathbb{U} \times \tilde{\mathbb{G}}$ через компоненты формы связности. Внешняя производная от формы связности (13.5) имеет вид

$$\begin{aligned} d\omega &= \left[\frac{1}{2} dx^\alpha \wedge dx^\beta (\partial_\alpha \omega_\beta^A - \partial_\beta \omega_\alpha^A) - dx^\alpha \wedge db^B \partial_B \omega_\alpha^A + d\omega^{*A} \right] L_A = \\ &= \left[\frac{1}{2} dx^\alpha \wedge dx^\beta (\partial_\alpha \omega_\beta^A - \partial_\beta \omega_\alpha^A) + dx^\alpha \wedge \omega^{*B} \omega_\alpha^C f_{BC}^A - \frac{1}{2} \omega^{*B} \wedge \omega^{*C} f_{BC}^A \right] L_A, \end{aligned}$$

где $\omega_\alpha^A(x, b) = \overset{\circ}{A}_\alpha^B(x) S_B^{-1A}(b)$. Кроме того мы воспользовались формулой Маурера–Картана (8.24) для структурной группы и правилом дифференцирования матрицы присоединенного представления (8.42). Теперь нетрудно вычислить значение формы кривизны на векторных полях

$$D\omega(\tilde{X}, \tilde{Y}) = d\omega(\mathbf{h}\tilde{X}, \mathbf{h}\tilde{Y}) = \tilde{X}^\alpha \tilde{Y}^\beta R_{\alpha\beta}^A L_A,$$

где

$$R_{\alpha\beta}^A = \partial_\alpha \omega_\beta^A - \partial_\beta \omega_\alpha^A - \omega_\alpha^B \omega_\beta^C f_{BC}^A \quad (13.32)$$

– компоненты формы кривизны. Эти компоненты можно выразить через компоненты, заданные на нулевом сечении

$$R_{\alpha\beta}^A(x, a) = \overset{\circ}{F}_{\alpha\beta}^B(x) S_B^{-1A}(a),$$

где

$$\overset{\circ}{F}_{\alpha\beta}^A = \partial_\alpha \overset{\circ}{A}_\beta^A - \partial_\beta \overset{\circ}{A}_\alpha^A - \overset{\circ}{A}_\alpha^B \overset{\circ}{A}_\beta^C f_{BC}^A \quad (13.33)$$

– компоненты тензора кривизны для нулевого сечения. Таким образом, форма кривизны имеет вид

$$R = \frac{1}{2} dx^\alpha \wedge dx^\beta R_{\alpha\beta}^A L_A, \quad (13.34)$$

где компоненты определены равенствами (13.32).

Теперь спустимся на базу $\mathbb{U}$. Для сечений (13.22) получаем следующие выражения для локальных форм кривизны

$$\begin{aligned} R_\sigma &:= \sigma^* R = \frac{1}{2} dx^\alpha \wedge dx^\beta F_{\alpha\beta}^A L_A, \\ R_{\sigma'} &:= \sigma'^* R = \frac{1}{2} dx^\alpha \wedge dx^\beta F'_{\alpha\beta}{}^A L_A, \end{aligned}$$

где

$$F_{\alpha\beta}^A = \partial_\alpha A_\beta^A - \partial_\beta A_\alpha^A - A_\alpha^B A_\beta^C f_{BC}^A \quad (13.35)$$

и такое же выражение для $F'_{\alpha\beta}{}^A$ через штрихованные компоненты A_α^A . Нетрудно проверить, что компоненты локальной формы кривизны преобразуются ковариантным образом,

$$F'_{\alpha\beta}{}^A = F_{\alpha\beta}^B S_B^{-1A}, \quad (13.36)$$

при калибровочном преобразовании (13.19).

Переходя к присоединенному представлению

$$F_{\alpha\beta}^A \mapsto F_{\alpha\beta B}^C := -F_{\alpha\beta}^A f_{AB}^C,$$

получаем следующее выражение для локальной формы кривизны

$$F_{\alpha\beta A}{}^B = \partial_\alpha A_{\beta A}{}^B - \partial_\beta A_{\alpha A}{}^B - A_{\alpha A}{}^C A_{\beta C}{}^B + A_{\beta A}{}^C A_{\alpha C}{}^B. \quad (13.37)$$

Это выражение совпадает с выражением для локальной формы кривизны в аффинной геометрии (5.57), которое было получено ранее, после замены $A_{\alpha A}{}^B \mapsto \omega_{\alpha a}{}^b$. Тем самым мы показали, что аффинная связность, которая была введена ранее независимым образом, является частным случаем связности на главном расслоении общего вида. Выражение для компонент локальной формы кривизны (13.37) можно записать в виде

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha - [A_\alpha, A_\beta], \quad (13.38)$$

где мы, для краткости, опустили матричные индексы и квадратные скобки обозначают коммутатор матриц. При калибровочном преобразовании (13.16) компоненты локальной формы кривизны преобразуются ковариантно:

$$F'_{\alpha\beta} = S F_{\alpha\beta} S^{-1},$$

как и следовало ожидать.

Форма кривизны играет важную роль в приложениях. Обращение в нуль ее компонент дает критерий локальной тривиальности связности. Действительно, при доказательстве локальной тривиальности линейной связности в разделе 5.5 конкретный вид структурной группы не был использован. Поэтому справедлива

Теорема 13.2.2. *Пусть в некоторой односвязной области $U \subset M$ заданы компоненты локальной формы связности $A_{\alpha A}{}^B$. Если соответствующая локальная форма кривизны равна нулю на U , то существует такое калибровочное преобразование, после которого компоненты локальной формы связности обратятся в нуль, возможно, в меньшей окрестности. Или, существует такая матрица калибровочного преобразования S , что компоненты локальной формы связности имеют вид чистой калибровки*

$$A_\alpha = \partial_\alpha S S^{-1}, \quad (13.39)$$

где мы, для краткости, опустили матричные индексы.

При проведении вычислений на пространстве главного расслоения $\mathbb{P}$, например, в моделях типа Калуцы–Клейна, в касательном расслоении $T(\mathbb{P})$ удобно использовать базис $\{D_\alpha, L_A^*\}$, состоящий из горизонтальных векторных полей

$$D_\alpha := \partial_\alpha - \omega_\alpha{}^A L_A^*, \quad (13.40)$$

и фундаментальных векторных полей L_A^* . Этот базис неголономен:

$$[D_\alpha, D_\beta] = -R_{\alpha\beta}{}^A L_A^*, \quad (13.41)$$

$$[D_\alpha, L_A^*] = 0, \quad (13.42)$$

$$[L_A^*, L_B^*] = f_{AB}{}^C L_C^*, \quad (13.43)$$

где $R_{\alpha\beta}{}^A$ – компоненты формы кривизны (13.32). Второе коммутационное соотношение (13.42) является следствием инвариантности распределения горизонтальных подпространств относительно действия группы справа (напомним, что левоинвариантные векторные поля генерируют действие группы справа, а правоинвариантные

– слева, раздел 8.3). Заметим, что ковариантную производную $D\varphi$ от тензориальной 0-формы типа $(\rho, \mathbb{V})$ можно записать в виде

$$D\varphi = dx^\alpha (D_\alpha \varphi^i) e_i,$$

где векторное поле D_α действует как дифференцирование. Это следует из определения (13.21). В приведенной формуле $\varphi^i = \varphi^i(x, a)$ в отличие от формул (13.25), где ковариантная производная берется от сечений $\varphi^i = \varphi^i(x, \sigma(x))$. Кроме того, справедливо равенство

$$[D_\alpha, D_\beta] \varphi^i = \varphi^j R_{\alpha\beta j}{}^i, \quad (13.44)$$

где $R_{\alpha\beta j}{}^i := R_{\alpha\beta}{}^\Lambda L_{\Lambda j}{}^i$. Эта формула является аналогом формулы (6.93), полученной в аффинной геометрии. В аффинной геометрии в правой части стоит дополнительное слагаемое с тензором кручения. $\square$

Для связи с моделями математической физики дадим

Определение. Компоненты локальной формы кривизны $F_{\alpha\beta}{}^\Lambda$ называются *напряженностью калибровочного поля* или *напряженностью поля Янга–Миллса*. В электродинамике калибровочной группой является абелева группа $\mathbb{U}(1)$, а компоненты локальной формы кривизны $F_{\alpha\beta}$ называются *напряженностью электромагнитного поля*. $\square$

Продолжим общее рассмотрение.

Теорема 13.2.3 (Тождества Бианки). Пусть на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана форма связности ω . Тогда форма кривизны $R := D\omega$ удовлетворяет тождествам Бианки:

$$DR = 0, \quad (13.45)$$

где D – внешняя ковариантная производная.

Доказательство. Из определения внешней ковариантной производной следует, что $DR(\tilde{X}, \tilde{Y}, \tilde{Z}) = 0$, если хотя бы один из векторов $\tilde{X}, \tilde{Y}, \tilde{Z}$ вертикален. Поэтому достаточно доказать, что $dR(\tilde{X}, \tilde{Y}, \tilde{Z}) = 0$, когда все три вектора горизонтальны. Возьмем внешнюю производную от структурного уравнения (13.30):

$$0 = dd\omega^\Lambda = -\frac{1}{2}d\omega^\mathbb{B} \wedge \omega^\mathbb{C} f_{\mathbb{B}\mathbb{C}}{}^\Lambda + \frac{1}{2}\omega^\mathbb{B} \wedge d\omega^\mathbb{C} f_{\mathbb{B}\mathbb{C}}{}^\Lambda + dR^\Lambda.$$

Поскольку $\omega^\Lambda(\tilde{X}) = 0$, если вектор $\tilde{X}$ горизонтален, то

$$dR^\Lambda(\tilde{X}, \tilde{Y}, \tilde{Z}) = 0$$

если все три вектора горизонтальны. $\square$

Теорема 13.2.4. Пусть ω – форма связности на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и φ – произвольная тензориальная 1-форма типа $\text{ad } \mathbb{G}$. Тогда

$$D\varphi(\tilde{X}, \tilde{Y}) = d\varphi(\tilde{X}, \tilde{Y}) + \frac{1}{2}[\varphi(\tilde{X}), \omega(\tilde{Y})] + \frac{1}{2}[\omega(\tilde{X}), \varphi(\tilde{Y})]$$

для всех $\tilde{X}, \tilde{Y} \in \mathbb{T}_p(\mathbb{P})$ и $p \in \mathbb{P}$.

Доказательство. Аналогично доказательству теоремы 13.2.1 о структурном уравнении. См., например, [48], глава II, предложение 5.5. $\square$

Пример 13.2.6 (Локальное рассмотрение). Запишем тождества Бианки в компонентах. Внешняя производная от компонент формы кривизны (13.34) имеет вид

$$\begin{aligned} dR^A &= \frac{1}{2} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \partial_\alpha R_{\beta\gamma}{}^A + \frac{1}{2} dx^\alpha \wedge dx^\beta \wedge \omega^{*B} L_B^* R_{\alpha\beta}{}^A = \\ &= \frac{1}{2} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \partial_\alpha R_{\beta\gamma}{}^A - \frac{1}{2} dx^\alpha \wedge dx^\beta \wedge \omega^{*B} R_{\alpha\beta}{}^C f_{BC}{}^A. \end{aligned}$$

Ее значение на горизонтальных векторных полях равно

$$dR^A(\mathfrak{h}\tilde{X}, \mathfrak{h}\tilde{Y}, \mathfrak{h}\tilde{Z}) = 3\tilde{X}^\alpha \tilde{Y}^\beta \tilde{Z}^\gamma D_{[\alpha} R_{\beta\gamma]}{}^A,$$

где квадратные скобки обозначают антисимметризацию по трем индексам, и

$$D_\alpha R_{\beta\gamma}{}^A = \partial_\alpha R_{\beta\gamma}{}^A + \omega_\alpha{}^B R_{\beta\gamma}{}^C f_{BC}{}^A.$$

Таким образом, тождества Бианки в компонентах имеют вид

$$D_\alpha R_{\beta\gamma}{}^A + D_\beta R_{\gamma\alpha}{}^A + D_\gamma R_{\alpha\beta}{}^A = 0. \quad (13.46)$$

Если задано локальное сечение расслоения, то эти тождества можно спустить на базу, используя возврат отображения. Тогда тождества Бианки для компонент локальных форм кривизны и связности примут следующий вид

$$\nabla_\alpha F_{\beta\gamma}{}^A + \nabla_\beta F_{\gamma\alpha}{}^A + \nabla_\gamma F_{\alpha\beta}{}^A = 0,$$

где

$$\nabla_\alpha F_{\beta\gamma}{}^A = \partial_\alpha F_{\beta\gamma}{}^A + A_\alpha{}^B F_{\beta\gamma}{}^C f_{BC}{}^A$$

и напряженность $F_{\alpha\beta}{}^A$ имеет вид (13.35). Именно в таком виде они, как правило, используются в приложениях. $\square$

13.2.1 Связность на $\mathbb{P}(\mathbb{R}^2, \pi, \mathbb{R}) = \mathbb{R}^3$

Чтобы лучше представить себе довольно сложное понятие связности на главном расслоении, в настоящем разделе мы рассмотрим простой и наглядный пример, когда пространство главного расслоения совпадает с обычным трехмерным евклидовым пространством.

Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовой системой координат x, y, z как главное расслоение. В качестве базы расслоения выберем плоскость $x, y \in \mathbb{R}^2$, а типичным слоем будем считать ось $z \in \mathbb{R}$, которая рассматривается, как группа трансляций. Выберем также естественную проекцию

$$\pi : \mathbb{R}^3 \ni (x, y, z) \mapsto (x, y) \in \mathbb{R}^2.$$

Таким образом, построено гладкое тривиальное главное расслоение $\mathbb{P}(\mathbb{R}^2, \pi, \mathbb{R}) = \mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$.

Вертикальные подпространства $\mathbb{V}(\mathbb{P})$ в касательном расслоении $\mathbb{T}(\mathbb{P})$ одномерны, и в каждой точке $p = (x, y, z) \in \mathbb{P}$ натянуты на векторы, параллельные оси z .

Зададим на главном расслоении связность, т.е. инвариантное распределение горизонтальных подпространств. Для этого можно задать достаточно гладкую поверхность $z(x, y)$, где $x, y \in \mathbb{R}^2$ – координаты на поверхности. Тем самым мы выбираем сечение главного расслоения $\mathbb{P}(\mathbb{R}^2, \pi, \mathbb{R})$. После этого мы сдвигаем поверхность на все возможные постоянные векторы вдоль оси z . Теперь отождествим распределение горизонтальных подпространств с касательными пространствами ко всем поверхностям. По построению, такое распределение будет инвариантно относительно трансляций. При этом мы требуем также выполнение неравенств

$$\frac{\partial z}{\partial x} \neq \infty, \quad \frac{\partial z}{\partial y} \neq \infty, \quad \forall (x, y) \in \mathbb{R}^2,$$

Это является необходимым и достаточным условием разложения касательного пространства

$$\mathbb{T}_p(\mathbb{P}) = \mathbb{V}_p(\mathbb{P}) \oplus \mathbb{H}_p(\mathbb{P})$$

в каждой точке $p \in \mathbb{P}$ в прямую сумму вертикальных $\mathbb{V}_p(\mathbb{P})$ и горизонтальных $\mathbb{H}_p(\mathbb{P})$ подпространств. Таким образом, мы построили связность Γ на главном расслоении $\mathbb{P}(\mathbb{R}^2, \pi, \mathbb{R})$.

Рассмотрим касательный вектор к главному расслоению в координатном базисе

$$\tilde{X} = \tilde{X}^x \partial_x + \tilde{X}^y \partial_y + \tilde{X}^z \partial_z \in \mathbb{T}_p(\mathbb{P}).$$

Он раскладывается единственным образом на вертикальную и горизонтальные составляющие

$$\tilde{X} = \mathbf{v}\tilde{X} + \mathbf{h}\tilde{X},$$

где

$$\mathbf{v}\tilde{X} = \tilde{X}^z \partial_z - \tilde{X}^x \frac{\partial z}{\partial x} \partial_z - \tilde{X}^y \frac{\partial z}{\partial y} \partial_z, \quad (13.47)$$

$$\mathbf{h}\tilde{X} = \tilde{X}^x \partial_x + \tilde{X}^y \partial_y + \tilde{X}^x \frac{\partial z}{\partial x} \partial_z + \tilde{X}^y \frac{\partial z}{\partial y} \partial_z. \quad (13.48)$$

Форма связности $\omega = dx\omega_x + dy\omega_y + dz$ в рассматриваемом случае имеет только две неизвестные компоненты: ω_x и ω_y , т.к. структурная группа одномерна, и $L_A \mapsto \partial_z$ и $\omega^{*A} \mapsto dz$. Сравнение разложения (13.48) с общей формулой (13.7) показывает, что компоненты связности имеют вид

$$\omega_x = -\frac{\partial z}{\partial x}, \quad \omega_y = -\frac{\partial z}{\partial y}.$$

В рассматриваемом случае компоненты формы связности не зависят от z в силу трансляционной инвариантности. Легко видеть, что значение формы связности на любом горизонтальном векторном поле равно нулю:

$$\omega(\mathbf{h}\tilde{X}) = 0. \quad (13.49)$$

Верно также и обратное утверждение: любое решение уравнения (13.49) имеет вид (13.47).

Единственная возможная нетривиальная компонента тензора кривизны равна нулю:

$$R_{xy} = \partial_x \omega_y - \partial_y \omega_x = -\frac{\partial^2 z}{\partial x \partial y} + \frac{\partial^2 z}{\partial y \partial x} = 0.$$

Рассмотрим произвольное сечение главного расслоения

$$\sigma : \mathbb{R}^2 \ni (x, y) \mapsto (x, y, w(x, y)) \in \mathbb{R}^3. \quad (13.50)$$

Дифференциал этого отображения отображает касательные векторы к базе в касательные векторы к пространству главного расслоения

$$\begin{aligned} \mathbb{T}(\mathbb{R}^2) \ni X = X^x \partial_x + X^y \partial_y &\mapsto \\ \mapsto \sigma_* X = X^z \partial_x + X^y \partial_y + X^x \frac{\partial w}{\partial x} \partial_z + X^y \frac{\partial w}{\partial y} \partial_z &\in \mathbb{T}(\mathbb{R}^2). \end{aligned}$$

Из равенства $\omega_\sigma(X) = \omega(\sigma_* X)$ следуют выражения для компонент локальной формы связности $\omega_\sigma = dx A_x + dy A_y$:

$$A_x = \frac{\partial(w - z)}{\partial x}, \quad A_y = \frac{\partial(w - z)}{\partial y}.$$

Если задано другое сечение $w'(x, y)$, то компоненты локальной формы связности преобразуются по правилу

$$A'_x = A_x + \partial_x a, \quad A'_y = A_y + \partial_y a, \quad (13.51)$$

где $a(x, y) := w'(x, y) - w(x, y)$.

Заметим, что сама форма связности вообще никак не преобразуется при изменении сечений, т.к. задана на пространстве расслоения и определена до рассмотрения каких-либо сечений.

Локальная форма кривизны для введенной выше связности также равна нулю.

Таким образом, построенная связность является плоской. Распределение горизонтальных векторных полей в данном случае находится в инволюции и, согласно теореме Фробениуса, через каждую точку проходит интегральное подмногообразие. Им является поверхность $z(x, y)$, проходящая через эту точку.

Выше мы рассмотрели связность специального вида, которая задается поверхностями $z(x, y)$. В общем случае распределение горизонтальных векторных полей задается произвольной формой связности

$$\omega = dx \omega_x + dy \omega_y + dz,$$

где ω_x и ω_y – некоторые достаточно гладкие функции от x, y . Теперь форма кривизны может иметь нетривиальную компоненту

$$R_{xy} = \partial_x \omega_y - \partial_y \omega_x.$$

Поэтому в общем случае теорема Фробениуса не выполняется, и интегральные подмногообразия отсутствуют.

Если задано сечение (13.50), то компоненты локальной формы связности $\omega_\sigma = dx A_x + dy A_y$ будут иметь вид

$$A_x = \omega_x + \partial_x w, \quad A_y = \omega_y + \partial_y w.$$

Калибровочное преобразование при переходе к другому сечению при этом останется прежним (13.51).

Локальная форма кривизны в общем случае имеет одну независимую компоненту

$$F_{xy} = \partial_x A_y - \partial_y A_x,$$

которая может быть нетривиальной.

13.3 Параллельный перенос

Пусть на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана связность $\Gamma : \mathbb{P} \ni p \mapsto \mathbb{H}_p \subset T_p(\mathbb{P})$. Определим понятие параллельного переноса слоя $\pi^{-1}(x_0)$ над точкой базы $x_0 \in \mathbb{M}$ вдоль произвольной кусочно дифференцируемой кривой

$$\gamma : [0, 1] \ni t \mapsto x(t) \in \mathbb{M} \quad (13.52)$$

с началом в точке x_0 . Для наших целей достаточно рассматривать кусочно дифференцируемые кривые класса $\mathcal{C}^1$.

Определение. *Горизонтальной кривой* в $\mathbb{P}$ называется кусочно дифференцируемая кривая, все касательные векторы которой горизонтальны. *Горизонтальным лифтом* или *подъемом* (или просто *лифтом*) кривой γ (13.52), заданной на базе $\mathbb{M}$, называется горизонтальная кривая в $\mathbb{P}$,

$$\tilde{\gamma} : [0, 1] \ni t \mapsto p(t) \in \mathbb{P},$$

такая, что $\pi \circ \tilde{\gamma} = \gamma$. □

Понятие горизонтального лифта кривой соответствует понятию лифта векторного поля. Действительно, если $\tilde{X} \in \mathcal{X}(\mathbb{P})$ – лифт дифференцируемого векторного поля X , заданного на базе $\mathbb{M}$, то интегральная кривая $\tilde{\gamma}$ векторного поля $\tilde{X}$, проходящая через точку $p_0 \in \mathbb{P}$, есть горизонтальный лифт интегральной кривой γ поля X , проходящей через точку $x_0 = \pi(p_0)$.

Предложение 13.3.1. *Пусть $\gamma = x(t)$, $t \in [0, 1]$, – кусочно дифференцируемая кривая класса $\mathcal{C}^1$ в $\mathbb{M}$ с началом в точке $x_0 \in \mathbb{M}$. Тогда для произвольной точки слоя $p_0 \in \pi^{-1}(x_0)$ существует единственный горизонтальный лифт $\tilde{\gamma} = p(t)$ кривой γ с началом в точке p_0 .*

Доказательство. Состоит в явном построении лифта $\tilde{\gamma}$. См., например, [48], глава II, предложение 3.1. □

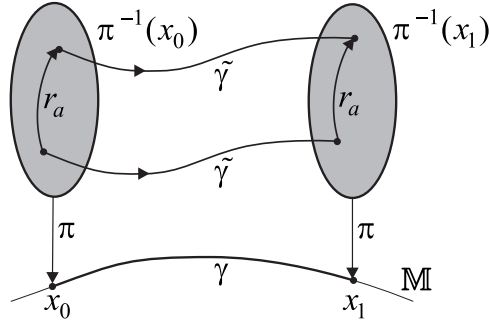
Используя предложение 13.3.1, определим параллельное перенесение слоев следующим образом.

Определение. Пусть $\gamma = x(t)$ – кривая в $\mathbb{M}$ с началом и концом в точках x_0 и x_1 . Пусть $\tilde{\gamma}$ – единственный горизонтальный лифт кривой γ , который начинается в точке p_0 , находящейся в слое $\pi^{-1}(x_0)$. Лифт $\tilde{\gamma}$ имеет конечную точку $p_1 \in \mathbb{P}$ такую, что $\pi(p_1) = x_1$. Меняя начальную точку p_0 в слое $\pi^{-1}(x_0)$, мы получаем отображение слоя $\pi^{-1}(x_0)$ в слой $\pi^{-1}(x_1)$, которое переводит точку p_0 в p_1 . Это отображение называется *параллельным переносом* слоя из точки x_0 в точку x_1 вдоль кривой γ . Параллельный перенос слоев обозначается той же буквой, что и кривая, $\gamma : \pi^{-1}(x_0) \rightarrow \pi^{-1}(x_1)$. □

Параллельный перенос слоев является изоморфизмом, что вытекает из следующего утверждения.

Предложение 13.3.2. *Параллельный перенос слоя $\gamma : \pi^{-1}(x_0) \rightarrow \pi^{-1}(x_1)$ вдоль любой кривой перестановочен с действием структурной группы $\mathbb{G}$ на $\mathbb{P}$: $\gamma \circ r_a = r_a \circ \gamma$ для всех $a \in \mathbb{G}$.*

Доказательство. Следует из того, что правое действие структурной группы r_a отображает каждую горизонтальную кривую в горизонтальную, см. рис.13.1. □

Рис. 13.1: Параллельный перенос слоев вдоль кривой γ .

Параллельный перенос слоя вдоль кривой γ не зависит от выбора параметризации кривой. Кроме того, если слой переносится из точки x_0 в точку x_1 параллельно вдоль кривой γ , то он параллельно переносится вдоль этой же кривой из точки x_0 в любую промежуточную точку $x(t)$, $t \in [0, 1]$, на γ .

Замечание. Если точки x_0 и x_1 в базе $\mathbb{M}$ фиксированы, то в общем случае параллельный перенос слоя $\pi^{-1}(x_0)$ в слой $\pi^{-1}(x_1)$ зависит от кривой γ , соединяющей эти точки. Для односвязных баз эта зависимость характеризуется формой кривизны R формы связности ω и будет обсуждаться в следующих разделах. $\square$

При рассмотрении фундаментальной группы в разделе 10.3 мы определили произведение путей (кривых) $\gamma_2 \circ \gamma_1$ как последовательный проход вдоль путей γ_1 и γ_2 (10.4) и обратный путь γ^{-1} как путь γ , проходимый в обратном направлении (10.5). Следующее предложение очевидно.

Предложение 13.3.3. 1) Если γ_1 – путь из x_0 в x_1 и γ_2 – путь из x_1 в x_2 , то параллельный перенос слоя $\pi^{-1}(x_0)$ в слой $\pi^{-1}(x_2)$ вдоль произведения путей $\gamma_2 \circ \gamma_1$ равен произведению отображений слоев $\gamma_2 \circ \gamma_1 : \pi^{-1}(x_0) \rightarrow \pi^{-1}(x_2)$.

2) Если γ^{-1} – обратный путь для пути γ из точки x_0 в точку x_1 , то параллельный перенос слоя $\pi^{-1}(x_1)$ в слой $\pi^{-1}(x_0)$ вдоль пути γ^{-1} является обратным отображением $\gamma^{-1} = \gamma^{-1} : \pi^{-1}(x_1) \rightarrow \pi^{-1}(x_0)$.

13.4 Группы голономии

Пусть задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ со связностью Γ . Используя понятие параллельного переноса, определим группу голономии данной связности Γ .

Обозначим через $\Omega(\mathbb{M}, x)$ множество замкнутых кусочно дифференцируемых кривых (петель) на базе $\mathbb{M}$ с началом и концом в точке $x \in \mathbb{M}$. Подмножество, состоящее из путей, гомотопных постоянному пути в точке x , обозначим $\Omega_0(\mathbb{M}, x) \subset \Omega(\mathbb{M}, x)$. Произведение и обратный путь для всех путей из $\Omega(\mathbb{M}, x)$ были определены в разделе 10.3. В разделе 13.3 было показано, что параллельный перенос слоя $\pi^{-1}(x)$ вдоль замкнутого пути $\gamma \in \Omega(\mathbb{M}, x)$ есть изоморфизм слоя $\pi^{-1}(x)$ на себя. В общем случае этот изоморфизм будет нетривиален, т.к. после параллельного переноса вдоль замкнутого пути слой может повернуться. Множество всех таких изоморфизмов образует группу в силу предложения 13.3.3.

Определение. Группа, состоящая из изоморфизмов слоя $\pi^{-1}(x)$, которые соответствуют параллельным переносам данного слоя вдоль всех замкнутых кусочно дифференцируемых путей $\gamma \in \Omega(\mathbb{M}, x)$, называется *группой голономии* $\Phi(x)$ связности Γ

в точке $x \in \mathbb{M}$. Подгруппа $\Phi_0(x) \subset \Phi(x)$, соответствующая параллельным переносам вдоль замкнутых путей, стягиваемых в точку, $\gamma \in \Omega_0(\mathbb{M}, x)$, называется *суженной группой голономии* связности Γ в точке $x \in \mathbb{M}$. $\square$

Группу голономии $\Phi(x)$ и суженную группу голономии $\Phi_0(x)$ можно считать подгруппами в структурной группе $\mathbb{G}$ следующим образом.

Определение. Зафиксируем некоторую точку слоя $p \in \pi^{-1}(x)$. После параллельного переноса слоя вдоль пути $\gamma \in \Omega(\mathbb{M}, x)$ эта точка отобразится в некоторую точку $\gamma(p) = pa \in \pi^{-1}(x)$, где $a \in \mathbb{G}$ – некоторый элемент структурной группы. Если задан другой путь $\gamma' \in \Omega(\mathbb{M}, x)$, которому соответствует элемент $b \in \mathbb{G}$, то произведение путей $\gamma' \circ \gamma$ определяет элемент $ba \in \mathbb{G}$, поскольку

$$\gamma' \circ \gamma(p) = \gamma'(pa) = (\gamma'(p))a = pba.$$

По предложению 13.3.3 множество элементов $a \in \mathbb{G}$, определенных всеми путями $\gamma \in \Omega(x)$, образует группу, которая называется *группой голономии* $\Phi(p)$ связности Γ в точке $p \in \mathbb{P}$. Замкнутым путям $\gamma \in \Omega_0(\mathbb{M}, x)$ соответствует *суженная группа голономии* $\Phi_0(p)$ в точке $p \in \mathbb{P}$. $\square$

Замечание. $\Phi(x)$ есть группа изоморфизмов слоя $\pi^{-1}(x)$ на себя, а $\Phi(p)$ есть подгруппа в $\mathbb{G}$. Выше мы построили единственный изоморфизм из $\Phi(x)$ на $\Phi(p)$, который делает коммутативной следующую диаграмму:

$$\begin{array}{ccc} \Omega(\mathbb{M}, x) & & \\ \downarrow & \searrow & \\ \Phi(x) & \longrightarrow & \Phi(p) \end{array}$$

$\square$

Группу голономии $\Phi(p)$ можно определить другим образом. Введем на пространстве главного расслоения отношение эквивалентности $p \sim q$, где $p, q \in \mathbb{P}$, если точки p и q можно соединить горизонтальной кривой. При этом точки p и q не обязательно лежат в одном слое. Нетрудно проверить, что это действительно отношение эквивалентности. Тогда группа голономии $\Phi(p)$ совпадает с множеством тех элементов $a \in \mathbb{G}$, для которых $p \sim pa$. Легко проверить, что это множество элементов образует подгруппу в $\mathbb{G}$, т.к. $p \sim q$ влечет за собой $pa \sim qa$.

Предложение 13.4.1. Пусть дано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ со связностью Γ . Тогда:

- 1) Если $q = pa$, $a \in \mathbb{G}$, то $\Phi(q) = \text{ad}(a^{-1})\Phi(p)$, т.е. группы голономии точек одного слоя $\Phi(q)$ и $\Phi(p)$ сопряжены в $\mathbb{G}$; аналогично, $\Phi_0(q) = \text{ad}(a^{-1})\Phi_0(p)$.
- 2) Если точки $p, q \in \mathbb{P}$ можно соединить горизонтальной кривой, т.е. $p \sim q$, то $\Phi(p) = \Phi(q)$ и $\Phi_0(p) = \Phi_0(q)$.

Доказательство. 1). Пусть $b \in \Phi(p)$, т.е. $p \sim pb$. Тогда $pa \sim pba$ и, следовательно, $q \sim qa^{-1}ba$. Поэтому $\text{ad}(a^{-1})b = a^{-1}ba \in \Phi(q)$. Отсюда вытекает, что $\Phi(q) = \text{ad}(a^{-1})\Phi(p)$ и $\Phi_0(q) = \text{ad}(a^{-1})\Phi_0(p)$.

2) Отношение $p \sim q$ влечет за собой $pb \sim qb$. Из транзитивности отношения эквивалентности $\sim$ следует, что $p \sim pb$ тогда и только тогда, когда $q \sim qb$, т.е.

$b \in \Phi(p)$ тогда и только тогда, когда $b \in \Phi(q)$. Тем самым $\Phi(p) = \Phi(q)$. Чтобы доказать равенство $\Phi_0(p) = \Phi_0(q)$, допустим, что $\tilde{\delta}$ – горизонтальная кривая в $\mathbb{P}$ из p в q . Если $b \in \Phi_0(p)$, то существует горизонтальная кривая $\tilde{\gamma}$ в $\mathbb{P}$ из p в pb такая, что кривая в базе $\pi\tilde{\gamma}$ является замкнутым путем с началом и концом в точке $\pi(p)$, которая гомотопна постоянному пути в точке $\pi(p)$. Тогда композиция $(r_b\tilde{\delta}) \circ \tilde{\gamma} \circ \delta^{-1}$ есть горизонтальная кривая в $\mathbb{P}$ из q в qb и ее проекция на базу $\mathbb{M}$ есть замкнутый путь с началом и концом в точке $\pi(q)$, который гомотопен постоянному пути. Поэтому $b \in \Phi_0(q)$. $\square$

Если база $\mathbb{M}$ связна, то для любой пары точек $p, q \in \mathbb{P}$ найдется элемент $a \in \mathbb{G}$ такой, что $q \sim pa$. Поэтому из предложения 13.4.1 следует, что группы голономии $\Phi(p)$ для всех точек $p \in \mathbb{P}$ сопряжены друг другу в $\mathbb{G}$ и поэтому изоморфны. По тем же причинам все суженные группы голономии $\Phi_0(p)$ также изоморфны друг другу.

Итак, мы определили группу голономии $\Phi(p)$, суженную группу голономии $\Phi_0(p)$ и показали, что с точностью до преобразования подобия они не зависят от точки расслоения $\mathbb{P}$. Теперь сформулируем несколько общих свойств групп голономий.

Теорема 13.4.1. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение со связной базой $\mathbb{M}$. Пусть $\Phi(p)$ и $\Phi_0(p)$ – группа голономии и суженная группа голономии связности Γ в точке $p \in \mathbb{P}$. Тогда:

- 1) $\Phi_0(p)$ есть связная подгруппа Ли в $\Phi(p)$;
- 2) $\Phi_0(p)$ есть нормальная подгруппа в $\Phi(p)$ и фактор группа $\Phi(p)/\Phi_0(p)$ счетна.

Доказательство. Используется паракомпактность базы $\mathbb{M}$. См., например, [48], глава II, теорема 4.2. $\square$

Следствие. Группа голономии $\Phi(p)$ является подгруппой Ли в структурной группе $\mathbb{G}$ с компонентой единицы $\Phi_0(p)$. В частности, $\dim \Phi(p) = \dim \Phi_0(p)$. $\square$

При определении групп голономий мы не оговорили класс дифференцируемости рассматриваемых кусочно дифференцируемых кривых $\gamma \in \Omega(\mathbb{M}, x)$. Чем ниже класс дифференцируемости, тем больше множество кривых. Поэтому могло бы оказаться так, что группы голономии зависят от класса дифференцируемости кривых. Это оказывается не верно. Пусть $\Omega^k(\mathbb{M}, x)$ – множество замкнутых кусочно дифференцируемых путей в $\mathbb{M}$ с началом и концом в точке $x \in \mathbb{M}$ класса $\mathcal{C}^k$. Обозначим соответствующую группу голономии через $\Phi^k(p)$. Очевидно, что $\Phi^1(p) \supset \Phi^2(p) \supset \dots \supset \Phi^\infty(p)$. Верны также и обратные включения.

Теорема 13.4.2 (Номидзу, Одзеки). Все группы голономии $\Phi^k(p)$, $1 \leq k \leq \infty$, совпадают.

Доказательство. См. [77]. $\square$

Следствие. Все суженные группы голономии $\Phi_0^k(p)$, $1 \leq k \leq \infty$, совпадают. $\square$

Доказательство. Согласно теореме 13.4.1 суженная группа голономии $\Phi_0^k(p)$ есть связная компонента единицы группы $\Phi^k(p)$. $\square$

Замечание. В случае, когда $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ является вещественно аналитическим главным расслоением с аналитической связностью Γ , можно определить группу голономии $\Phi^\omega(p)$, используя только кусочно аналитические кривые. Можно доказать, что $\Phi^\omega(p) = \Phi^1(p)$ и $\Phi_0^\omega(p) = \Phi_0^1(p)$ [48]. $\square$

Поскольку группы голономии не зависят от класса дифференцируемости путей, то в дальнейшем мы не будем его указывать.

13.5 Петля Вильсона

Настоящий раздел посвящен одному из способов вычисления группы голономии. Пусть задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ со связностью Γ . Рассмотрим окрестность $\mathbb{U} \subset \mathbb{M}$ с координатами x^α , $\alpha = 1, \dots, n$, содержащую точку базы $x_0 \in \mathbb{M}$, и отождествим подрасслоение $\pi^{-1}(\mathbb{U})$ с прямым произведением $\mathbb{U} \times \mathbb{G}$. Тогда точка подрасслоения задается парой элементов $p = (x, a) \in \mathbb{U} \times \mathbb{G}$. Пусть $\gamma = x(t)$ – произвольная кривая в $\mathbb{U}$ с началом в точке x_0 и $\tilde{\gamma} = (x(t), a(t))$ – ее единственный горизонтальный лифт с началом в точке $p_0 = (x_0, a_0)$. При этом функция $a(t)$ определяет некоторую кривую в структурной группе $\mathbb{G}$ с началом в точке a_0 . В инвариантном виде мы пишем

$$\tilde{\gamma} = \overset{\circ}{\gamma} \circ a(t), \quad (13.53)$$

где $\overset{\circ}{\gamma} = (x(t), e)$ – опорная кривая в $\pi^{-1}(\mathbb{U})$, лежащая в нулевом сечении, для которой $\pi(\overset{\circ}{\gamma}) = \gamma$. При этом мы рассматриваем равенство (13.53) как уравнение на $a(t)$ при заданной кривой γ , которая однозначно определяет опорную кривую $\overset{\circ}{\gamma}$.

Касательный вектор к горизонтальной кривой $\tilde{\gamma}$ имеет вид $\dot{\tilde{\gamma}} = (\dot{x}, \dot{a})$. Уравнение для $a(t)$ получается из условия горизонтальности. Рассмотрим окрестность единицы группы, где форма связности имеет вид (13.5). Горизонтальность касательного вектора к кривой записывается в виде равенства

$$\omega(\dot{\tilde{\gamma}}) = (\dot{x}^\alpha \omega_\alpha^A + \dot{a}^B L_B^{-1A}) L_A = 0.$$

Отсюда следует система уравнений на $a^A(t)$:

$$\dot{x}^\alpha \overset{\circ}{A}_\alpha^B S_B^{-1A} + \dot{a}^B L_B^{-1A} = 0, \quad (13.54)$$

где $\overset{\circ}{A}_\alpha^A(x)$ – компоненты локальной формы связности на нулевом сечении (поле Янга–Миллса) и $S_B^A(a)$ – матрица присоединенного представления, соответствующая элементу $a \in \mathbb{G}$.

Уравнение (13.54) неудобно для приложений, т.к. содержит матрицу $L_B^{-1A}(a)$, которая определена только в окрестности единицы группы. Чтобы устранить это неудобство, перейдем к какому-либо представлению структурной группы

$$\rho: \quad \mathbb{G} \ni \quad a \mapsto (S_j^{-1i}(a)) \quad \in \text{aut } \mathbb{V}, \quad i, j = 1, \dots, \dim \mathbb{V}.$$

Мы выбрали представление в виде обратных матриц, чтобы не менять общепринятого определения упорядоченного (хронологического) произведения, которое будет дано ниже. Теперь заметим, что

$$\dot{S}_i^{-1j} = \dot{a}^A \partial_A S_i^{-1j} = \dot{a}^B L_B^{-1A} L_A S_i^{-1j} = \dot{a}^B L_B^{-1A} S_i^{-1k} L_{Ak}^j,$$

где L_{Ai}^j – представление генераторов L_A алгебры Ли $\mathfrak{g}$, и мы воспользовались правилом дифференцирования матриц представления (8.66). Умножив уравнение (13.54) на $S_i^{-1k} L_{Ak}^j$ и воспользовавшись инвариантностью генераторов L_{Ai}^j (8.68), получаем уравнение на матрицу представления

$$\dot{S}_i^{-1j} - \dot{x}^\alpha \overset{\circ}{A}_{\alpha i}^k S_k^{-1j} = 0,$$

где $\overset{\circ}{A}_{\alpha i}^j := -\overset{\circ}{A}_\alpha^A L_{Ai}^j$. Это уравнение можно переписать в почти ковариантном виде

$$\dot{x}^\alpha \left(\partial_\alpha S_i^{-1j} - \overset{\circ}{A}_{\alpha i}^k S_k^{-1j} \right) = 0.$$

Для ковариантности не хватает одного слагаемого с калибровочным полем для индекса j . Опустив, для краткости, матричные индексы, получаем уравнение

$$\dot{S}^{-1} = \dot{x}^\alpha \overset{\circ}{A}_\alpha S^{-1}, \quad (13.55)$$

которое можно записать в интегральном виде:

$$S^{-1}(t) = S_0^{-1} + \int_0^t ds \dot{x}^\alpha \overset{\circ}{A}_\alpha S^{-1} = S_0^{-1} + \int_{x(0)}^{x(t)} dx^\alpha \overset{\circ}{A}_\alpha S^{-1}.$$

где $S_0^{-1} := S^{-1}(a_0)$, точка обозначает дифференцирование по s и второй интеграл берется вдоль кривой γ от точки $x(0)$ до точки $x(t)$. Решение этого уравнения записывается в виде упорядоченной Р-экспоненты

$$S^{-1}(t) = \text{P exp} \left(\int_0^t ds \dot{x}^\alpha \overset{\circ}{A}_\alpha \right) S_0^{-1}, \quad (13.56)$$

которая определена разложением в ряд Тейлора:

$$\text{P exp} \left(\int_0^t ds \dot{x}^\alpha \overset{\circ}{A}_\alpha \right) := 1 + \int_0^t ds \dot{x}^\alpha \overset{\circ}{A}_\alpha + \int_0^t ds_1 \int_0^{s_1} ds_2 \dot{x}^{\alpha_1} \overset{\circ}{A}_{\alpha_1} \dot{x}^{\alpha_2} \overset{\circ}{A}_{\alpha_2} + \dots, \quad (13.57)$$

где

$$\dot{x}^{\alpha_k} \overset{\circ}{A}_{\alpha_k} := \frac{dx^{\alpha_k}(s_k)}{ds_k} \overset{\circ}{A}_{\alpha_k}(s_k), \quad \forall k = 1, 2, \dots$$

Напомним общее определение Р-произведения.

Определение. Пусть задано семейство операторов $A(t)$, непрерывно зависящих от вещественного параметра t , тогда

$$\text{P}[A(t_1)A(t_2)] := \begin{cases} A(t_1)A(t_2), & t_1 \geq t_2, \\ A(t_2)A(t_1), & t_1 < t_2 \end{cases} \quad (13.58)$$

называется Р-произведением или упорядоченным произведением операторов $A(t_1)$ и $A(t_2)$. $\square$

Если операторы для различных точек коммутируют, то Р-произведение совпадает с обычным произведением.

Замечание. В квантовой теории поля роль параметра t часто играет время. Поэтому Р-произведение называют также *хронологическим произведением*. $\square$

Нетрудно проверить равенство

$$\begin{aligned} \int_0^t ds_1 \int_0^{s_1} ds_2 \dots \int_0^{s_{k-1}} ds_k A(s_1)A(s_2) \dots A(s_k) = \\ = \frac{1}{k!} \int_0^t ds_1 \int_0^t ds_2 \dots \int_0^t ds_k \text{P}[A(s_1)A(s_2) \dots A(s_k)]. \end{aligned}$$

Поэтому разложение (13.57) имеет место.

Известно, что ряд (13.57) равномерно сходится в шаре произвольного радиуса.

Продолжим общее построение. Если представление структурной группы является точным, то матрица представления $S^{-1}(a)$ однозначно определяет элемент структурной группы $a = \rho^{-1}(S^{-1})$. В этом случае решение (13.56) определяет кривую $a(t)$ с началом в точке $a_0 \in \mathbb{G}$. Меняя точку a_0 , мы получаем отображение

$$\pi^{-1}(x_0) \ni (x_0, a_0) \mapsto (x(t), a(t)) \in \pi^{-1}(x(t)), \quad \forall a_0 \in \mathbb{G}. \quad (13.59)$$

Поскольку кривая $\tilde{\gamma}$ горизонтальна, то это отображение задает параллельный перенос слоя $\pi^{-1}(x_0)$ главного расслоения $\mathbb{P}$ вдоль пути $\gamma \in \mathbb{M}$.

Таким образом, решение (13.56) определяет параллельный перенос слоев (13.59) через компоненты связности $\overset{\circ}{A}_\alpha$ на опорном нулевом сечении. Отображение (13.59) можно построить для произвольного опорного сечения следующим образом. Пусть задано произвольное сечение $\sigma(x) = (x, b(x))$. Тогда компоненты локальной формы связности для рассматриваемого представления на этом сечении имеют вид (13.15)

$$A_\alpha = S_b \overset{\circ}{A}_\alpha S_b^{-1} + \partial_\alpha S_b S_b^{-1}, \quad S_b := S(b).$$

Для горизонтальной кривой $\tilde{\gamma}$ имеем равенство

$$\tilde{\gamma} = (x(t), a(t)) = (x(t), b(t)c(t)), \quad (13.60)$$

где $c(t)$ – некоторая новая кривая в $\mathbb{G}$, связывающая кривую $(x(t), b(t))$ на сечении σ с горизонтальной кривой $\tilde{\gamma}$. Поскольку $S_a^{-1} = S_b^{-1} S_c^{-1}$, то простые вычисления приводят (13.55) к уравнению

$$\dot{S}_c^{-1} = \dot{x}^\alpha A_\alpha S_c^{-1},$$

определяющему кривую $c(t)$ с начальным условием $c_0 = b_0^{-1} a_0$. Решение этого уравнения также дается упорядоченной экспонентой

$$S_c^{-1}(t) = \text{P exp} \left(\int_0^t ds \dot{x}^\alpha A_\alpha \right) S_{c(0)}^{-1}.$$

Эта формула также определяет параллельный перенос (13.59), что следует из (13.60).

Поскольку $S_0^{-1} = S_{b(0)}^{-1} S_{c(0)}^{-1}$, то простые вычисления приводят к следующему правилу преобразования упорядоченной экспоненты при изменении сечения $\sigma_0 \mapsto \sigma$

$$\text{P exp} \left(\int_0^t ds \dot{x}^\alpha \overset{\circ}{A}_\alpha \right) = S(b_0) \text{P exp} \left(\int_0^t ds \dot{x}^\alpha A_\alpha \right) S^{-1}(b_t). \quad (13.61)$$

Этот закон преобразования похож на тензорный, однако таковым не является, потому что слева и справа стоят матрицы преобразования, взятые в различных точках: в начале и конце пути.

Если задано два произвольных сечения (13.10) то имеет место аналогичная формула.

Рассмотрим замкнутый путь в базе $\gamma \in \Omega(\mathbb{U}, x_0)$. Поскольку мы ограничились координатной окрестностью $\mathbb{U}$, то все пути стягиваемы к точке x_0 (гомотопны постоянному пути в x_0). Для этих путей $S(1) = S(0)$ и P -экспонента при калибровочном преобразовании $\sigma_0 \mapsto \sigma$ преобразуется по тензорному закону

$$\text{P exp} \left(\oint_\gamma dx^\alpha \overset{\circ}{A}_\alpha \right) = S(b_0) \text{P exp} \left(\oint_\gamma dx^\alpha A_\alpha \right) S^{-1}(b_0). \quad (13.62)$$

При параллельном переносе слоя $\pi^{-1}(x_0)$ вдоль замкнутого пути γ точка $p_0 = (x_0, a_0)$ отобразится в точку $p_1 = (x_0, a_1)$. Пусть $a_0 = e$, тогда $S_0 = \mathbb{1}$. В этом случае

$$\hat{a} = \rho^{-1} \left[\text{P exp} \left(\oint_{\gamma} dx^{\alpha} \overset{\circ}{A}_{\alpha} \right) \right] \in \Phi_0(\sigma_0(x_0)) \quad (13.63)$$

– элемент суженной группы голономии $\Phi_0(\sigma_0(x_0))$. Здесь мы предполагаем, что представление ρ является точным. Этот элемент определяется калибровочным полем, соответствующим нулевому сечению σ_0 , которое проходит через начальную точку кривой (x_0, e) . Если $a_0 \neq e$, то элемент группы голономии имеет вид $a_0^{-1}a_1 \in \Phi_0(p_0)$ и равен

$$\rho^{-1} \left[S_0^{-1} \text{P exp} \left(\oint_{\gamma} dx^{\alpha} \overset{\circ}{A}_{\alpha} \right) S_0 \right] = a_0^{-1} \hat{a} a_0 \in \Phi_0(\sigma(x_0)).$$

То есть он сопряжен элементу $\hat{a}$ в соответствии с утверждением 1) предложения 13.4.1.

Поскольку P -экспонента для замкнутого пути при изменении сечения преобразуется по тензорному закону (13.62), то для произвольного сечения σ P -экспонента определяет элемент суженной группы голономии $\Phi_0(\sigma(x_0))$,

$$\rho^{-1} \left[\text{P exp} \left(\oint_{\gamma} dx^{\alpha} A_{\alpha} \right) \right] \in \Phi_0(\sigma(x_0)), \quad (13.64)$$

где A_{α} – калибровочное поле, соответствующее произвольному сечению, проходящему через точку $\sigma(x_0) \in \pi^{-1}(\mathbb{U})$. Таким образом, рассматривая все замкнутые пути $\gamma \in \mathbb{M}$ с началом в точке x_0 , можно определить суженную группу голономии.

Определение. След от P -экспоненты

$$W_{\gamma}[A] := \text{tr} \left[\text{P exp} \left(\oint_{\gamma} dx^{\alpha} A_{\alpha} \right) \right]. \quad (13.65)$$

называется *петлей Вильсона*. □

Предложение 13.5.1. *Петля Вильсона инвариантна относительно калибровочных преобразований.*

Доказательство. Прямое следствие равенства (13.62). □

Замечание. Петля Вильсона играет важную роль в решеточной формулировке квантовых калибровочных моделей. □

Полученное выражение для упорядоченной экспоненты (13.56) не зависит от выбора координат на $\mathbb{U} \subset \mathbb{M}$. Это значит, что понятие упорядоченной экспоненты, которое было получено в одной карте, без труда переносится на произвольные пути в $\mathbb{M}$, которые в общем случае не покрываются одной картой. Для этого весь путь надо разбить на отрезки, каждый из которых покрывается одной картой, и взять сумму интегралов вдоль каждого отрезка.

13.6 Отображение связностей

В разделе 12.3 мы изучили отображение расслоений. В частности, было определено вложение (инъекция) расслоений, редукция структурной группы, а также индуцированное расслоение. Ниже мы изучим вопрос о том, как ведут себя связности и соответствующие им группы голономий при отображении расслоений. В дальнейшем эти результаты будут использованы при изучении групп голономий.

Предложение 13.6.1. Пусть $f : \mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1) \rightarrow \mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$ – гомоморфизм главных расслоений, состоящий из дифференцируемого отображения пространств расслоений $f_{\mathbb{P}} : \mathbb{P}_1 \rightarrow \mathbb{P}_2$ и гомоморфизма структурных групп $f_{\mathbb{G}} : \mathbb{G}_1 \rightarrow \mathbb{G}_2$, такой, что индуцированное отображение баз $f_{\mathbb{M}} : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ есть диффеоморфизм. Пусть Γ_1 – связность на $\mathbb{P}_1$ с формой связности ω_1 и формой кривизны R_1 . Тогда:

- 1) Существует единственная связность Γ_2 на $\mathbb{P}_2$ такая, что $f_{\mathbb{P}}$ отображает горизонтальные подпространства связности Γ_1 в горизонтальные подпространства связности Γ_2 .
- 2) Если ω_2 и R_2 – формы связности и кривизны для Γ_2 , то

$$f_{\mathbb{P}}^* \omega_2 = f_{\mathfrak{g}} \omega_1 \quad \text{и} \quad f_{\mathbb{P}}^* R_2 = f_{\mathfrak{g}} R_1,$$

где правые части $f_{\mathfrak{g}} \omega_1$ и $f_{\mathfrak{g}} R_1$ обозначают $\mathfrak{g}_2$ -значные формы на $\mathbb{P}_1$, определенные соотношениями:

$$(f_{\mathfrak{g}} \omega_1)(\tilde{X}) = f_{\mathfrak{g}}(\omega_1(\tilde{X})) \quad \text{и} \quad (f_{\mathfrak{g}} R_1)(\tilde{X}, \tilde{Y}) = f_{\mathfrak{g}}(R_1(\tilde{X}, \tilde{Y})) \quad \forall \tilde{X}, \tilde{Y} \in \mathcal{X}(\mathbb{P}_1),$$

где $f_{\mathfrak{g}} := f_{\mathbb{G}*}$ – гомоморфизм алгебр Ли $\mathfrak{g}_1 \rightarrow \mathfrak{g}_2$, индуцированный отображением структурных групп $f_{\mathbb{G}} : \mathbb{G}_1 \rightarrow \mathbb{G}_2$ (дифференциал отображения $f_{\mathbb{G}}$).

- 3) Если $p_2 = f_{\mathbb{P}}(p_1) \in \mathbb{P}_2$ – образ точки $p_1 \in \mathbb{P}_1$, то $f_{\mathbb{G}}$ гомоморфно отображает группу голономии $\Phi(p_1)$ в точке p_1 на $\Phi(p_2)$ и ограниченную группу голономии $\Phi_0(p_1)$ на $\Phi_0(p_2)$.

Доказательство. Проводится путем явного построения связности Γ_2 на $\mathbb{P}_2$. См., например, [48], глава II, предложение 6.1. $\square$

Определение. В ситуации, описанной в предложении 13.6.1, говорят, что f отображает связность Γ_1 в связность Γ_2 . В частном случае, если $\mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1)$ – редуцированное подрасслоение в $\mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$, т.е. $f_{\mathbb{G}}$ – мономорфизм, $\mathbb{M}_1 = \mathbb{M}_2 = \mathbb{M}$ и $f_{\mathbb{M}} = \text{id}_{\mathbb{M}}$, то говорят, что связность Γ_2 на $\mathbb{P}_2$ *редуцируема* к связности Γ_1 на $\mathbb{P}_1$. Автоморфизм f главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ называется *автоморфизмом связности* Γ на $\mathbb{P}$, если он отображает Γ в Γ . В этом случае говорят, что связность Γ *инвариантна* относительно f . $\square$

Предложение 13.6.2. Любая связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ инвариантна относительно вертикальных автоморфизмов (пример 12.3.1). $\square$

Доказательство. Прямое следствие свойства 2) в определении связности. $\square$

Следствие. Пусть $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$ – подрасслоение в $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, где $\mathbb{H}$ – подгруппа Ли в $\mathbb{G}$. Пусть Γ – связность на $\mathbb{P}$ с формой связности ω . Тогда связность Γ на $\mathbb{P}$ редуцируема к связности Γ' на $\mathbb{Q}$ тогда и только тогда, когда сужение формы связности ω на $\mathbb{Q}$ является $\mathfrak{h}$ -значным, где $\mathfrak{h}$ – подалгебра Ли в $\mathfrak{g}$, соответствующая подгруппе Ли $\mathbb{H} \subset \mathbb{G}$. Если связность Γ редуцируема к Γ' , то группы голономии Φ и суженные группы голономии Φ_0 для $\mathbb{Q}$ и $\mathbb{P}$ изоморфны. $\square$

Доказательство. Пусть Γ' – связность на $\mathbb{Q}$. Ее форма связности, по определению, $\mathfrak{h}$ -значна и по предложению 13.6.1 продолжается единственным образом до связности Γ на $\mathbb{P}$. Обратно. Если связность Γ на $\mathbb{P}$ редуцируема к связности на $\mathbb{Q}$, то сужение ω на $\mathbb{Q}$ $\mathfrak{h}$ -значно. Пусть связность Γ редуцируема к Γ' . Отождествим множество точек $\mathbb{Q}$ с его образом $f_{\mathbb{P}}(\mathbb{Q})$ в $\mathbb{P}$. Тогда любая горизонтальная кривая в $\mathbb{P}$ с началом в произвольной точке $p' \in f_{\mathbb{P}}(\mathbb{Q})$ будет целиком лежать в $f_{\mathbb{P}}(\mathbb{Q})$, т.к. сужение распределения горизонтальных подпространств в $\mathbb{P}$ на $\mathbb{Q}$ совпадает со связностью на $\mathbb{Q}$. Поскольку группы голономии для всех точек $p \in \mathbb{P}$ изоморфны, то изоморфны также все группы голономии для $\mathbb{Q}$ и $\mathbb{P}$. $\square$

Предложение 13.6.3. Пусть $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$ – подрасслоение в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, где $\mathbb{H}$ – подгруппа Ли в $\mathbb{G}$. Допустим, что алгебра Ли $\mathfrak{g}$ для $\mathbb{G}$ допускает подпространство $\mathfrak{m}$ такое, что $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ и $\text{ad}(\mathbb{H})\mathfrak{m} = \mathfrak{m}$, где $\mathfrak{h}$ подалгебра Ли для $\mathbb{H}$. Тогда для каждой формы связности ω на $\mathbb{P}$ $\mathfrak{h}$ -компонента ω' формы связности ω , суженная на $\mathbb{Q}$ является формой связности на $\mathbb{Q}$.

Доказательство. См., например, [48], глава II, предложение 6.4. $\square$

Замечание. В силу следствия из предложения 13.6.1 форма связности ω' совпадает с ω на $\mathbb{Q}$. $\square$

В предложении 13.6.1 мы рассматривали отображение связности Γ_1 на $\mathbb{P}_1$ в некоторую связность Γ_2 на $\mathbb{P}_2$. При определенных условиях справедливо также обратное утверждение, и связность с $\mathbb{P}_2$ можно перенести на $\mathbb{P}_1$.

Предложение 13.6.4. Пусть $f : \mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1) \rightarrow \mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$ – гомоморфизм главных расслоений такой, что гомоморфизм структурных групп $f_{\mathbb{G}} : \mathbb{G}_1 \rightarrow \mathbb{G}_2$ является изоморфизмом. Пусть Γ_2 – связность на $\mathbb{P}_2$ с формой связности ω_2 и формой кривизны R_2 . Тогда:

- 1) Существует единственная связность Γ_1 на $\mathbb{P}_1$ такая, что $f_{\mathbb{P}}$ отображает горизонтальные подпространства связности Γ_1 в горизонтальные подпространства связности Γ_2 .
- 2) Если ω_1 и R_1 – формы связности и кривизны для Γ_1 , то

$$f_{\mathbb{P}}^* \omega_2 = f_{\mathfrak{g}} \omega_1 \quad \text{и} \quad f_{\mathbb{P}}^* R_2 = f_{\mathfrak{g}} R_1,$$

где правые части определены в предложении 13.6.1.

- 3) Если $p_2 = f_{\mathbb{P}}(p_1) \in \mathbb{P}_2$ – образ точки $p_1 \in \mathbb{P}_1$, то изоморфизм $f_{\mathbb{G}}$ гомоморфно отображает группу голономии $\Phi(p_1)$ в точке p_1 в $\Phi(p_2)$ и ограниченную группу голономии $\Phi_0(p_1)$ в $\Phi_0(p_2)$.

Доказательство. Проводится путем явного построения связности Γ_1 на $\mathbb{P}_1$. См., например, [48], глава II, предложение 6.2. $\square$

Определение. В ситуации, описанной в предложении 13.6.4, говорят, что связность Γ_1 индуцирована гомоморфизмом f из связности Γ_2 . $\square$

Следствие. Пусть $f : \mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}) \rightarrow \mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G})$ – отображение расслоений с одинаковой структурной группой такое, что $f_{\mathbb{G}} = \text{id}_{\mathbb{G}}$ – тождественный автоморфизм. Если ω_2 – форма связности на $\mathbb{P}_2$, то отображение $f_{\mathbb{P}}$ индуцирует связность на $\mathbb{P}_1$: $\omega_1 = f_{\mathbb{P}}^* \omega_2$. В частности, для данного главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и отображения баз $f_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{M}$ каждая связность на $\mathbb{P}$ индуцирует связность на $f_{\mathbb{N}}^{-1}(\mathbb{P})$. В частном случае, если $\mathbb{U} \subset \mathbb{M}$ – открытое подмножество, то связность Γ на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ индуцирует связность на индуцированном подрасслоении $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$. $\square$

Замечание. В предложении 13.6.1 связность Γ_2 на $\mathbb{P}_2$ строилась таким образом, что гомоморфное отображение групп голономий является сюръективным. В предложении 13.6.4 утверждается, что отображение групп голономий является только гомоморфизмом. $\square$

13.7 Связность на ассоциированном расслоении

Пусть дано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и ассоциированное с ним расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ с типичным слоем $\mathbb{F}$ (см. раздел 12.2). Если на $\mathbb{P}$ задана связность, то она определяет связность на $\mathbb{E}$ следующим образом.

Определение. Пусть $u \in \mathbb{E}$ – произвольная точка ассоциированного расслоения. *Вертикальным подпространством* $\mathbb{V}_u(\mathbb{E})$ в касательном пространстве $\mathbb{T}_u(\mathbb{E})$ называется касательное пространство к слою $\pi_{\mathbb{E}}^{-1}(\pi_{\mathbb{E}}(u))$, которое лежит в $\mathbb{T}_u(\mathbb{E})$. $\square$

Чтобы определить горизонтальное подпространство, вспомним, что ассоциированное расслоение строилось с помощью естественной проекции на фактор пространство

$$\mathbb{P} \times \mathbb{F} \rightarrow \mathbb{E} = \mathbb{P} \times_{\mathbb{G}} \mathbb{F}.$$

Выберем точку $(p, v) \in \mathbb{P} \times \mathbb{F}$, которая проектируется на $u \in \mathbb{E}$. Теперь зафиксируем точку типичного слоя $v \in \mathbb{F}$ и рассмотрим отображение

$$v : \mathbb{P} \ni p \mapsto v(p) = u \in \mathbb{E}, \quad (13.66)$$

отображающее точку $p \in \mathbb{P}$ в $v(p) \in \mathbb{E}$. То есть каждой точке типичного слоя v ставится в соответствие отображение (13.66), которое мы обозначаем той же буквой.

Замечание. В отличие от отображения $p : \mathbb{F} \rightarrow \mathbb{F}_x$, определенного ранее (12.15), это отображение в общем случае не является диффеоморфизмом, так как размерности главного и ассоциированного расслоения могут отличаться. Даже если размерности совпадают, $\dim \mathbb{G} = \dim \mathbb{F}$, то этого недостаточно для того, чтобы отображение (13.66) было диффеоморфизмом. Действительно, если v_0 – неподвижная точка группы преобразований, то отображение $v_0(p)$ переводит все точки слоя $\pi^{-1}(x)$ в одну фиксированную точку ассоциированного расслоения $u_0 \in \pi_{\mathbb{E}}^{-1}(x)$. $\square$

Поскольку точки (p, v) и (pa, va) из $\mathbb{P} \times \mathbb{F}$ проектируются в одну и ту же точку ассоциированного расслоения $u \in \mathbb{E}$, то $(va^{-1})(p) = v(pa)$, т.е. диаграмма

$$\begin{array}{ccc} \mathbb{P} & \xrightarrow{a} & \mathbb{P} \\ & \searrow va^{-1} & \downarrow v \\ & & \mathbb{E} \end{array}$$

коммутативна. Используя построенное отображение, определим горизонтальные подпространства в ассоциированном расслоении.

Определение. *Горизонтальным подпространством* в точке $u \in \mathbb{E}$ называется образ $\mathbb{H}_u(\mathbb{E}) = v_*(\mathbb{H}_p(\mathbb{P}))$, где v_* – дифференциал отображения (13.66). $\square$

Легко видеть, что подпространство $\mathbb{H}_u(\mathbb{E})$ не зависит от выбора точек $(p, v) \in \mathbb{P} \times \mathbb{F}$, которые проектируются в точку $u \in \mathbb{E}$. Действительно, поскольку $(va)(p) = v(pa^{-1})$, то $(va)_* = v_* \circ r_{a*}^{-1}$. Поэтому для точки (pa, va) , которая проектируется в ту же точку $u \in \mathbb{E}$, что и (p, v) , справедливо равенство

$$\mathbb{H}_u = (va)_* \mathbb{H}_{pa} = v_* \circ r_{a*}^{-1} \circ r_{a*} \mathbb{H}_p = v_* \mathbb{H}_p,$$

где мы использовали инвариантность (13.3) распределения горизонтальных подпространств на $\mathbb{P}$. В следующем примере мы докажем, что касательное пространство к ассоциированному расслоению представляет собой прямую сумму,

$$\mathbb{T}_u(\mathbb{E}) = \mathbb{V}_u(\mathbb{E}) \oplus \mathbb{H}_u(\mathbb{E}). \quad (13.67)$$

Этого достаточно для определения связности на ассоциированном расслоении. Тем не менее построенная связность на ассоциированном расслоении обладает дополнительным свойством, которое наследуется из главного расслоения: она инвариантна относительно действия группы справа. Действительно, по построению, отображение (13.66) перестановочно с групповым действием. Поэтому перестановочны также дифференциалы этих отображений,

$$v_* \circ r_{a*} = r_{a*} \circ v_*.$$

Тогда из правой инвариантности связности на главном расслоении (13.3) следует инвариантность связности на ассоциированном расслоении:

$$r_{a*} \mathbb{H}_u(\mathbb{E}) = \mathbb{H}_{ua}(\mathbb{E}). \quad (13.68)$$

Пример 13.7.1 (Локальное рассмотрение). Рассмотрим достаточно малую координатную окрестность на базе, $\mathbb{U} \subset \mathbb{M}$, с координатами x^α , $\alpha = 1, \dots, n$, такую, что выполнены условия локальной тривиализации расслоений, $\pi^{-1}(\mathbb{U}) \approx \mathbb{U} \times \mathbb{G}$ и $\pi_{\mathbb{E}}^{-1}(\mathbb{U}) \approx \mathbb{U} \times \mathbb{F}$. отождествим $\pi^{-1}(\mathbb{U})$ с $\mathbb{U} \times \mathbb{G}$ и $\pi_{\mathbb{E}}^{-1}(\mathbb{U})$ с $\mathbb{U} \times \mathbb{F}$. Ограничим наше рассмотрение окрестностью единицы группы Ли, $\mathbb{U}_{\mathbb{G}} \subset \mathbb{G}$, где определены координаты a^A , $A = 1, \dots, N$, и функция композиции (см. раздел 8.1). Выберем также некоторую координатную окрестность в типичном слое, $\mathbb{U}_{\mathbb{F}} \subset \mathbb{F}$, где определены координаты v^i , $i = 1, \dots, \dim \mathbb{F}$, и задано отображение $v \mapsto va$ в координатной форме. Тогда точки расслоений будут иметь координаты $p = \{x^\alpha, a^A\} \in \mathbb{P}$ и $u = \{x^\alpha, v^i\} \in \mathbb{E}$. По определению, векторы ∂_i касательны к слою $\pi_{\mathbb{E}}^{-1}(x)$ и, следовательно, образуют базис вертикальных подпространств $\mathbb{V}_u(\mathbb{E})$ для всех $u \in \mathbb{E}$. Пусть $v_0 = \{v_0^i\} \in \mathbb{U}_{\mathbb{F}}$ – фиксированная точка типичного слоя. Тогда соответствующее этой точке отображение (13.66) в координатах имеет вид

$$v_0 : \quad \mathbb{U} \times \mathbb{U}_{\mathbb{G}} \ni \quad \{x^\alpha, a^A\} \mapsto \{x^\alpha, v^i(v_0, a)\} \quad \in \mathbb{U} \times \mathbb{U}_{\mathbb{F}},$$

где $v^i(v_0, a)$ – некоторая функция координат v_0^i и a^A . Пусть $D_\alpha = \partial_\alpha - \omega_\alpha^A L_A^*$, где ω_α^A есть компоненты формы связности (13.4), – базис горизонтальных подпространств $\mathbb{H}_p(\mathbb{P})$ в главном расслоении, который был построен ранее (13.40). Тогда он отображается в касательное пространство к ассоциированному расслоению,

$$v_{0*} : \quad \mathbb{H}_p(\mathbb{P}) \ni \quad D_\alpha \mapsto \hat{D}_\alpha = \partial_\alpha - \omega_\alpha^A L_A^* v^i \partial_i \quad \in \mathbb{H}_u(\mathbb{E}),$$

где L_A^* действует на $v^i(v_0, a)$ как дифференцирование по a . Конкретный вид функций $v^i(v_0, a)$ зависит от типичного слоя и действия на нем структурной группы. Независимо от вида функций $v^i(v_0, a)$ векторы $\hat{D}_\alpha$ линейно независимы, т.к. содержат ∂_α ,

и поэтому образуют базис горизонтального подпространства $\mathbb{H}_u(\mathbb{E})$. Таким образом, совокупность векторов $\{\hat{D}_\alpha, \partial_i\}$ образует базис касательного пространства $\mathbb{T}_u(\mathbb{E})$, который соответствует разложению (13.67). Аналогичное построение можно выполнить в окрестности произвольной точки ассоциированного расслоения $u \in \mathbb{E}$. Следовательно, разложение касательного пространства $\mathbb{T}_u(\mathbb{E})$ в прямую сумму (13.67) имеет место в общем случае.

Если типичный слой – это векторное пространство $\mathbb{V}$, на котором задано представление ρ (см., раздел 8.8), то

$$v^i(v_0, a) = v_0^j S_j^{-1i}(a),$$

где $S_j^i(a)$ – матрица представления элемента $a \in \mathbb{G}$. В этом случае

$$\omega_\alpha^A L_A^* v^i = \omega_\alpha^A v_0^j S_j^k L_{Ak}^i = -v^j \omega_{\alpha j}^i,$$

где введено обозначение $\omega_{\alpha j}^i := -\omega_\alpha^A L_{Aj}^i$ и мы воспользовались правилом дифференцирования матриц представления (8.67). Тогда базис горизонтальных векторных полей на ассоциированном расслоении имеет вид

$$\hat{D}_\alpha = \partial_\alpha + v^j \omega_{\alpha j}^i \partial_i. \quad (13.69)$$

Этот базис инвариантен относительно действия структурной группы $\mathbb{G}$ на ассоциированном расслоении $\mathbb{E}$ справа. Действительно, действие элемента $b \in \mathbb{G}$ на базисный вектор (13.69) имеет вид

$$r_{b*} \hat{D}_\alpha = \partial_\alpha + v^i \omega_{\alpha i}^j(u) S_j^{-1k}(b) \partial_k,$$

т.к. $v^i(v_0, ab) = v^j(v_0, a) S_j^{-1i}(b)$. Поскольку $v^i(ub) = v^j(u) S_j^{-1i}(b)$ и $\omega_\alpha^A(pb) = \omega_\alpha^B S_B^{-1A}(b)$, то

$$r_{b*} \hat{D}_\alpha|_u = \hat{D}_\alpha|_{ub},$$

где мы воспользовались инвариантностью (8.68) матриц представления генераторов структурной группы. Это соответствует инвариантности распределения горизонтальных подпространств на ассоциированном расслоении (13.68). $\square$

Определения горизонтального лифта и параллельного переноса для ассоциированных расслоений дословно повторяют определения, данные для главных расслоений.

Определение. Кривая $\hat{\gamma}$ в ассоциированном расслоении $\mathbb{E}$ называется *горизонтальной*, если касательный к ней вектор горизонтален в каждой точке. Если задана кривая γ в базе $\mathbb{M}$, то *горизонтальным лифтом* или просто *лифтом* этой кривой называется такая горизонтальная кривая $\hat{\gamma}$ в $\mathbb{E}$, что $\pi_{\mathbb{E}}(\hat{\gamma}) = \gamma$. $\square$

Предложение 13.7.1. Пусть $\gamma = x(t)$, $t \in [0, 1]$, – кусочно дифференцируемая кривая класса C^1 в $\mathbb{M}$ с началом в точке $x_0 \in \mathbb{M}$. Тогда для произвольной точки слоя $u_0 \in \pi_{\mathbb{E}}^{-1}(x_0)$ существует единственный горизонтальный лифт $\hat{\gamma} = u(t)$ кривой γ с началом в точке u_0 .

Доказательство. Сначала докажем существование горизонтального лифта. Выберем точку $(p_0, v_0) \in \mathbb{P} \times \mathbb{F}$ такую, что $p_0(v_0) = u_0$, где отображение p определено

формулой (12.15). Согласно предложению 13.3.1 существует единственный горизонтальный лифт $\tilde{\gamma}$ кривой γ в главное расслоение $\mathbb{P}$ с началом в точке $p_0 \in \mathbb{P}$. Тогда кривая $\hat{\gamma} = \tilde{\gamma}(v_0)$ является горизонтальным лифтом кривой γ в ассоциированное расслоение $\mathbb{E}$. Действительно, касательный вектор к кривой $\hat{\gamma}$ лежит в $\mathbb{H}_\gamma(\mathbb{E})$, что сразу следует из определения дифференциала отображения (2.72). Единственность горизонтального лифта следует из единственности решения системы линейных дифференциальных с заданными начальными условиями. $\square$

Для формулировки следующего утверждения нам понадобится естественное

Определение. Локальное сечение ассоциированного расслоения $\sigma : \mathbb{U} \rightarrow \mathbb{E}$, определенное на открытом подмножестве $\mathbb{U} \subset \mathbb{M}$, называется *параллельным или горизонтальным*, если образ $\sigma_* T_x(\mathbb{M})$, где σ_* – дифференциал сечения, горизонтален при всех $x \in \mathbb{U}$, т.е. для любой кривой γ в $\mathbb{U}$, соединяющей точки x_0 и x_1 , точка слоя $\sigma(x_0)$ при параллельном переносе слоя вдоль кривой γ переходит в точку $\sigma(x_1)$. $\square$

В предложении 12.2.3 ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ мы отождествили с фактор пространством $\mathbb{P}/\mathbb{H}$. Затем в теореме 12.3.2 привели критерий редуцируемости структурной группы $\mathbb{G}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ к подгруппе $\mathbb{H}$, который заключается в существовании глобального сечения σ ассоциированного расслоения $\mathbb{E}$. Кроме того, была установлена естественная взаимно однозначная связь между сечениями σ ассоциированного расслоения $\mathbb{E}$ и редуцированными главными расслоениями $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$. В примере 12.3.2 эта теорема была использована для доказательства существования римановой метрики на произвольном многообразии. Возникает вопрос о том, в каком случае связность, заданная на главном расслоении $\mathbb{P}$, редуцируема к связности на редуцированном расслоении $\mathbb{Q}$. Ответ дает следующее утверждение.

Предложение 13.7.2. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение и $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ ассоциированное расслоение со стандартным слоем $\mathbb{G}/\mathbb{H}$, где $\mathbb{H}$ – замкнутая подгруппа в $\mathbb{G}$. Пусть $\sigma : \mathbb{M} \rightarrow \mathbb{E}$ – глобальное сечение ассоциированного расслоения и $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$ редуцированное подрасслоение в $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, соответствующее сечению σ . Связность Γ на $\mathbb{P}$ редуцируема к связности Γ' на $\mathbb{Q}$ тогда и только тогда, когда сечение σ параллельно относительно Γ .

Доказательство. См., например, [48], глава II, предложение 7.4. $\square$

13.8 Свойства групп голономий

Продолжим изучение свойств групп голономий, которое было начато в разделе 13.4.

Теорема 13.8.1 (Теорема редукции). Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение со связностью Γ и p – произвольная точка в $\mathbb{P}$. Обозначим через $\mathbb{P}(p)$ множество точек в $\mathbb{P}$, которые можно соединить с точкой p горизонтальными кусочно дифференцируемыми кривыми. Тогда:

- 1) $\mathbb{P}(p)$ – редуцированное главное расслоение со структурной группой $\Phi(p)$;
- 2) Связность Γ редуцируема к связности на $\mathbb{P}(p)$.

Доказательство. См., например, [48], глава II, теорема 7.1. $\square$

Эта теорема оправдывает следующее

Определение. Главное расслоение $\mathbb{P}(p)$, с базой $\mathbb{M}$, проекцией π и структурной группой $\Phi(p)$, состоящее из множества точек в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, которые можно соединить с точкой p горизонтальными кусочно дифференцируемыми кривыми, называется *расслоением голономии* через p . $\square$

Очевидно, что $\mathbb{P}(p) = \mathbb{P}(q)$ тогда и только тогда, когда точки p и q можно соединить горизонтальной кривой. В разделе 13.4 было введено отношение эквивалентности: $p \sim q$, если p и q можно соединить горизонтальной кривой. Поэтому для каждой пары точек из главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ либо $\mathbb{P}(p) = \mathbb{P}(q)$, либо $\mathbb{P}(p) \cap \mathbb{P}(q) = \emptyset$. Другими словами, главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ разлагается в объединение,

$$\mathbb{P}(\mathbb{M}, \pi, \mathbb{G}) = \bigcup_{p \in \mathbb{P}} \mathbb{P}(p),$$

попарно непересекающихся расслоений голономии. Так как каждый элемент $a \in \mathbb{G}$ отображает каждую горизонтальную кривую в горизонтальную, то $r_a \mathbb{P}(p) = \mathbb{P}(pa)$ и отображение

$$r_a : \mathbb{P}(p) \rightarrow \mathbb{P}(pa)$$

индуцирует изоморфизм расслоений $f = (f_{\mathbb{P}}, f_{\mathbb{G}})$, где $f_{\mathbb{P}} = r_a$, с соответствующим изоморфизмом структурных групп:

$$f_{\mathbb{G}} = \text{ad}(a^{-1}) : \Phi(p) \ni b \mapsto a^{-1}ba \in \Phi(pa),$$

т.к. группы голономии в различных точках сопряжены друг другу (предложение 13.4.1). Легко видеть, что для двух произвольных точек $p, q \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ существует такой элемент $a \in \mathbb{G}$, что $\mathbb{P}(p) = \mathbb{P}(qa)$. Поэтому все расслоения голономий $\mathbb{P}(p)$ изоморфны друг другу.

Теорема 13.8.2 (Амброз–Зингер). Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение со связной базой $\mathbb{M}$. Пусть Γ – связность на $\mathbb{P}$ с формой кривизны R , $\Phi(p_0)$ – группа голономии в точке $p_0 \in \mathbb{P}$ и $\mathbb{P}(p_0)$ – расслоение голономии через p_0 . Тогда алгебра Ли группы голономии $\Phi(p_0)$ совпадает с подпространством в алгебре Ли $\mathfrak{g}$ структурной группы $\mathbb{G}$, порожденной всеми элементами вида $R_p(\tilde{X}, \tilde{Y})$ для всех $p \in \mathbb{P}(p_0)$ и всех горизонтальных векторных полей $\tilde{X}, \tilde{Y}$ в точке p .

Доказательство. Используется паракомпактность $\mathbb{M}$ [78]. $\square$

Теорема 13.8.3. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение со связным пространством расслоения $\mathbb{P}$. Если $\dim \mathbb{M} \geq 2$, то существует связность на $\mathbb{P}$ такая, что расслоения голономии $\mathbb{P}(p)$ для всех $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ совпадают с главным расслоением $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$.

Доказательство. Явное построение связности. При этом используется паракомпактность $\mathbb{M}$ [48], глава II, теорема 8.2. Для линейных связностей это утверждение было доказано в [79]. В общем случае доказательство дано в [80]. $\square$

Следствие. Любая связная группа Ли $\mathbb{G}$ может быть реализована как группа голономии некоторой связности в тривиальном главном расслоении $\mathbb{P} = \mathbb{M} \times \mathbb{G}$, где $\mathbb{M}$ – произвольное дифференцируемое многообразие размерности $\dim \mathbb{M} \geq 2$. $\square$

Доказательство. Выберем связную окрестность $\mathbb{U} \subset \mathbb{M}$. Тогда расслоение $\mathbb{U} \times \mathbb{G}$ связно и мы попадаем в зону деятельности теоремы 13.8.3. Связность с $\mathbb{U} \times \mathbb{G}$ продолжается на связность на $\mathbb{M} \times \mathbb{G}$ согласно предложению 13.6.1. $\square$

13.9 Плоские связности

Рассмотрим тривиальное главное расслоение $\mathbb{P} = \mathbb{M} \times \mathbb{G}$. Для каждого элемента структурной группы $a \in \mathbb{G}$ множество $\mathbb{M} \times \{a\}$ есть подмногообразие в $\mathbb{P}$. В частности, $\mathbb{M} \times \{e\}$, где e – единица группы, есть редуцированное подрасслоение в $\mathbb{P}$.

Определение. *Канонической плоской связностью* на тривиальном главном расслоении $\mathbb{P} = \mathbb{M} \times \mathbb{G}$ называется распределение горизонтальных подпространств $\mathbb{H}_p(\mathbb{P})$, образованное касательными пространствами к $\mathbb{M} \times \{a\}$ для всех точек $p = (x, a) \in \mathbb{M} \times \mathbb{G}$. $\square$

Предложение 13.9.1. *Связность Γ на тривиальном главном расслоении $\mathbb{P} = \mathbb{M} \times \mathbb{G}$ является канонической плоской тогда и только тогда, когда она редуцируема к единственной связности на $\mathbb{M} \times \{e\}$.*

Доказательство. Выберем нулевое сечение $\sigma_0 : x \mapsto (x, e)$. Это сечение является главным расслоением $\mathbb{P}(\mathbb{M}, \pi, e)$, на котором существует единственная связность. Эта связность взаимно однозначно определяет каноническую плоскую связность на $\mathbb{P} = \mathbb{M} \times \mathbb{G}$. $\square$

Пусть θ – каноническая форма на группе Ли $\mathbb{G}$, определенная в разделе 8.2. Обозначим через $\text{pr} : \mathbb{M} \times \mathbb{G} \rightarrow \mathbb{G}$ естественную проекцию и положим

$$\omega := \text{pr}^* \theta = \omega^{*A} L_A. \quad (13.70)$$

Эта 1-форма является частным случаем формы связности (13.5) и определяет каноническую плоскую связность на $\mathbb{P}$. Формула Маурера–Картана для канонической 1-формы (8.28) влечет, что каноническая плоская связность имеет нулевую кривизну, т.к.

$$d\omega = d(\text{pr}^* \theta) = \text{pr}^*(d\theta) = \text{pr}^* \left(-\frac{1}{2} [\theta, \theta] \right) = -\frac{1}{2} [\text{pr}^* \theta, \text{pr}^* \theta] = -\frac{1}{2} [\omega, \omega].$$

Сравнивая полученное равенство со структурным уравнением (13.27), заключаем, что форма кривизны канонической плоской связности тождественно равна нулю, $R = 0$.

Теперь рассмотрим случай произвольного главного расслоения. Дадим общее

Определение. Связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ называется *плоской*, если каждая точка базы $x \in \mathbb{M}$ имеет окрестность $\mathbb{U}$ такую, что индуцированная связность на $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$ изоморфна канонической плоской связности на $\mathbb{U} \times \mathbb{G}$. Другими словами, существует изоморфизм $\chi : \pi^{-1}(\mathbb{U}) \rightarrow \mathbb{U} \times \mathbb{G}$, отображающий горизонтальное подпространство в каждой точке $p \in \pi^{-1}(\mathbb{U})$ в горизонтальное подпространство канонической плоской связности на $\mathbb{U} \times \mathbb{G}$ в точке $\chi(p)$. $\square$

Предложение 13.9.2. *Пусть задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Тогда плоская связность на $\mathbb{P}$ существует и единственна с точностью до изоморфизма.*

Доказательство. Если задано главное расслоение, то определен атлас на базе $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и семейство функций перехода (12.7), которые по теореме 12.1.2 с точностью до изоморфизма определяют главное расслоение. Выбрав координатное покрытие базы достаточно малым, можно считать, что все координатные окрестности $\mathbb{U}_i$ соответствуют окрестностям, входящим в определение плоской связности. Согласно теореме 13.1.1 для однозначного задания связности на $\mathbb{P}$ достаточно задать семейство

локальных форм связности на каком-либо атласе базы. Это означает, что плоская связность на произвольном главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ существует и единственна с точностью до изоморфизма. $\square$

Теорема 13.9.1. *Связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ является плоской тогда и только тогда, когда ее форма кривизны равна нулю, $R = 0$.*

Доказательство. Необходимость очевидна. Обратно. Допустим, что форма кривизны равна нулю. Пусть $\mathbb{U}$ – односвязная окрестность точки $x \in \mathbb{P}$ и рассмотрим индуцированную связность на $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$. По теоремам 13.4.1 и Амброза–Зингера 13.8.2 группа голономии индуцированной связности на $\mathbb{P}|_{\mathbb{U}}$ состоит только из единицы. Применяя теорему редукции 13.8.1, мы видим, что индуцированная связность на $\mathbb{P}|_{\mathbb{U}}$ изоморфна канонической плоской связности на $\mathbb{U} \times \mathbb{G}$. $\square$

Следствие. Любая связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с одномерной базой $\mathbb{M}$ является плоской. $\square$

Доказательство. Любая 2-форма на одномерном многообразии равна нулю. Отсюда следует, что все локальные формы кривизны тоже равны нулю. Так как для формы кривизны только горизонтальные компоненты являются нетривиальными, то она также обращается в нуль. $\square$

Следствие. Пусть Γ – связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ такая, что ее форма кривизны равна нулю, $R = 0$. Если база $\mathbb{M}$ односвязна, то главное расслоение $\mathbb{P}$ изоморфно тривиальному расслоению $\mathbb{M} \times \mathbb{G}$ и связность Γ изоморфна канонической плоской связности на $\mathbb{M} \times \mathbb{G}$. $\square$

Доказательство. Группа голономии в рассматриваемом случае состоит из единственного элемента – единицы. Поэтому расслоение голономии $\mathbb{P}(p)$ пересекает каждый слой ровно в одной точке. Следовательно, каждое расслоение голономии задает глобальное сечение, и поэтому главное расслоение тривиально. При этом горизонтальные подпространства касательны к расслоению голономии. Пусть $q \in \mathbb{P}(p)$ – произвольная точка расслоения голономии через p и $\sigma_0 = (x, e)$ – нулевое сечение главного расслоения $\mathbb{P} = \mathbb{M} \times \mathbb{G}$. Тогда для каждой точки базы $x \in \mathbb{M}$ существует единственный элемент $a(x) \in \mathbb{G}$ такой, что $\sigma_0 = qa$, где $x = \pi(q)$. При этом вертикальный автоморфизм $p \mapsto pa$ переводит связность Γ на $\mathbb{P} = \mathbb{M} \times \mathbb{G}$ в каноническую плоскую связность. $\square$

Если форма кривизны равна нулю, то распределение горизонтальных векторных полей находится в инволюции. Это сразу следует из (13.41), т.к. векторы D_α образуют базис распределения горизонтальных векторных полей. Согласно теореме Фробениуса для плоской связности через каждую точку p главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ проходит интегральное подмногообразие. Все касательные векторы к интегральным подмногообразиям горизонтальны и интегральное подмногообразие, проходящее через точку p , совпадает с расслоением голономии $\mathbb{P}(p)$ через p .

Пример 13.9.1 (Локальное рассмотрение). Пусть $\mathbb{Q} = \mathbb{U} \times \mathbb{G}$ – тривиальное главное расслоение, база $\mathbb{U}$ которого покрыта одной картой. Общий вид формы связности на $\mathbb{Q}$ был найден ранее, см. формулу (13.5). Сравнение этого выражения с выражением (13.70) показывает, что связность на $\mathbb{Q}$ является канонической плоской тогда и только тогда, когда часть ее компонент, описывающих произвол в выборе связности на нулевом сечении, обращаются в нуль, $\overset{\circ}{A}_\alpha^\wedge(x) = 0$. Из равенства (13.13) следует, что

компоненты локальной формы канонической плоской связности для произвольного сечения $\sigma : x \mapsto (x, b(x))$ имеют вид

$$A_\alpha{}^A(x) = \partial_\alpha b^B L^{-1}{}_B{}^A(b).$$

После перехода к какому-либо представлению структурной группы

$$\rho : \mathbb{G} \ni a \mapsto \{S_i^j(a)\} \in \mathbf{aut} \mathbb{V},$$

для компонент локальной формы связности справедливо равенство (13.15). Поскольку $\overset{\circ}{A}_\alpha{}^A = 0$, то компоненты плоской связности в общем случае имеют вид

$$A_\alpha = \partial_\alpha S S^{-1}, \quad (13.71)$$

где мы, для простоты, опустили матричные индексы.

Определение. Калибровочное поле A_α , заданное равенством (13.71) на координатной окрестности U , называется *чистой калибровкой*. $\square$

Теперь рассмотрим случай, когда база $\mathbb{M}$ главного расслоения не является односвязной. Пусть Γ – связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ со связной базой $\mathbb{M}$. Выберем произвольную точку $p_0 \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и обозначим через $\tilde{\mathbb{M}} = \mathbb{P}(p_0)$ расслоение голономии через p_0 . Тогда $\tilde{\mathbb{M}}(\mathbb{M}, \pi, \Phi(p_0))$ есть главное расслоение над $\mathbb{M}$ со структурной группой $\Phi(p_0)$. Так как суженная группа голономии $\Phi_0(p_0)$ для плоской связности всегда тривиальна, то по теоремам 13.4.1 и Амброза–Зингера 13.8.2 группа голономии $\Phi(p_0)$ при неодносвязной базе дискретна. Поэтому отображение $\pi : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ является накрытием со связным накрывающим пространством.

Пусть $x_0 = \pi(p_0) \in \mathbb{M}$. Каждая замкнутая кривая в базе $\mathbb{M}$, исходящая из x_0 , при помощи параллельного переноса слоев вдоль нее определяет некоторый элемент группы голономии $\Phi(p_0)$. Поскольку суженная группа голономии тривиальна, то любые две замкнутые и гомотопные относительно начала кривые, представляющие один и тот же элемент фундаментальной группы $\pi_1(\mathbb{M}, x_0)$, порождают один и тот же элемент из $\Phi(p_0)$. Таким образом, мы получаем сюръективное отображение фундаментальной группы $\pi_1(\mathbb{M}, p_0)$ на группу голономии $\Phi(p_0)$. Легко видеть, что это отображение является гомоморфизмом групп. Пусть $\mathbb{H}$ – нормальная подгруппа в $\Phi(p_0)$ и положим $\mathbb{M}' = \tilde{\mathbb{M}}/\mathbb{H}$. Тогда $\mathbb{M}'$ – главное расслоение над $\mathbb{M}$ со структурной группой $\Phi(p_0)/\mathbb{H}$. В частности, отображение $\pi : \mathbb{M}' \rightarrow \mathbb{M}$ – накрытие. Пусть $\mathbb{P}'(\mathbb{M}', \pi, \mathbb{G})$ – главное расслоение, индуцированное из $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ накрывающей проекцией $f_{\mathbb{M}} = \pi : \mathbb{M}' \rightarrow \mathbb{M}$. Пусть $f : \mathbb{P}'(\mathbb{M}', \pi, \mathbb{G}) \rightarrow \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – естественный гомоморфизм главных расслоений, см. теорему 12.3.4, тогда справедливо

Предложение 13.9.3. *Существует единственная связность Γ' на главном расслоении $\mathbb{P}'(\mathbb{M}', \pi, \mathbb{G})$, которая отображается на связность Γ на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ гомоморфизмом $f : \mathbb{P}' \rightarrow \mathbb{P}$. Связность Γ' плоская. Если точка $p'_0 \in \mathbb{P}'$ такова, что $f_{\mathbb{M}}(p'_0) = p_0$, то группа голономии $\Phi(p'_0)$ для связности Γ' изоморфно отображается на $\mathbb{H}$ гомоморфизмом $f_{\mathbb{G}}$.*

Доказательство. См., например, [48], глава II, предложение 9.3. $\square$

В данном утверждении всегда можно выбрать в качестве нормальной подгруппы единицу группы голономии $e \in \Phi(p_0)$.

Следствие. Пусть $\tilde{\mathbb{P}}(\tilde{\mathbb{M}}, \pi, \mathbb{G})$ – главное расслоение, индуцированное из $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ накрывающей проекцией $f_{\mathbb{M}} = \pi : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$. Тогда существует единственная связность $\tilde{\Gamma}$ на главном расслоении $\tilde{\mathbb{P}}(\tilde{\mathbb{M}}, \pi, \mathbb{G})$, которая отображается на связность Γ на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ гомоморфизмом $f : \tilde{\mathbb{P}} \rightarrow \mathbb{P}$. Связность $\tilde{\Gamma}$ плоская, и соответствующая ей группа голономии тривиальна.

13.10 Локальные и инфинитезимальные группы голономии

Группа голономии $\Phi(p)$ в точке $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, которая является важнейшей глобальной характеристикой связности, была определена при помощи множества всех замкнутых путей в базе $\Omega(\mathbb{M}, \pi(p))$. Это крайне неудобно для практических вычислений, т.к. зачастую структура многообразия $\mathbb{M}$ просто неизвестна. Возникает вопрос, можно ли каким-либо образом вычислить группу голономии, исходя из локальных характеристик связности? В некоторых случаях суженная группа голономии $\Phi_0(p)$ действительно определяется локальными свойствами связности. В настоящем разделе мы опишем два подхода к этой проблеме, которые основаны на понятиях локальной и инфинитезимальной групп голономии.

Начнем рассмотрение с локальной группы голономии. Пусть на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ со связной базой $\mathbb{M}$ задана связность Γ . Для каждого связного и односвязного открытого подмножества базы $\mathbb{U} \subset \mathbb{M}$ обозначим через $\Gamma_{\mathbb{U}}$ связность на $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$, которая индуцирована из связности Γ . В силу следствия из предложения 13.6.4 связность $\Gamma_{\mathbb{U}}$ существует и единственна – это сужение связности Γ на $\pi^{-1}(\mathbb{U})$. Для каждой точки $p \in \pi^{-1}(\mathbb{U})$ обозначим через $\Phi_0(p, \mathbb{U})$ и $\mathbb{P}(p, \mathbb{U})$ суженную группу голономии с опорной точкой p и расслоение голономии через точку p для связности $\Gamma_{\mathbb{U}}$, соответственно. Напомним, что расслоение голономии $\mathbb{P}(p, \mathbb{U})$ состоит из тех точек $q \in \pi^{-1}(\mathbb{U})$, которые можно соединить с точкой p горизонтальной кривой, целиком лежащей в $\pi^{-1}(\mathbb{U})$.

Рассмотрим две окрестности $\mathbb{U}_2 \subset \mathbb{U}_1$, которые содержат точку $x = \pi(p)$. Тогда всякая замкнутая петля, целиком лежащая в $\mathbb{U}_2$, будет также петлей в $\mathbb{U}_1$. Поэтому справедливо включение

$$\Phi_0(p, \mathbb{U}_2) \subset \Phi_0(p, \mathbb{U}_1).$$

Как подгруппа суженной группы голономии $\Phi_0(p)$ группа $\Phi_0(p, \mathbb{U}_1)$ однозначно определяется своей алгеброй Ли. Поэтому из равенства размерностей $\dim \Phi_0(p, \mathbb{U}_2) = \dim \Phi_0(p, \mathbb{U}_1)$ следует совпадение групп голономии $\Phi_0(p, \mathbb{U}_2) = \Phi_0(p, \mathbb{U}_1)$. Это наблюдение приводит к следующему понятию.

Определение. Локальной группой голономии $\Phi_{\text{loc}}(p)$ в точке p называется пересечение

$$\Phi_{\text{loc}}(p) := \bigcap_{\mathbb{U}} \Phi_0(p, \mathbb{U}) \quad (13.72)$$

по всем связным и односвязным открытым окрестностям $\mathbb{U}$ точки $x = \pi(p)$. $\square$

Замечание. В настоящем разделе мы будем рассматривать только связные и односвязные открытые окрестности точек. Поэтому в дальнейшем, для краткости, мы будем говорить просто окрестности. $\square$

Пусть $\{\mathbb{U}_k\}$, $k = 1, 2, \dots$, – последовательность окрестностей, сходящихся к точке x , т.е. $\mathbb{U}_k \supset \overline{\mathbb{U}_{k+1}}$ и $\bigcap_{k=1}^{\infty} \mathbb{U}_k = \{x\}$. Тогда, очевидно, имеют место включения

$$\Phi_0(p, \mathbb{U}_1) \supset \Phi_0(p, \mathbb{U}_2) \supset \Phi_0(p, \mathbb{U}_3) \supset \dots$$

Поскольку для каждой окрестности $\mathbb{U}$ точки x существует целое число $k_{\mathbb{U}}$ такое, что $\mathbb{U}_k \subset \mathbb{U}$ для всех $k > k_{\mathbb{U}}$, то локальная группа голономии представима в виде

$$\Phi_{\text{loc}}(p) = \bigcap_{k=1}^{\infty} \Phi_0(p, \mathbb{U}_k).$$

Так как каждая суженная группа голономии $\Phi_0(p, \mathbb{U}_k)$ есть связная подгруппа Ли в структурной группе $\mathbb{G}$ (теорема 13.4.1), то отсюда следует, что размерность суженной группы голономии $\dim \Phi_0(p, \mathbb{U}_k)$ постоянна для достаточно больших k . Поэтому для достаточно больших значений k справедливо равенство $\Phi_{\text{loc}}(p) = \Phi_0(p, \mathbb{U}_k)$.

Замечание. Конечно, суженная группа голономии $\Phi_0(p, x)$ для “окрестности”, состоящей из одной точки x , состоит ровно из одного элемента – единицы, и ее размерность равна нулю. Допустим, что параметр k в последовательности $\{\mathbb{U}_k\}$ непрерывен. Это может быть, например, радиус шара, если последовательность $\mathbb{U}_k$ состоит из шаров. Тогда функция $\dim \Phi_0(p, \mathbb{U}_k)$ от k принимает значения в целых числах и не может быть непрерывной, если суженная группа голономии $\Phi_0(p)$ для всего главного расслоения нетривиальна. В дальнейшем мы увидим, что при определенных условиях локальная группа голономии $\Phi_{\text{loc}}(p)$ совпадает с суженной группой $\Phi_0(p)$. $\square$

Предложение 13.10.1. *Локальные группы голономии имеют следующие свойства:*

- 1) *локальная группа голономии $\Phi_{\text{loc}}(p)$ есть связная подгруппа Ли в структурной группе $\mathbb{G}$, содержащаяся в суженной группе голономии $\Phi_0(p)$;*
- 2) *каждая точка $x = \pi(p)$ имеет окрестность $\mathbb{U}$ такую, что $\Phi_{\text{loc}}(p) = \Phi_0(p, \mathbb{V})$ для любой окрестности $\mathbb{V} \ni x$, содержащейся в $\mathbb{U}$;*
- 3) *если $\mathbb{U}$ – окрестность точки $x = \pi(p)$, о которой говорится в свойстве 2), то $\Phi_{\text{loc}}(p) \supset \Phi_{\text{loc}}(q)$ для всех точек $q \in \mathbb{P}(p, \mathbb{U})$;*
- 4) *для каждого $a \in \mathbb{G}$ справедливо равенство*

$$\Phi_{\text{loc}}(pa) = \text{ad}(a^{-1})\Phi_{\text{loc}}(p);$$

- 5) *для каждого целого t множество точек базы*

$$\{\pi(p) = x \in \mathbb{M} : \dim \Phi_{\text{loc}}(p) \leq t\}$$

открыто в $\mathbb{M}$.

Доказательство. Свойства 1)–4) очевидны. Докажем свойство 5). Из свойства 4) следует, что функция $\dim \Phi_{\text{loc}}(p)$ постоянна на каждом слое и поэтому ее можно рассматривать как функцию на базе $\mathbb{M}$, принимающую целые значения. Из свойств 3) и 4) вытекает, что

$$\dim \Phi_{\text{loc}}(q) \leq \dim \Phi_{\text{loc}}(p),$$

для всех точек $q \in \pi^{-1}(\mathbb{U})$. Отсюда вытекает свойство 5). $\square$

Теорема 13.10.1. *Пусть $\mathfrak{g}_0(p)$ и $\mathfrak{g}_{\text{loc}}(p)$ – алгебры Ли для групп голономий $\Phi_0(p)$ и $\Phi_{\text{loc}}(p)$, соответственно. Тогда $\Phi_0(p)$ и $\mathfrak{g}_0(p)$ порождаются, соответственно, всеми $\Phi_{\text{loc}}(q)$ и $\mathfrak{g}_{\text{loc}}(q)$ для всех точек $q \in \mathbb{P}(p)$.*

Доказательство. См., например, [48], глава II, теорема 10.2. $\square$

Теорема 13.10.2. Если $\dim \Phi_{\text{loc}}(p)$ постоянна на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, то локальная и суженная группы голономии совпадают, $\Phi_{\text{loc}}(p) = \Phi_0(p)$, для всех $p \in \mathbb{P}$.

Доказательство. По свойству 3) предложения 13.10.1 каждая точка $x = \pi(p)$ имеет окрестность $\mathbb{U}$ такую, что $\Phi_{\text{loc}}(p) \supset \Phi_{\text{loc}}(q)$ для каждой точки q из расслоения голономии $\mathbb{P}(p, \mathbb{U})$. Так как $\dim \Phi_{\text{loc}}(p) = \dim \Phi_{\text{loc}}(q)$, то сами группы голономии в различных точках совпадают, $\Phi_{\text{loc}}(p) = \Phi_{\text{loc}}(q)$. Отсюда следует, что если $q \in \mathbb{P}(p)$, то $\Phi_{\text{loc}}(p) = \Phi_{\text{loc}}(q)$ для всех $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Из теоремы 13.10.1 вытекает равенство $\Phi_0(p) = \Phi_{\text{loc}}(p)$. $\square$

Теперь перейдем к определению инфинитезимальной группы голономии и изучим ее связь с локальной группой голономии. Инфинитезимальная группа голономии может быть определена только для гладких C^∞ главных расслоений $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с гладкой C^∞ связностью Γ . В дальнейшем мы будем считать, что условие гладкости выполнено.

Инфинитезимальная группа голономии в точке главного расслоения $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ определяется при помощи формы кривизны R связности Γ , заданной на $\mathbb{P}$. Сначала определим индуктивно семейство подпространств $\mathfrak{m}_k(p)$, $k = 0, 1, 2, \dots$, в алгебре Ли $\mathfrak{g}$ структурной группы $\mathbb{G}$. Пусть $\mathfrak{m}_0(p)$ подпространство в $\mathfrak{g}$, порожденное всеми элементами вида $R_p(X, Y)$, где X, Y – произвольные горизонтальные векторы в точке $p \in \mathbb{P}$. Рассмотрим $\mathfrak{g}$ -значную функцию на $\mathbb{P}$ вида

$$f_k := Z_k \dots Z_1(R(X, Y)), \quad (13.73)$$

где $X, Y, Z_1, \dots, Z_k$ – произвольные горизонтальные векторные поля на $\mathbb{P}$ и векторы $Z_1, \dots, Z_k$ действуют на $\mathfrak{g}$ -значную функцию $R(X, Y)$ как дифференцирования. Пусть $\mathfrak{m}_k(p)$ – подпространство в алгебре Ли $\mathfrak{g}$, порожденное подпространством $\mathfrak{m}_{k-1}$ и значениями в точке p всех функций f_k вида (13.73). Положим

$$\mathfrak{g}_{\text{inf}}(p) := \bigcup_{k=0}^{\infty} \mathfrak{m}_k(p). \quad (13.74)$$

По сути дела, это то подмножество в алгебре Ли структурной группы, которое порождается формой кривизны и всеми ее производными.

Предложение 13.10.2. Подпространство $\mathfrak{g}_{\text{inf}}(p)$ в $\mathfrak{g}$ есть подалгебра Ли в алгебре Ли $\mathfrak{g}_{\text{loc}}(p)$ локальной группы голономии.

Доказательство. См. [81]. $\square$

Это предложение позволяет ввести новое понятие.

Определение. Связная подгруппа Ли $\Phi_{\text{inf}}(p)$ в структурной группе $\mathbb{G}$, порожденная подалгеброй $\mathfrak{g}_{\text{inf}}(p)$ называется *инфинитезимальной группой голономии* связности Γ в точке $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. $\square$

Предложение 13.10.3. Инфинитезимальная группа голономии $\Phi_{\text{inf}}(p)$ имеет следующие свойства:

- 1) $\Phi_{\text{inf}}(p)$ является связной подгруппой Ли локальной группы голономии $\Phi_{\text{loc}}(p)$;
- 2) $\Phi_{\text{inf}}(pa) = \text{ad}(a^{-1})\Phi_{\text{inf}}(p)$ и $\mathfrak{g}_{\text{inf}}(pa) = \text{ad}(a^{-1})\mathfrak{g}_{\text{inf}}(p)$;

3) для каждого целого m множество точек базы

$$\{\pi(p) = x \in \mathbb{M} : \dim \Phi_{\text{inf}}(p) \geq m\} \quad (13.75)$$

открыто в $\mathbb{M}$;

4) если $\Phi_{\text{inf}}(p) = \Phi_{\text{loc}}(p)$ в точке p , то существует окрестность $\mathbb{U}$ точки $x = \pi(p)$ такая, что

$$\Phi_{\text{inf}}(q) = \Phi_{\text{loc}}(q) = \Phi_{\text{inf}}(p) = \Phi_{\text{loc}}(p)$$

для всех $q \in \mathbb{U}$.

Доказательство. Свойство 1) следует из предложения 13.10.2.

Свойство 2) вытекает из интуитивно понятного равенства

$$\mathfrak{m}_k(pa) = \text{ad}(a^{-1})\mathfrak{m}_k(p), \quad \forall k.$$

Детали доказательства приведены в [48], глава II, предложение 10.5.

3). Размерность инфинитезимальной группы голономии $\dim \Phi_{\text{inf}}(p)$ в силу свойства 2) можно рассматривать как функцию на $\mathbb{M}$ со значениями в целых числах. Если значения конечного числа функций f_k вида (13.73) линейно независимы в точке p , то они таковы и в любой точке из некоторой окрестности точки p . Поэтому, если свойство (13.75) выполнено в точке x , то оно имеет место и в некоторой окрестности x .

4). Допустим, что $\Phi_{\text{inf}}(p) = \Phi_{\text{loc}}(p)$. Из свойства 3) предложения 13.10 и свойства 5) предложения 13.10.1 следует, что точка $x = \pi(p)$ имеет окрестность $\mathbb{U}$ такую, что

$$\dim \Phi_{\text{inf}}(q) \geq \dim \Phi_{\text{inf}}(p) \quad \text{и} \quad \dim \Phi_{\text{loc}}(q) \leq \dim \Phi_{\text{loc}}(p), \quad \forall q \in \pi^{-1}(\mathbb{U}).$$

С другой стороны, $\Phi_{\text{loc}}(q) \supset \Phi_{\text{inf}}(q)$ для каждого $q \in \pi^{-1}(\mathbb{U})$. Отсюда вытекают равенства

$$\dim \Phi_{\text{loc}}(q) = \dim \Phi_{\text{inf}}(q) = \dim \Phi_{\text{loc}}(p) = \dim \Phi_{\text{inf}}(p)$$

и, следовательно, $\Phi_{\text{loc}}(q) = \Phi_{\text{inf}}(q)$ для каждого $q \in \pi^{-1}(\mathbb{U})$. Применяя теорему 13.10.2 к индуцированному расслоению $\mathbb{P}|_{\mathbb{U}}$, видим, что $\Phi_0(p, \mathbb{U}) = \Phi_{\text{loc}}(p)$ и $\Phi_0(q, \mathbb{U}) = \Phi_{\text{loc}}(q)$. Если $q \in \mathbb{P}(p, \mathbb{U})$, то $\Phi_0(p, \mathbb{U}) = \Phi_0(q, \mathbb{U})$. Поэтому $\Phi_{\text{loc}}(p) = \Phi_{\text{loc}}(q)$. $\square$

Теорема 13.10.3. Если размерность инфинитезимальной группы голономии $\dim \Phi_{\text{inf}}(q)$ постоянна в некоторой окрестности точки $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, то локальная и инфинитезимальная группы голономий в точке p совпадают, $\Phi_{\text{loc}}(p) = \Phi_{\text{inf}}(p)$.

Доказательство. См. [81] $\square$

Следствие. Если размерность инфинитезимальной группы голономии $\dim \Phi_{\text{inf}}(p)$ постоянна на всем главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, то

$$\Phi_0(p) = \Phi_{\text{loc}}(p) = \Phi_{\text{inf}}(p) \quad (13.76)$$

для всех $p \in \mathbb{P}$. $\square$

Доказательство. Вытекает из теорем 13.10.2 и 13.10.3. $\square$

Теорема 13.10.4. Для вещественно аналитической связности Γ на вещественно аналитическом главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ следующие группы голономии равны

$$\Phi_0(p) = \Phi_{\text{loc}}(p) = \Phi_{\text{inf}}(p)$$

для всех $p \in \mathbb{P}$.

Доказательство. См. [81]. $\square$

Равенство групп голономий (13.76) позволяет вычислить суженную группу голономии $\Phi_0(p)$. Действительно, алгебра Ли инфинитезимальной группы голономии (13.74) порождена всеми функциями вида (13.73). Пусть x^α , $\alpha = 1, \dots, n$, – система координат в окрестности $\mathbb{U}$ точки $x = \pi(p)$ и $\sigma : \mathbb{U} \rightarrow \mathbb{P}$ – локальное сечение. Тогда алгебра инфинитезимальной группы голономии порождается значениями компонент локальной формы кривизны $F_{\alpha\beta}^A$ и всех ее ковариантных производных:

$$F_{\alpha\beta}^A, \nabla_{\gamma_1} F_{\alpha\beta}^A, \nabla_{\gamma_2} \nabla_{\gamma_1} F_{\alpha\beta}^A, \dots$$

Ясно, что подпространство $\mathfrak{m}_0 \subset \mathfrak{g}$ порождено всеми компонентами тензора кривизны $F_{\alpha\beta}^A$, подпространство $\mathfrak{m}_1 \subset \mathfrak{g}$ – всеми компонентами тензора кривизны $F_{\alpha\beta}^A$ и их первых ковариантных производных $\nabla_{\gamma_1} F_{\alpha\beta}^A$ и так далее. Таким образом, при заданной связности на главном расслоении, инфинитезимальная группа голономии позволяет, в принципе, вычислить суженную группу голономии. Конечно, после этого необходимо проверить, что условие следствия выполнено.

13.11 Инвариантные связности

Прежде чем рассматривать инвариантные связности общего вида, мы опишем важный частный случай.

Теорема 13.11.1. Пусть $\mathbb{G}$ – связная группа Ли, и $\mathbb{H}$ – ее замкнутая подгруппа Ли. Пусть $\mathfrak{g}$ и $\mathfrak{h}$ – алгебры Ли для $\mathbb{G}$ и $\mathbb{H}$, соответственно.

- 1) Если существует линейное подпространство $\mathfrak{m}$ в $\mathfrak{g}$ такое, что $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ и $\text{ad}(\mathbb{H})\mathfrak{m} = \mathfrak{m}$, то $\mathfrak{h}$ -компонента ω канонической 1-формы θ в $\mathbb{G}$ определяет связность на главном расслоении $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, инвариантную относительно действия левых сдвигов из $\mathbb{G}$.
- 2) Обратно, любая связность на главном расслоении $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, инвариантная относительно действия левых сдвигов из $\mathbb{G}$ (если она существует), определяет разложение $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ и может быть получена так, как это описано в пункте 1).
- 3) Форма кривизны R инвариантной связности, определенная формой ω из пункта 1), равна

$$R(X, Y) = -\frac{1}{2}[X, Y]_{\mathfrak{h}},$$

где $X, Y \in \mathfrak{m}$ – произвольные левоинвариантные векторные поля на $\mathbb{G}$ из $\mathfrak{m}$ и в правой части равенства взята $\mathfrak{h}$ -компонента коммутатора.

- 4) Пусть $\mathfrak{g}(e)$ – алгебра Ли группы голономии $\Phi(e)$ в единице e группы Ли $\mathbb{G}$ для инвариантной связности, определенной в пункте 1). Тогда $\mathfrak{g}(e)$ порождается всеми элементами вида $[X, Y]_{\mathfrak{h}}$, где $X, Y \in \mathfrak{m}$.

Доказательство. 1). Пусть Z^* – фундаментальное векторное поле, соответствующее элементу подалгебры $Z \in \mathfrak{h}$. Из определения канонической 1-формы (см. раздел 8.2) следует, что $\omega(Z^*) = \theta(Z^*) = Z$. Пусть $\theta_{\mathfrak{m}}$ есть $\mathfrak{m}$ -компонента канонической формы θ . Для любого $a \in \mathbb{H}$ и $X \in \mathbb{T}_p(\mathbb{G})$ справедливы равенства

$$\begin{aligned} \theta(r_{a*}X) &= \omega(r_{a*}X) + \theta_{\mathfrak{m}}(r_{a*}X), \\ \text{ad}(a^{-1})\theta(X) &= \text{ad}(a^{-1})\omega(X) + \text{ad}(a^{-1})\theta_{\mathfrak{m}}(X). \end{aligned}$$

Левые части этих равенств совпадают. Поскольку по условию теоремы $\text{ad}(a^{-1})\mathfrak{m} = \mathfrak{m}$, то сравнение $\mathfrak{h}$ -компонент правых частей приводит к равенству

$$\omega(r_{a*}X) = \text{ad}(a^{-1})\omega(X),$$

т.е. 1-форма ω определяет связность на $\mathbb{G}$. Эта связность инвариантна относительно действия группы слева по построению.

2). Пусть ω – форма связности на главном расслоении $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, инвариантная относительно действия левых сдвигов из $\mathbb{G}$. Пусть $X \in \mathfrak{m}$ – множество левоинвариантных векторных полей на $\mathbb{G}$ таких, что $\omega(X) = 0$. Тогда алгебра Ли на $\mathbb{G}$ разлагается в прямую сумму, $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$.

3). Левоинвариантное векторное поле горизонтально тогда и только тогда, когда оно лежит в $\mathfrak{m}$. Поэтому утверждение 3) следует из равенства (13.28).

4). Пусть $\mathfrak{g}_1$ – подпространство в $\mathfrak{g}$, порожденное множеством элементов вида $R_e(X, Y)$, где $X, Y \in \mathfrak{m}$. Пусть $\mathfrak{g}_2$ – подпространство в $\mathfrak{g}$, порожденное множеством элементов $R_a(X, Y)$, где $X, Y \in \mathfrak{m}$ для всех $a \in \mathbb{G}$. Тогда по теореме Амброза–Зингера $\mathfrak{g}_1 \subset \mathfrak{g}(e) \subset \mathfrak{g}_2$. С другой стороны, $\mathfrak{g}_1 = \mathfrak{g}_2$, т.к. $R_a(X, Y) = R_e(X, Y)$ для любых $X, Y \in \mathfrak{m}$ и $a \in \mathbb{G}$. Теперь 4) следует из 3). $\square$

Замечание. Линейное подпространство алгебры Ли $\mathfrak{m} \subset \mathfrak{g}$ представляет собой распределение горизонтальных векторных полей на $\mathbb{G}$, т.е. связность Γ на главном расслоении $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, инвариантную относительно действия группы слева. В общем случае $\mathfrak{m} \subset \mathfrak{g}$ является только линейным подпространством, а не подалгеброй. $\square$

Замечание. Утверждение 1) теоремы 13.11.1 можно рассматривать как частный случай предложения 13.6.3. Пусть $\mathbb{P} = (\mathbb{G}/\mathbb{H}) \times \mathbb{G}$ – тривиальное главное расслоение над базой $\mathbb{G}/\mathbb{H}$ со структурной группой $\mathbb{G}$. Вложим расслоение $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$ в $\mathbb{P}$ при помощи отображения

$$f(a) = (\pi(a), a), \quad a \in \mathbb{G},$$

где $\pi : \mathbb{G} \rightarrow \mathbb{G}/\mathbb{H}$ – естественная проекция. Пусть ϕ – форма связности, определяющая каноническую плоскую связность на $\mathbb{P}$. Ее $\mathfrak{h}$ -компонента, суженная на подрасслоение $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, по предложению 13.6.3 определяет связность и совпадает с формой связности ω в утверждении 1. $\square$

Возвращаясь к общему случаю, сначала докажем предложение, которое является основой для многих приложений.

Предложение 13.11.1. Пусть s_t – однопараметрическая группа автоморфизмов главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и $\tilde{X}$ – векторное поле на $\mathbb{P}$, индуцированное s_t . Пусть Γ – связность на $\mathbb{P}$, инвариантная относительно действия s_t . Для произвольной точки главного расслоения $p_0 \in \mathbb{P}$ определим четыре кривые p_t , x_t , $\tilde{x}_t$ и a_t следующим образом:

$$p_t := s_t(p_0) \in \mathbb{P}, \quad x_t := \pi(p_t) \in \mathbb{M},$$

$\tilde{x}_t$ есть горизонтальный лифт x_t такой, что $\tilde{x}_0 = p_0$ и

$$p_t := \tilde{x}_t a_t, \quad a_t \in \mathbb{G}. \quad (13.77)$$

Тогда a_t является однопараметрической подгруппой в структурной группе $\mathbb{G}$, порожденной элементом $X = \omega_{p_0}(\tilde{X})$, где ω – форма связности для Γ .

Доказательство. Поскольку

$$\dot{p}_t = r_{a_t*} \dot{\tilde{x}}_t + \tilde{x}_{t*} \dot{a}_t,$$

то справедливо равенство

$$\omega(\dot{p}_t) = \text{ad}(a_t^{-1})\omega(\dot{\tilde{x}}_t) + a_{t*}^{-1}\dot{a}_t,$$

где $\tilde{x}_{t*}$ и a_{t*} – дифференциалы отображений $\tilde{x}_t : \mathbb{G} \rightarrow \pi^{-1}(x_t)$ и $a_t : \mathbb{G} \rightarrow \mathbb{G}$, соответственно. Первое слагаемое в правой части равно нулю, т.к. кривая $\tilde{x}_t$ горизонтальна. Следовательно, $\omega(\dot{p}_t) = a_{t*}^{-1}\dot{a}_t$. С другой стороны, $\dot{p}_t = s_{t*}\tilde{X}_{p_0}$, где s_{t*} – дифференциал отображения $s_t : \mathbb{P} \rightarrow \mathbb{P}$, и поэтому $\omega(\dot{p}_t) = \omega(\tilde{X}_{p_0}) = X$, т.к. форма связности инвариантна относительно однопараметрической группы преобразований s_t . Отсюда следует равенство $X = a_{t*}^{-1}\dot{a}_t$. $\square$

Определение. Кривая a_t в структурной группе Ли $\mathbb{G}$ из условия предложения 13.11.1 называется *разверткой* кривой p_t в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. $\square$

Замечание. В этом определении кривая p_t может быть произвольной кривой в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, не обязательно связанной с группой симметрии связности. $\square$

Определение. Пусть $\mathbb{K}$ – группа Ли автоморфизмов главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с алгеброй Ли $\mathfrak{k}$. Выберем в главном расслоении опорную точку $p_0 \in \mathbb{P}$. Каждый элемент из $\mathbb{K}$ индуцирует некоторое преобразование базы при помощи проекции. Множество $\mathbb{J}$ всех элементов из $\mathbb{K}$, которые оставляют неподвижной точку $x_0 = \pi(p_0)$, образуют замкнутую подгруппу в $\mathbb{K}$, которая называется *подгруппой изотропии* в $\mathbb{K}$ для точки $x_0 \in \mathbb{M}$. Определим гомоморфизм групп Ли

$$\lambda : \mathbb{J} \ni j \mapsto \lambda(j) = a \in \mathbb{G} \quad (13.78)$$

следующим образом. Каждый автоморфизм $j \in \mathbb{J}$ переводит p_0 в точку $j(p_0)$, которая принадлежит тому же слою $\pi^{-1}(\pi(p_0))$, т.к. точка x_0 неподвижна. Следовательно, $j(p_0) = p_0 a$ для некоторого $a \in \mathbb{G}$. Положим $\lambda(j) = a$. Тогда для двух элементов подгруппы изотропии, $j_1, j_2 \in \mathbb{J}$, справедливы равенства

$$p_0 \lambda(j_1 j_2) = (j_1 j_2)(p_0) = j_1(p_0 \lambda(j_2)) = (j_1(p_0)) \lambda(j_2) = p_0 \lambda(j_1) \lambda(j_2).$$

Тем самым $\lambda(j_1 j_2) = \lambda(j_1) \lambda(j_2)$ и, следовательно, построенное отображение является гомоморфизмом групп. Нетрудно проверить, что отображение $\lambda : \mathbb{J} \rightarrow \mathbb{G}$ дифференцируемо. Гомоморфизм групп λ индуцирует гомоморфизм алгебр Ли,

$$\lambda : \mathfrak{j} \rightarrow \mathfrak{g},$$

который мы обозначили той же буквой. $\square$

Замечание. Отображение (13.78) зависит от выбора опорной точки $p_0 \in \mathbb{P}$. В настоящем разделе мы будем считать, что опорная точка p_0 выбрана и зафиксирована. $\square$

Предложение 13.11.2. Пусть $\mathbb{K}$ – группа Ли автоморфизмов главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с алгеброй Ли $\mathfrak{k}$ и Γ – связность на $\mathbb{P}$ с формой связности ω и формой

кривизны R , инвариантная относительно автоморфизмов $\mathbb{K}$. Пусть $\mathbb{J} \subset \mathbb{K}$ – подгруппа изотропии в $\mathbb{K}$ для точки $x_0 := \pi(p_0)$ с алгеброй Ли $\mathfrak{j}$. Определим линейное отображение

$$\Lambda : \mathfrak{k} \ni X \mapsto \Lambda(X) = \omega_{p_0}(\tilde{X}) \in \mathfrak{g},$$

где $\tilde{X}$ – векторное поле на $\mathbb{P}$, индуцированное полем X . Тогда:

- 1) $\Lambda(X) = \lambda(X)$, $\forall X \in \mathfrak{j}$;
- 2) $\Lambda(\text{ad}(j)X) = \text{ad}(\lambda(j))\Lambda(X)$, $\forall j \in \mathbb{J}$ и $\forall X \in \mathfrak{k}$, где $\text{ad}(j)$ обозначает присоединенное представление подгруппы изотропии $\mathbb{J}$ в $\mathfrak{k}$ и $\text{ad}(\lambda(j))$ – присоединенное представление структурной группы $\mathbb{G}$ в $\mathfrak{g}$;
- 3) $2R_{p_0}(\tilde{X}, \tilde{Y}) = [\Lambda(X), \Lambda(Y)] - \Lambda([X, Y])$, $\forall X, Y \in \mathfrak{k}$.

Доказательство. См., например, [48], глава II, предложение 11.3. $\square$

Замечание. Геометрический смысл отображения Λ дается предложением 13.11.1. $\Lambda(X)$ – это тот элемент алгебры Ли $\mathfrak{g}$, который порождает однопараметрическую подгруппу a_t в структурной группе $\mathbb{G}$, определенную равенством (13.77). $\square$

Замечание. Отображение Λ в предложении 13.11.2 является только линейным. В общем случае оно не является гомоморфизмом алгебр Ли. $\square$

Определение. Группа автоморфизмов $\mathbb{K}$ действует на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ слой-транзитивно, если для любых двух слоев из $\mathbb{P}$ существует элемент в $\mathbb{K}$, отображающий один слой в другой, т.е. если действие $\mathbb{K}$ на базе $\mathbb{M}$ транзитивно. $\square$

Предложение 13.11.3. Если $\mathbb{J}$ – подгруппа изотропии для слой-транзитивной группы автоморфизмов $\mathbb{K}$ в точке $x_0 := \pi(p_0)$, то база $\mathbb{M}$ является однородным пространством, $\mathbb{M} \approx \mathbb{K}/\mathbb{J}$.

Доказательство. Вытекает из теоремы 9.1.2. $\square$

Описание $\mathbb{K}$ -инвариантных связностей на главном расслоении в случае слой-транзитивного действия групп автоморфизмов дает

Теорема 13.11.2. Если связная группа Ли $\mathbb{K}$ является слой-транзитивной группой автоморфизмов главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и если $\mathbb{J}$ – подгруппа изотропии в $\mathbb{K}$ для точки $x_0 := \pi(p_0)$, то существует взаимно однозначное соответствие между множеством $\mathbb{K}$ -инвариантных связностей на $\mathbb{P}$ и множеством линейных отображений $\Lambda : \mathfrak{k} \rightarrow \mathfrak{g}$, которые удовлетворяют условиям 1) и 2) предложения 13.11.2. Соответствие задается следующим образом:

$$\Lambda : \mathfrak{k} \ni X \mapsto \Lambda(X) = \omega_{p_0}(\tilde{X}) \in \mathfrak{g},$$

где $\tilde{X}$ – векторное поле на $\mathbb{P}$, индуцированное полем X .

Доказательство. См. [82]. $\square$

Замечание. Из теоремы 13.1.1 следует, что связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ однозначно определяется заданием семейства локальных форм связности на произвольном покрытии базы $\mathbb{M}$. Это происходит потому что при помощи действия структурной группы распределение горизонтальных подпространств с локальных сечений можно разнести на все пространство главного расслоения $\mathbb{P}$. Если

известно, что связность инвариантна также относительно слой-транзитивных автоморфизмов $\mathbb{K}$, то локальную форму связности достаточно задать в одной точке базы $\mathbb{M}$. Действительно, при помощи автоморфизмов $\mathbb{K}$ она разносится по всей базе $\mathbb{M}$, а затем, действуя структурной группой, ее можно разнести по всему пространству расслоения $\mathbb{P}$.

Если группа автоморфизмов $\mathbb{K}$ не является слой-транзитивной, тогда в задании инвариантной связности появляется значительный произвол. В этом случае форму связности можно определить на каждой орбите действия группы $\mathbb{K}$ на $\mathbb{P}$. Для этого достаточно задать ее в какой-либо одной точке на каждой орбите и при помощи $\mathbb{K}$ разнести ее по орбитам и, наконец, по всему пространству главного расслоения $\mathbb{P}$, действуя структурной группой $\mathbb{G}$. После этого необходимо проверить дифференцируемость полученного распределения горизонтальных подпространств. $\square$

Если группа автоморфизмов $\mathbb{K}$ действует на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ слой-транзитивно, то форма кривизны R , которая является тензориальной формой типа $\text{ad } \mathbb{G}$, инвариантная относительно $\mathbb{K}$, полностью определяется своими значениями в опорной точке $R_{p_0}(\tilde{X}, \tilde{Y})$, где $\tilde{X}$ и $\tilde{Y}$ – векторные поля на $\mathbb{P}$, индуцированные элементами $X, Y \in \mathfrak{k}$ алгебры Ли группы автоморфизмов $\mathbb{K}$. В этом случае утверждение 3) предложения 13.11.2 выражает форму кривизны $R_{p_0}(\tilde{X}, \tilde{Y})$ в терминах Λ .

Из предложения 13.11.2 и теорем 13.9.1 и 13.11.2 получаем

Следствие. $\mathbb{K}$ -инвариантная связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, определенная отображением Λ , является плоской тогда и только тогда, когда отображение $\Lambda : \mathfrak{k} \rightarrow \mathfrak{g}$ есть гомоморфизм алгебр Ли. $\square$

Теорема 13.11.3. *Допустим, что в теореме 13.11.2 алгебра Ли $\mathfrak{k}$ содержит линейное подпространство $\mathfrak{m}$ такое, что $\mathfrak{k} = \mathfrak{j} \oplus \mathfrak{m}$ и $\text{ad}(\mathbb{J})\mathfrak{m} = \mathfrak{m}$, где $\text{ad}(\mathbb{J})$ – присоединенное представление $\mathbb{J}$ в $\mathfrak{k}$. Тогда:*

- 1) *Существует взаимно однозначное соответствие между множеством $\mathbb{K}$ -инвариантных связностей на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и множеством линейных отображений $\Lambda_{\mathfrak{m}} : \mathfrak{m} \rightarrow \mathfrak{g}$ таких, что*

$$\Lambda_{\mathfrak{m}}(\text{ad}(j)X) = \text{ad}(\lambda(j))\Lambda_{\mathfrak{m}}(X), \quad \forall X \in \mathfrak{m} \text{ и } \forall j \in \mathbb{J};$$

соответствие задается теоремой 13.11.2 следующим образом

$$\Lambda(X) = \begin{cases} \lambda(X), & \text{если } X \in \mathfrak{j}, \\ \Lambda_{\mathfrak{m}}(X), & \text{если } X \in \mathfrak{m}. \end{cases}$$

- 2) *Форма кривизны R для $\mathbb{K}$ -инвариантной связности, определяемой при помощи отображения $\Lambda_{\mathfrak{m}}$, удовлетворяет следующему равенству*

$$2R_{p_0}(\tilde{X}, \tilde{Y}) = [\Lambda_{\mathfrak{m}}(X), \Lambda_{\mathfrak{m}}(Y)] - \Lambda_{\mathfrak{m}}([X, Y]_{\mathfrak{m}}) - \lambda([X, Y]_{\mathfrak{j}}), \quad \forall X, Y \in \mathfrak{m},$$

где $[X, Y]_{\mathfrak{m}}$ и $[X, Y]_{\mathfrak{j}}$ обозначают, соответственно, $\mathfrak{m}$ - и $\mathfrak{j}$ -компоненту коммутатора $[X, Y] \in \mathfrak{k}$.

Доказательство. Пусть $\Lambda : \mathfrak{k} \rightarrow \mathfrak{g}$ – линейное отображение, удовлетворяющее утверждениям 1) и 2) предложения 13.11.2. Пусть $\Lambda_{\mathfrak{m}}$ – сужение отображения Λ на $\mathfrak{m}$. Нетрудно проверить, что отображение $\Lambda \rightarrow \Lambda_{\mathfrak{m}}$ является взаимно однозначным и согласно теореме 13.11.2 дает желаемое соответствие. Утверждение 2) следует из утверждения 3) предложения 13.11.2. $\square$

Определение. В теореме 13.11.3 $\mathbb{K}$ -инвариантная связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, определяемая условием $\Lambda_{\mathfrak{m}} = 0$, называется *канонической инвариантной связностью* относительно разложения $\mathfrak{g} = \mathfrak{j} \oplus \mathfrak{m}$. $\square$

Следующая теорема определяет алгебру Ли группы голономии $\mathbb{K}$ -инвариантной связности.

Теорема 13.11.4. В предположениях и обозначениях теоремы 13.11.2 алгебра Ли $\mathfrak{g}(p_0)$ группы голономии $\Phi(p_0)$ для $\mathbb{K}$ -инвариантной связности, определяемой при помощи линейного отображения $\Lambda : \mathfrak{k} \rightarrow \mathfrak{g}$, задается суммой

$$\mathfrak{m}_0 + [\Lambda(\mathfrak{k}), \mathfrak{m}_0] + [\Lambda(\mathfrak{k}), [\Lambda(\mathfrak{k}), \mathfrak{m}_0]] + \dots,$$

где $\mathfrak{m}_0$ – линейное подпространство в $\mathfrak{g}$, порожденное множеством

$$\{[\Lambda(X), \Lambda(Y)] - \Lambda([X, Y])\}, \quad \forall X, Y \in \mathfrak{k}.$$

Доказательство. См., например, [48], глава II, теорема 11.8. $\square$

Замечание. Утверждения 1) и 3) теоремы 13.11.1 следуют из теоремы 13.11.3, если в качестве главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ выбрать $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$ и положить $\mathbb{K} = \mathbb{G}$. Тогда инвариантная связность из теоремы 13.11.1 является канонической инвариантной связностью. Утверждение 4) теоремы 13.11.1 следует из теоремы 13.11.4. $\square$

Глава 14

Приложения в квантовой механике

В настоящей главе рассмотрены некоторые приложения дифференциальной геометрии в нерелятивистской квантовой механике. Нетривиальные геометрические структуры, а речь идет о нетривиальной связности на главном расслоении, часто возникают при решении уравнений математической физики. В настоящей главе будет показано, как возникает нетривиальная связность на главном расслоении со структурной группой $U(1)$ или $U(N)$ в нерелятивистской квантовой механике при решении уравнения Шредингера. Удивительно не столько то, что главное расслоение возникает естественным образом, а то, что предсказанные эффекты были подтверждены экспериментально.

Сначала мы дадим геометрическую интерпретацию нерелятивистской квантовой механике в конечномерном случае. Будет показано, что гамильтониан квантовой системы задает компоненты локальной формы связности на главном расслоении, а уравнение Шредингера определяет параллельный перенос слоев. При этом базой является одномерное многообразие, соответствующее времени, а структурной группой – унитарная группа $U(N)$, где N – размерность гильбертова пространства состояний квантовомеханической системы. Решение квантовомеханической задачи не зависит от выбора базиса в гильбертовом пространстве, и его выбирают из соображений удобства. Использование базиса, состоящего из собственных векторов гамильтониана, позволяет упростить доказательство адиабатической теоремы и сделать его более прозрачным.

В качестве приложения адиабатической теоремы рассмотрена фаза Берри [83]. В заключительном разделе настоящей главы рассмотрен эффект Ааронова–Бома [84], который, хотя и не имеет прямого отношения к адиабатической теореме, с геометрической точки зрения аналогичен фазе Берри.

Эффект Ааронова–Бома и фаза Берри привлекают большое внимание теоретиков и экспериментаторов в течение многих лет. Интерес вызван двумя обстоятельствами. Во-первых, в обоих случаях при решении уравнения Шредингера естественным образом возникает $U(1)$ связность. Во-вторых, в теории калибровочных полей распространено мнение, что к наблюдаемым эффектам может приводить только нетривиальная напряженность поля, а не сами потенциалы, которые не являются калибровочно инвариантными. Вопреки этому мнению Ааронов и Бом, а также Берри показали, что интеграл от калибровочного поля вдоль замкнутой кривой может привести к наблюдаемым эффектам. Эти выводы вскоре были подтверждены экспериментально.

Понятие фазы Берри было обобщено на неабелев случай, соответствующий вырожденным уровням энергии гамильтониана, Вилчеком и Зи [85]. В этом случае при решении уравнения Шредингера естественным образом возникают неабелевы калиб-

ровочные поля.

Во всех перечисленных выше случаях к наблюдаемым эффектам приводят элементы группы голономии (см. раздел 13.4) соответствующих связностей. Элементы группы голономии в общем случае являются ковариантными объектами, а для абелевой группы $\mathbb{U}(1)$ – инвариантными. Мы покажем, что главное расслоение может быть тривиальным, но связность, которая на нем возникает, в общем случае имеет нетривиальную группу голономии и приводит к наблюдаемым эффектам. Отсюда следует, что фаза Берри и эффект Ааронова–Бома имеют геометрическую природу.

14.1 Адиабатическая теорема

Адиабатическая теорема [86] занимает одно из центральных мест в нерелятивистской квантовой механике, т.к. позволяет находить приближенное решение уравнения Шредингера при медленном изменении гамильтониана во времени. Первоначально она была доказана для дискретного (возможно, бесконечного) спектра гамильтониана при некоторых ограничениях на возможное пересечение уровней энергии [86]. Ниже приведено доказательство адиабатической теоремы в наиболее простом конечномерном случае.

В нерелятивистской квантовой механике состояние системы описывается вектором гильбертова пространства (*волновой функцией*) $\psi(t) \in \mathbb{H}$, зависящим от времени $t \in \mathbb{R}$ и некоторого набора других переменных, который определяется рассматриваемой задачей. Эволюция квантовой системы во времени описывается *уравнением Шредингера* [87, 88]

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi, \quad (14.1)$$

где H – самосопряженный линейный оператор, действующий в гильбертовом пространстве $\mathbb{H}$, который называется *гамильтонианом* системы, и $\hbar$ – постоянная Планка.

Для уравнения Шредингера, как правило, ставится задача Коши с начальным условием

$$\psi(0) = \psi_0, \quad (14.2)$$

где $\psi_0 \in \mathbb{H}$ – некоторый фиксированный вектор гильбертова пространства.

В дальнейшем, для простоты, положим $\hbar = 1$ и обозначим частную производную по времени точкой, $\dot{\psi} := \partial_t \psi$.

Предположим, для простоты, что гильбертово пространство представляет собой конечномерное комплексное пространство $\mathbb{H} = \mathbb{C}^n$ комплексной размерности $\dim_{\mathbb{C}} \mathbb{H} = n$. В гильбертовом пространстве задано скалярное произведение, которое обозначим круглыми скобками,

$$\mathbb{H} \times \mathbb{H} \ni \psi, \phi \mapsto (\psi, \phi) \in \mathbb{C}.$$

По определению, скалярное произведение линейно по первому аргументу ψ и выполнено равенство: $(\psi, \phi)^\dagger = (\phi, \psi)$, где символ $\dagger$ обозначает комплексное сопряжение. Квадрат вектора гильбертова пространства (ψ, ψ) является вещественным числом, при этом мы требуем, чтобы квадратичная форма (ψ, ψ) была строго положительно определена, т.е. $(\psi, \psi) \geq 0$, причем $(\psi, \psi) = 0$ тогда и только тогда, когда $\psi = 0$. Тогда скалярное произведение определяет норму вектора гильбертова пространства:

$$\|\psi\| := \sqrt{(\psi, \psi)}.$$

Поскольку уравнение Шредингера линейно по ψ , а гамильтониан самосопряжен, то норма произвольного решения уравнений Шредингера сохраняется во времени. Отсюда следует, что векторы состояния можно нормировать. Обычно предполагается, что векторы состояния нормированы на единицу,

$$\|\psi\| = 1. \quad (14.3)$$

Нормировка вектора состояния не устраняет полностью произвол в выборе вектора гильбертова пространства, т.к. остается произвол в выборе постоянного фазового множителя.

Замечание. Мы используем символ $\dagger$ для обозначения эрмитова сопряжения матриц, т.е. транспонирования матрицы и комплексного сопряжения всех элементов. В частном случае, когда матрица состоит из одного элемента, эрмитово сопряжение совпадает с комплексным. $\square$

Пусть в гильбертовом пространстве $\mathbb{H}$ выбран некоторый базис e_k , $k = 1, \dots, N$. Тогда гамильтониан квантовомеханической системы задается эрмитовой $N \times N$ матрицей, а вектор состояния $\psi = \psi^k(t)e_k$ – строкой из N компонент,

$$\psi = (\psi^1, \dots, \psi^N),$$

где $\psi^1(t), \dots, \psi^N(t)$ – комплекснозначные компоненты вектора. Если базис гильбертова пространства ортонормирован,

$$(e_k, e_l) = \delta_{kl},$$

то скалярное произведение задается равенством

$$(\psi, \phi) = \psi \phi^\dagger = \psi^1 \phi_1^\dagger + \dots + \psi^N \phi_N^\dagger.$$

Рассмотрим задачу Коши (14.1), (14.2) в общем случае, когда гамильтониан системы зависит от времени $H = H(t)$. Для решения этой задачи необходимо выбрать базис в гильбертовом пространстве $\mathbb{H}$. Конечно, решение задачи от выбора базиса не зависит, и его выбирают из соображений удобства. Рассмотрим два случая.

Пусть базис $e_k \in \mathbb{H}$, $k = 1, \dots, N$, ортонормирован и фиксирован, $\dot{e}_k = 0$. Произвольный вектор можно разложить по этому базису $\psi = \psi^k e_k$. При этом гамильтониан задается эрмитовой $N \times N$ матрицей H_l^k , а задача Коши для уравнения Шредингера в компонентах примет вид системы обыкновенных дифференциальных уравнений с некоторыми начальными условиями:

$$\begin{aligned} i\dot{\psi}^k &= \psi^l H_l^k, \\ \psi^k(0) &= \psi_0^k. \end{aligned} \quad (14.4)$$

Замечание. Мы записываем действие гамильтониана в конечномерном случае справа, чтобы согласовать наши обозначения с обозначениями, принятыми в дифференциальной геометрии. Напомним, что в дифференциальной геометрии для векторного поля в координатном базисе принята запись $X = X^\alpha \partial_\alpha$. (Суммирование с десяти до четырех по циферблату часов.) Альтернативная запись $X = \partial_\alpha X^\alpha$ используется для обозначения дивергенции векторного поля. Преобразование координат мы записываем в виде

$$X = X^\alpha \partial_\alpha = X^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x^{\alpha'}} \partial_\beta.$$

То есть действие матрицы преобразования координат записывается справа. С другой стороны, действие операторов в квантовой механике общепринято записывать слева. Поэтому, если мы хотим использовать единообразные обозначения, то необходимо чем то пожертвовать. В конечномерном случае мы будем записывать матрицу, задающую линейный оператор, справа. Это вопрос соглашения, и к нему легко привыкнуть. $\square$

Рассмотрим теперь другой ортонормированный базис b_k , который может зависеть от времени, $b_k = b_k(t)$. Такой базис может оказаться более удобным для решения некоторых задач. Вектор гильбертова пространства ψ можно разложить также по этому базису $\psi = \psi'^k b_k$. Тогда задача Коши (14.4) будет выглядеть по другому:

$$\begin{aligned} i\dot{\psi}'^k &= \psi'^l H_l'^k, \\ \psi'^k(0) &= \psi_0'^k, \end{aligned} \quad (14.5)$$

где $H_l'^k$ – компоненты гамильтониана относительно нового базиса, которые будут определены ниже. Поскольку базисы ортонормированы, то они связаны между собой некоторым унитарным преобразованием:

$$b_k = S_k^l e_l, \quad S \in \mathbb{U}(N), \quad (14.6)$$

которое в общем случае зависит от времени, $S = S(t)$. При этом компоненты вектора гильбертова пространства преобразуются с помощью обратной матрицы,

$$\psi'^k = \psi^l S^{-1}_l{}^k.$$

Отсюда следует выражение для компонент начального вектора гильбертова пространства $\psi_0'^k = \psi_0^l S^{-1}_l{}^k(0)$. Переписав уравнение Шредингера (14.4) в базисе b_k , получим компоненты гамильтониана относительно нового базиса:

$$H' = SHS^{-1} + iS\dot{S}^{-1} = SHS^{-1} - i\dot{S}S^{-1}, \quad (14.7)$$

где мы, для краткости, опустили матричные индексы. Мы видим, что компоненты гамильтониана преобразуются так же, как компоненты локальной формы $\mathbb{U}(N)$ связности (13.16).

Теперь можно дать геометрическую интерпретацию нерелятивистской квантовой механике в конечномерном случае. Пусть время пробегает всю вещественную прямую, $t \in \mathbb{R}$. Тогда мы имеем главное расслоение $\mathbb{P}(\mathbb{R}, \pi, \mathbb{U}(N)) \approx \mathbb{R} \times \mathbb{U}(N)$ с базой $\mathbb{R}$, типичным слоем $\mathbb{U}(N)$ и проекцией $\pi : \mathbb{P} \rightarrow \mathbb{R}$ (см. раздел 12.1). Это расслоение тривиально, т.к. базой является вещественная прямая. Гамильтониан квантовой системы задает компоненты локальной формы $\mathbb{U}(N)$ связности (1-форма на $\mathbb{R}$ со значениями в алгебре Ли):

$$A_t = (iH_t^k) \in \mathfrak{u}(N),$$

где t – координатный ковариантный индекс, принимающий одно значение, который раньше обозначался греческой буквой α . Вектор гильбертова пространства $\psi \in \mathbb{H}$ – это сечение ассоциированного расслоения $\mathbb{E}(\mathbb{R}, \pi_E, \mathbb{H}, \mathbb{U}(N), \mathbb{P})$, типичным слоем которого является гильбертово пространство $\mathbb{H}$. Уравнение Шредингера имеет вид равенства нулю ковариантной производной,

$$\nabla_t \psi = \dot{\psi} + \psi A_t = 0,$$

т.е. задает параллельный перенос вектора гильбертова пространства. При изменении сечения компоненты связности преобразуются по правилу

$$A'_t = SA_t S^{-1} + \dot{S} S^{-1},$$

как и положено компонентам локальной формы связности (13.16). Кривизна этой связности равна нулю, поскольку база одномерна. Поэтому связность является плоской согласно теореме 13.9.1.

Решение задачи Коши для уравнения Шредингера (14.1), (14.2) не зависит от выбора базиса. Поэтому его выбирают из соображений удобства.

Пример 14.1.1. Пусть в фиксированном базисе e_k компоненты вектора гильбертова пространства имеют вид $\psi^k = \psi_0^l U_l^k$, где унитарная матрица $U_l^k(t)$ задает оператор эволюции квантовой системы, который, по определению, удовлетворяет дифференциальному уравнению

$$i\dot{U} = UH,$$

с начальным условием $U_l^k(0) = \delta_l^k$. Тогда нетрудно проверить, что оператор эволюции задает переход к такому базису гильбертова пространства $b_k := U_k^{-1l} e_l$, в котором гамильтониан равен нулю, $H' = 0$. Следовательно, вектор гильбертова пространства, описывающий эволюцию квантовой системы, в этом базисе имеет постоянные компоненты ψ_0^k , которые определяются начальным состоянием. $\square$

Замечание. При преобразовании базиса, которое зависит от времени, эрмитова матрица, соответствующая гамильтониану, испытывает калибровочное преобразование (14.7). При этом преобразовании в общем случае собственные значения матрицы меняются. Рассмотренный выше пример показывает, что если гамильтониан, заданный в постоянном базисе, имел некоторый спектр, то после перехода к новому базису, заданному оператором эволюции, гамильтониан H' обращается в нуль, и имеет только нулевые собственные значения. В квантовой механике уравнение Шредингера обычно задают, определив гамильтониан в постоянном базисе, исходя из физических соображений. Затем, если это удобнее, можно перейти к новому базису, зависящему от времени. $\square$

Перейдем к определению адиабатического предела и описанию базиса $b_k(t)$, который будет использован при доказательстве адиабатической теоремы. Адиабатическая теорема справедлива для гамильтонианов, которые медленно меняются со временем. А именно, мы предполагаем, что гамильтониан некоторой квантовой механической системы достаточно гладко зависит от вещественного параметра $\nu = \epsilon t$, где $\epsilon > 0$, который меняется на конечном отрезке, $\nu \in [0, \nu_0]$. Тогда медленное изменение гамильтониана означает, что параметр ν меняется на конечную величину при малых ϵ и больших временах t .

Определение. Двойной предел в решении задачи Коши для уравнения Шредингера (14.1) и (14.2) на отрезке $[0, t]$

$$\epsilon \rightarrow 0, \quad t \rightarrow \infty, \quad \text{при условии} \quad \epsilon t = \nu = \text{const.} \quad (14.8)$$

называется *адиабатическим*. $\square$

При исследовании адиабатического предела время t в уравнении Шредингера удобно заменить на параметр ν :

$$i\epsilon \frac{\partial \psi}{\partial \nu} = \psi H(\nu). \quad (14.9)$$

Вектор состояния $\psi(\nu, \epsilon)$ в таком случае зависит также от параметра ϵ , а адиабатический предел соответствует простому пределу $\epsilon \rightarrow 0$ при каждом значении параметра $\nu \in [0, \nu_0]$.

Асимптотическое решение уравнения вида (14.9) в общем случае построено в [89, 90].

Для доказательства адиабатической теоремы нам понадобится специальный базис $b_k(\nu)$, зависящий от ν . Пусть исходный гамильтониан $H(\nu)$ квантовой системы задан в некотором фиксированном базисе e_k . Тогда существует унитарная матрица $S(\nu)$, которая диагонализует гамильтониан:

$$SH(\nu)S^{-1} = H_D(\nu) = \text{diag}(E_1(\nu), \dots, E_N(\nu)), \quad (14.10)$$

где $E_1 \leq E_2 \leq \dots \leq E_N$ – уровни энергии собственных состояний гамильтониана H , которые будем считать упорядоченными. Пусть b_k – собственные векторы исходного гамильтониана:

$$b_k H = E_k b_k, \quad (14.11)$$

для всех ν . Как известно, строками матрицы преобразования S_k^l , где индекс k фиксирован и $l = 1, \dots, N$, являются компоненты собственных векторов $b_k = b_k^l e_l$ гамильтониана H : $S_k^l = b_k^l$. То есть гамильтониан $H(\nu)$ в базисе b_k диагонален. Унитарная матрица S определена неоднозначно, и произвол в ее выборе в дальнейшем рассмотрении будет использован.

Мы допускаем, что часть уровней энергии может быть вырождена. Обозначим через Υ_n множество индексов, для которых $E_j(\nu) = E_n(\nu)$ при $j \in \Upsilon_n$. Конечно, в качестве индекса n можно выбрать любой индекс, принадлежащий Υ_n . Если уровень E_n невырожден, то множество индексов состоит из одного элемента: $\Upsilon_n = \{n\}$. Мы докажем адиабатическую теорему в случае, когда множество индексов Υ_n для всех n не меняется со временем, т.е. уровни энергии не пересекаются.

Мы также предполагаем, что гамильтониан H , уровни энергии $E_1, \dots, E_N$ и матрица преобразования S достаточно гладко зависят от ν на конечном отрезке $[0, \nu_0]$.

Для доказательства адиабатической теоремы нам понадобится

Лемма 14.1.1. *Существует унитарная матрица S в (14.10) такая, что выполнено условие*

$$\left(\frac{dS}{d\nu} S^{-1} \right)_k^j = 0, \quad \forall k \in \Upsilon_j. \quad (14.12)$$

Доказательство. Рассмотрим два случая. Пусть уровень энергии E_k невырожден. Матрица преобразования S в формуле (14.10) определена с точностью до умножения каждой строки на фазовый множитель: $S_k^j \mapsto S_k^j e^{i\alpha_k(\nu)}$ для всех $j = 1, \dots, N$. Это соответствует произволу в выборе фазового множителя у собственного вектора состояния (14.11). Пусть фазовый множитель удовлетворяет уравнению

$$\frac{d\alpha_k}{d\nu} = i \sum_{j=1}^N \frac{dS_k^j}{d\nu} S_j^{-1k},$$

где суммирование по k в правой части отсутствует. Тогда нетрудно проверить, что после преобразования для любого решения этого уравнения выполнено равенство

$$\left(\frac{dS}{d\nu} S^{-1} \right)_k^k = 0. \quad (14.13)$$

Это можно проделать для всех невырожденных уровней одновременно, выбрав подходящим образом фазы $\alpha_k(\nu)$.

Теперь предположим, что все уровни энергии вырождены, $E_1 = \dots = E_N$. Тогда матрица S формуле (14.10) определена с точностью до унитарного преобразования:

$$S \mapsto WS, \quad W(\nu) \in \mathbb{U}(N).$$

Пусть матрица W удовлетворяет уравнению

$$\frac{dW}{d\nu} + W \frac{dS}{d\nu} S^{-1} = 0,$$

которое всегда имеет решение. Тогда после преобразования для любого решения будет выполнено равенство (14.12) для всех j, k .

Если вырождена только часть уровней, то соответствующее унитарное преобразование необходимо проделать только с этими уровнями. Таким образом, равенство (14.12) будет выполнено для всех уровней с $E_j = E_k$. $\square$

Доказательство адиабатической теоремы будет проведено в ортонормированном базисе (14.6), где матрица S выбрана таким образом, как описано в лемме 14.1.1. Этот базис состоит из собственных векторов исходного гамильтониана и гамильтониан $H(\nu)$ в нем диагонален (14.10). Компоненты вектора состояния в базисе b_k , как и ранее, пометим штрихом, $\psi = \psi'^k b_k$. Поскольку гамильтониан H в этом базисе диагонален, то квадрат модуля k -той компоненты вектора состояния

$$|(\psi, b_k)|^2 = |\psi'^k|^2,$$

где круглые скобки обозначают скалярное произведение в $\mathbb{H}$, равен вероятности обнаружить квантовую систему в состоянии E_k .

Для формулировки теоремы нам понадобится функция

$$\Delta E_n(\nu) = \min_{j, \sigma} |E_j(\sigma) - E_n(\sigma)|, \quad \forall \sigma \in [0, \nu],$$

где минимум $|E_j - E_n|$ берется по всем j , для которых $E_j \neq E_n$, и всем $\sigma \in [0, \nu]$. Поскольку уровни энергии не пересекаются, то для каждого значения параметра ν функция $\Delta E_n(\nu)$ конечна и равна минимальному расстоянию от уровня энергии E_n до остальных уровней энергии.

Теорема 14.1.1 (Адиабатическая теорема). Пусть гамильтониан $H = H(\nu)$, его собственные состояния $b_k(\nu)$ и уровни энергии $E_k(\nu)$ достаточно гладко зависят от ν на конечном отрезке $\nu \in [0, \nu_0]$. Предположим, что число вырожденных собственных состояний постоянно во времени. Пусть $\psi_{(n)}(\nu, \epsilon)$ – решение уравнения Шредингера, которое в начальный момент времени совпадает с собственным состоянием $b_n(0)$ гамильтониана $H(0)$, соответствующим уровню энергии $E_n(0)$. Тогда в адиабатическом пределе (14.8) справедлива следующая оценка нормы

$$1 - \sum_{j \in \Upsilon_n} |(\psi_{(n)}, b_j)|^2 = \frac{O(\epsilon^2)}{\Delta E_n^2(\nu)}, \quad \forall \nu \in [0, \nu_0]. \quad (14.14)$$

То есть в процессе эволюции квантовая система будет оставаться в собственном состоянии гамильтониана $H(\nu)$, соответствующим уровню энергии $E_n(\nu)$, с точностью порядка ϵ^2 .

Доказательство. Будем решать задачу Коши (14.5) в базисе (14.6). Гамильтониан, который входит в уравнение Шредингера, в этом базисе диагонален с точностью до линейных членов по ϵ ,

$$H' = H_D - i\epsilon \frac{dS}{d\nu} S^{-1}.$$

Пусть матрица S выбрана таким образом, как описано в лемме 14.1.1. Предположим, что в начальный момент времени система находится в собственном состоянии гамильтониана H_D и, следовательно, в собственном состоянии исходного гамильтониана $H = S^{-1}H_DS$. Это значит, что начальное условие в базисе b_k имеет вид

$$\psi_{(n)}(0, \epsilon) = b_n(0) = \underbrace{(0, \dots, 0}_{n-1}, 1, 0, \dots, 0).$$

Любое решение уравнения Шредингера представимо в виде

$$\psi_{(n)}(\nu, \epsilon) = \phi_{(n)}(\nu, \epsilon) \exp \left(-\frac{i}{\epsilon} \int_0^\nu d\sigma H_D(\sigma) \right), \quad (14.15)$$

где $\phi_{(n)}$ – некоторый вектор гильбертова пространства $\mathbb{H}$. Тогда для вектора $\phi_{(n)}$ получаем уравнение

$$\frac{\partial \phi_{(n)}}{\partial \nu} = -\phi_{(n)} \exp \left(-\frac{i}{\epsilon} \int_0^\nu d\sigma H_D \right) \frac{dS}{d\nu} S^{-1} \exp \left(\frac{i}{\epsilon} \int_0^\nu d\sigma H_D \right),$$

Это уравнение вместе с начальным условием перепишем в виде интегрального уравнения

$$\phi_{(n)}(\nu, \epsilon) = \psi'_{(n)}(0) - \int_0^\nu d\sigma \phi_{(n)} \exp \left(-\frac{i}{\epsilon} \int_0^\sigma d\lambda H_D \right) \frac{dS}{d\sigma} S^{-1} \exp \left(\frac{i}{\epsilon} \int_0^\sigma d\lambda H_D \right). \quad (14.16)$$

При $\epsilon \rightarrow 0$ подынтегральное выражение содержит быстро осциллирующий множитель и его легко оценить. Рассмотрим модуль компоненты решения $\psi'_{(n)}^j$, которая соответствует собственному состоянию гамильтониана H с энергией E_j , где $E_j \neq E_n$,

$$\left| \psi'_{(n)}^j \right| = \left| \phi_{(n)}^j \right| = \left| \sum_{k=1}^N \int_0^\nu d\sigma \exp \left(\frac{i}{\epsilon} \int_0^\sigma d\lambda (E_j - E_k) \right) \phi_{(n)}^k \left(\frac{dS}{d\nu} S^{-1} \right)_k^j \right|. \quad (14.17)$$

В сумме справа слагаемые с $E_k = E_j$ вклада не дают в силу равенства (14.12). При $E_k \neq E_j$ каждое слагаемое проинтегрируем по частям:

$$\begin{aligned} & \frac{\epsilon}{i(E_j - E_k)} \exp \left(\frac{i}{\epsilon} \int_0^\sigma d\lambda (E_j - E_k) \right) \phi_{(n)}^k \left(\frac{dS}{d\nu} S^{-1} \right)_k^j \Big|_0^\nu - \\ & - \frac{\epsilon}{i} \int_0^\nu d\sigma \exp \left(\frac{i}{\epsilon} \int_0^\sigma d\lambda (E_j - E_k) \right) \frac{1}{E_j - E_k} \frac{d}{d\sigma} \left[\phi_{(n)}^k \left(\frac{dS}{d\nu} S^{-1} \right)_k^j \right]. \end{aligned} \quad (14.18)$$

По предположению подынтегральное выражение во втором слагаемом является дифференцируемой функцией и его снова можно проинтегрировать по частям. В результате получим, что оно имеет порядок ϵ^2 , и им можно пренебречь. Модуль первого слагаемого, очевидно, ограничен. Таким образом, получаем оценку

$$\left| \psi'_{(n)}^j(\nu, \epsilon) \right| = \frac{O(\epsilon)}{\min |E_j(\sigma) - E_k(\sigma)|}, \quad \forall j \notin \Upsilon_n, \quad (14.19)$$

где минимум берется по всем k , для которых $E_k \neq E_j$, и всем $\sigma \in [0, \nu]$.

Теперь снова вернемся к выражению (14.18). Из оценки (14.19) вытекает, что $|\phi_{(n)}^k|$ для всех k при $E_k \neq E_n$ имеет порядок не ниже ϵ . Поэтому в сумме (14.17) все слагаемые с индексом $k \notin \Upsilon_n$ дают вклад не ниже ϵ^2 , и ими можно пренебречь. Поэтому оценку (14.19) можно улучшить

$$\left| \psi'_{(n)}{}^j(\nu, \epsilon) \right| = \frac{O(\epsilon)}{\min_{\sigma} |E_j(\sigma) - E_n(\sigma)|}, \quad \forall j \notin \Upsilon_n,$$

где минимум берется только по $\sigma \in [0, \nu]$.

Норма произвольного решения сохраняется во времени и равна единице. Следовательно,

$$1 - \sum_{j \in \Upsilon_n} |\psi'_{(n)}{}^j(\nu, \epsilon)|^2 = \sum_{j \notin \Upsilon_n} |\psi'_{(n)}{}^j(\nu, \epsilon)|^2.$$

Поскольку число уровней конечно, то отсюда вытекает оценка (14.14). $\square$

Замечание. В теореме функция $\Delta E_n(\nu)$ для каждого ν равна константе и ее можно включить в $O(\epsilon^2)$. Тем не менее мы выделили множитель ΔE_n с тем, чтобы показать, что предположение о том, что уровни энергии не пересекаются, является существенным. При пересечении уровней энергии знаменатель в (14.14) обращается в нуль и доказательство не проходит. В этом случае требуются дополнительные предположения о степени касания уровней энергии и дополнительное исследование. В своей оригинальной статье [86] Борн и Фок рассмотрели случай, когда спектр гамильтониана дискретен, но может быть неограничен. Неявно ими было сделано предположение о невырожденности спектра почти для всех моментов времени. Кроме того, допускалась возможность определенного пересечения уровней энергии с течением времени. Мы рассмотрели более простой конечномерный случай, когда уровни энергии не пересекаются. Это позволило упростить доказательство и выявить наиболее существенные моменты. Оценка (14.14) согласуется с оценкой, приведенной в [86]. $\square$

Адиабатическая теорема утверждает, что, если в начальный момент времени система находилась в собственном состоянии гамильтониана, соответствующем уровню энергии $E_n(0)$, и этот уровень невырожден, то в адиабатическом пределе она будет оставаться в собственном состоянии $E_n(\nu)$ с точностью порядка ϵ^2 при конечных значениях параметра ν . Если уровень энергии E_n вырожден, то система будет находиться в одном из собственных состояний E_j , где $j \in \Upsilon_n$, с той же точностью. В следующем разделе мы увидим, что оценка (14.14) неулучшаема, а в процессе эволюции система может оказаться в любом из вырожденных состояний E_j , $j \in \Upsilon_n$, с вероятностью порядка единицы. Эти утверждения, естественно, не зависят от выбора базиса, который использовался при доказательстве адиабатической теоремы.

Рассмотрим теперь, как выглядит в адиабатическом пределе решение задачи Коши (14.4) в фиксированном базисе в невырожденном случае. Пусть $\varphi(\nu)$ – собственная функция гамильтониана $H(\nu)$, отвечающая невырожденному собственному значению энергии $E(\nu)$,

$$\varphi H = E \varphi, \quad \forall \nu \in [0, \nu_0].$$

Эти собственные функции определены с точностью до фазового множителя, который может зависеть от ν . Пусть в начальный момент времени система находилась в собственном состоянии $\psi_0 = \varphi(0)$. В адиабатическом приближении она будет находиться в собственном состоянии, соответствующем энергии $E(\nu)$. Поскольку собственное состояние невырождено, то решение задачи Коши (14.4) может отличаться от φ не

более, чем на фазовый множитель. Поэтому будем искать решение в виде $\psi = e^{i\Theta}\varphi$, где $\Theta(t)$ – неизвестная функция времени. Тогда из уравнения Шредингера следует уравнение на фазу

$$\dot{\Theta} = i(\dot{\varphi}, \varphi) - E(\epsilon t), \quad (14.20)$$

где точка обозначает дифференцирование по времени. Поскольку в начальный момент времени $\Theta(0) = 0$, то фаза имеет вид

$$\Theta(t) = i \int_0^t ds \left(\frac{d\varphi}{ds}, \varphi \right) - \int_0^t ds E(\epsilon s) = i \int_0^\nu d\sigma \left(\frac{d\varphi}{d\sigma}, \varphi \right) - \int_0^t ds E(\epsilon s), \quad (14.21)$$

где $\sigma := \epsilon s$.

Покажем, что фазу собственной функции φ всегда можно выбрать таким образом, что будет выполнено равенство

$$\left(\frac{d\varphi}{d\nu}, \varphi \right) = 0, \quad (14.22)$$

если $\nu \in [0, \infty)$. Действительно, пусть $\varphi = e^{i\beta}\chi$, где функция $\beta(\nu)$ удовлетворяет уравнению

$$i \frac{d\beta}{d\nu} = \left(\frac{d\varphi}{d\nu}, \varphi \right) \quad (14.23)$$

с некоторым начальным условием, например, $\beta(0) = 0$. Тогда нетрудно проверить, что для новых собственных функций выполнено равенство $(d\chi/d\nu, \chi) = 0$. Поскольку уравнение (14.23) всегда имеет решение на полупрямой, то собственные функции φ гамильтониана всегда можно выбрать таким образом, что будет выполнено равенство (14.22).

Однако уравнение (14.23) может не иметь решения на окружности $\mathbb{S}^1$. Будем считать, что на окружности $\nu \in [0, 2\pi]$. Тогда необходимым условием существования решения является равенство

$$i \int_0^{2\pi} d\nu \left(\frac{d\varphi}{d\nu}, \varphi \right) = 2\pi m, \quad m = 0, \pm 1, \pm 2, \dots$$

Ясно, что это условие в общем случае не выполняется. Поэтому уравнение (14.23) может не иметь решения на окружности. В этом случае первое слагаемое в выражении для фазы (14.21) устранить нельзя. По сути дела это и есть фаза Берри.

Решение задачи Коши на окружности $\nu \in \mathbb{S}^1$ означает наличие машины времени. Эти решения можно отбросить как нефизические. Однако Берри предложил другую схему рассуждений, которая будет рассмотрена в разделе 14.2.1.

14.1.1 Двухуровневая система

В настоящем разделе в качестве примера мы рассмотрим двухуровневую квантово-механическую систему, для которой уравнение Шредингера решается явно. Будет показано, что оценка, данная в адиабатической теореме, является неулучшаемой.

Чтобы упростить задачу, поступим следующим образом. Вместо того, чтобы задать исходный гамильтониан в фиксированном базисе, а затем его диагонализировать, мы зададим диагональную матрицу H_D и унитарную матрицу S , которые определяют исходный гамильтониан $H = S^{-1}H_DS$. Пусть диагонализированный гамильтониан имеет вид

$$H_D = \begin{pmatrix} E_1(\nu) & 0 \\ 0 & E_2(\nu) \end{pmatrix},$$

где $E_{1,2}(\nu)$ – некоторые заданные функции. Унитарную матрицу S в (14.10) выберем в виде

$$S = \begin{pmatrix} \cos \frac{\alpha}{2} & i \sin \frac{\alpha}{2} \\ i \sin \frac{\alpha}{2} & \cos \frac{\alpha}{2} \end{pmatrix},$$

где $\alpha(\nu) \in \mathbb{R}$ – также некоторая заданная функция. Следовательно, исходный гамильтониан задачи имеет вид

$$H = S^{-1} H_D S = \begin{pmatrix} E_1 \cos^2 \frac{\alpha}{2} + E_2 \sin^2 \frac{\alpha}{2} & -\frac{i}{2}(E_2 - E_1) \sin \alpha \\ \frac{i}{2}(E_2 - E_1) \sin \alpha & E_1 \sin^2 \frac{\alpha}{2} + E_2 \cos^2 \frac{\alpha}{2} \end{pmatrix}$$

и зависит от трех, пока произвольных, функций $E_1(\nu)$, $E_2(\nu)$ и $\alpha(\nu)$ параметра ν , которые предполагаются достаточно гладкими.

Будем решать уравнение Шредингера в базисе (14.6), в котором гамильтониан имеет вид (14.7). Простые вычисления приводят к гамильтониану

$$H' = \begin{pmatrix} E_1(\nu) & \frac{\dot{\alpha}}{2} \\ \frac{\dot{\alpha}}{2} & E_2(\nu) \end{pmatrix},$$

где точка обозначает дифференцирование по времени t . Ищем решение уравнения Шредингера (14.5) в виде строки

$$\psi' = \left(\exp \left(-i \int_0^t ds E_1 \right) \phi, \exp \left(-i \int_0^t ds E_2 \right) \chi \right),$$

где $\phi(t)$ и $\chi(t)$ – неизвестные функции. Подстановка этого выражения в уравнение Шредингера приводит к системе уравнений для компонент:

$$\begin{aligned} i\dot{\phi} &= \frac{\dot{\alpha}}{2} \exp \left(-i \int_0^t ds (E_2 - E_1) \right) \chi, \\ i\dot{\chi} &= \frac{\dot{\alpha}}{2} \exp \left(i \int_0^t ds (E_2 - E_1) \right) \phi. \end{aligned} \tag{14.24}$$

При $\dot{\alpha} \neq 0$ из первого уравнения следует равенство

$$\chi = \frac{2i}{\dot{\alpha}} \exp \left(i \int_0^t ds (E_2 - E_1) \right) \dot{\phi}. \tag{14.25}$$

Продифференцируем это равенство по времени и подставим во второе уравнение. В результате получим уравнение второго порядка для ϕ :

$$\ddot{\phi} + \left(i(E_2 - E_1) - \frac{\ddot{\alpha}}{\dot{\alpha}} \right) \dot{\phi} + \left(\frac{\dot{\alpha}}{2} \right)^2 \phi = 0. \tag{14.26}$$

Для того, чтобы решить это уравнение в явном виде зафиксируем произвольные функции, которые входят в задачу:

$$\begin{aligned} E_1 &= E_1^{(0)} + \epsilon t, & E_{1(0)} &= \text{const}, \\ E_2 &= E_2^{(0)} + \epsilon t, & E_{2(0)} &= \text{const}, \\ \alpha &= 2\epsilon t. \end{aligned} \tag{14.27}$$

Тогда уравнение (14.26) примет простой вид

$$\ddot{\phi} + 2i\Delta E \dot{\phi} + \epsilon^2 \phi = 0, \quad (14.28)$$

где $\Delta E := E_2^{(0)} - E_1^{(0)}$ – расстояние между уровнями энергии. Общее решение этого уравнения зависит от двух постоянных интегрирования $C_{1,2}$ и имеет вид

$$\phi = e^{-i\Delta E t} (C_1 e^{i\omega_\epsilon t} + C_2 e^{-i\omega_\epsilon t}),$$

где

$$\omega_\epsilon := \sqrt{\Delta E^2 + \epsilon^2}.$$

Компонента χ определяется по формуле (14.25).

Предположим, что в начальный момент времени система находилась в состоянии E_1 , т.е.

$$\phi(0) = 1, \quad \chi(0) = 0. \quad (14.29)$$

Простые вычисления дают решение задачи Коши для системы уравнений (14.24):

$$\begin{aligned} \phi &= e^{-i\Delta E t} \left[\cos(\omega_\epsilon t) + \frac{i\Delta E}{\omega_\epsilon} \sin(\omega_\epsilon t) \right], \\ \chi &= e^{i\Delta E t} \left[-\frac{i\epsilon}{\omega_\epsilon} \sin(\omega_\epsilon t) \right]. \end{aligned} \quad (14.30)$$

Выпишем также компоненты соответствующего вектора состояния

$$\begin{aligned} \psi'^1 &= e^{-i\left(\frac{\nu^2}{2\epsilon} + E_1^{(0)} \frac{\nu}{\epsilon} - \Delta E \frac{\nu}{\epsilon}\right)} \left[\cos \frac{\omega_\epsilon \nu}{\epsilon} + \frac{i\Delta E}{\omega_\epsilon} \sin \frac{\omega_\epsilon \nu}{\epsilon} \right], \\ \psi'^2 &= e^{-i\left(\frac{\nu^2}{2\epsilon} + E_2^{(0)} \frac{\nu}{\epsilon} + \Delta E \frac{\nu}{\epsilon}\right)} \left[-\frac{i\epsilon}{\omega_\epsilon} \sin \frac{\omega_\epsilon \nu}{\epsilon} \right], \end{aligned} \quad (14.31)$$

где мы перешли от t, ϵ к переменным ν, ϵ . Отсюда следует, что адиабатический предел для самого вектора состояния неопределен, т.к. его фаза стремится к бесконечности. Однако оценку квадрата модуля компонент можно дать. Для решения (14.31) следует оценка

$$1 - |\psi'^1(\nu, \epsilon)|^2 = \frac{O(\epsilon^2)}{(\Delta E)^2}, \quad |\psi'^2(\nu, \epsilon)|^2 = \frac{O(\epsilon^2)}{(\Delta E)^2},$$

которая совпадает с оценкой в адиабатической теореме. Отсюда следует, что данная оценка неумлучшаема.

Теперь рассмотрим случай вырожденных состояний, $E_1 = E_2$, при заданных ранее функциях (14.27). В этом случае уравнение (14.28) сводится к уравнению для свободного осциллятора:

$$\ddot{\phi} + \epsilon^2 \phi = 0,$$

и легко интегрируется. Выпишем соответствующее решение задачи Коши (14.29) для компонент вектора состояния:

$$\begin{aligned} \psi'^1 &= e^{-i\left(\frac{\nu^2}{2\epsilon} - E_1^{(0)} \frac{\nu}{\epsilon}\right)} \cos \nu, \\ \psi'^2 &= -i e^{-i\left(\frac{\nu^2}{2\epsilon} - E_1^{(0)} \frac{\nu}{\epsilon}\right)} \sin \nu. \end{aligned}$$

Мы снова видим, что адиабатический предел у вектора состояния отсутствует. Однако квадраты модулей компонент хорошо определены:

$$|\psi'^1|^2 = \cos^2 \nu, \quad |\psi'^2|^2 = \sin^2 \nu.$$

Отсюда следует, что по мере увеличения параметра ν вектор состояния ψ' осциллирует между вырожденными состояниями. Это показывает, что, если в начальный момент времени система находится в одном из вырожденных состояний, то в процессе эволюции она может оказаться в любом из вырожденных состояний с вероятностью порядка единицы.

14.2 Фаза Берри

Перейдем к задаче, которую рассмотрел М. Берри [83], в ее простейшем варианте.

Пусть гильбертово пространство конечномерно и гамильтониан $H = H(\lambda)$ достаточно гладко зависит от точки некоторого многообразия $\lambda \in \mathbb{M}$ размерности $\dim \mathbb{M} = n$. Если на $\mathbb{M}$ выбрать координатную окрестность $\mathbb{U} \subset \mathbb{M}$, то гамильтониан будет зависеть от n параметров λ^k , $k = 1, \dots, n$, (координат точки λ). Будем считать, что положение точки λ на $\mathbb{M}$ зависит от времени t некоторым наперед заданным образом, т.е. гамильтониан зависит от точки некоторой кривой $\gamma = \lambda(t)$, $t \in [0, t_0]$. Мы будем решать уравнение Шредингера в адиабатическом приближении, т.е. величина t_0 должна быть достаточно велика. Предположим также, что гамильтониан зависит от времени только через точку $\lambda(t) \in \mathbb{M}$ как сложная функция.

14.2.1 Абелев случай: невырожденное состояние

Рассмотрим задачу на собственные значения

$$\phi H = E \phi, \quad E = \text{const},$$

где $\phi \in \mathbb{H}$ при всех $\lambda \in \mathbb{M}$. Предположим, что существует невырожденное собственное значение энергии E и соответствующее собственное состояние ϕ , которые достаточно гладко зависят от $\lambda \in \mathbb{M}$. Не ограничивая общности, предположим, что собственная функция ϕ нормирована на единицу, $(\phi, \phi) = 1$. Тогда она единственна с точностью до умножения на фазовый множитель, который может зависеть от λ . Зафиксируем этот фазовый множитель каким-либо образом.

Теперь будем решать задачу Коши для уравнения Шредингера (14.1) с начальным условием

$$\psi|_{t=0} = \phi_0, \quad (14.32)$$

где $\phi_0 := \phi(\lambda(0))$. В адиабатическом приближении квантовая система в процессе эволюции будет оставаться в собственном состоянии, соответствующем уровню энергии $E(\lambda)$. Поэтому ищем решение в виде

$$\psi = e^{i\Theta} \phi,$$

где $\Theta = \Theta(t)$ – некоторая неизвестная функция от времени. Тогда из уравнения Шредингера следует уравнение на фазу (14.20) с начальным условием $\Theta|_{t=0} = 0$. Поскольку $\dot{\phi} = \dot{\lambda}^k \partial_k \phi$, то решение задачи Коши для уравнения (14.20) имеет вид

$$\Theta = \int_0^t dt \dot{\lambda}^k A_k - \int_0^t ds E(\lambda(s)) = \int_{\lambda(0)}^{\lambda(t)} d\lambda^k A_k - \int_0^t ds E(\lambda(s)), \quad (14.33)$$

где введено обозначение

$$A_k(\lambda) := i(\phi, \partial_k \phi) \quad (14.34)$$

и интеграл по λ берется вдоль кривой $\lambda(t)$.

Таким образом, интеграл (14.33) в адиабатическом приближении дает решение задачи Коши для уравнения Шредингера (14.1) с начальным условием (14.32). Первое слагаемое в (14.33) называется *геометрической фазой* или *фазой Берри*, а второе – *динамической фазой*.

Заметим, что компоненты (14.34) вещественны вследствие нормировки волновой функции. Действительно, дифференцируя условие нормировки $(\phi, \phi) = 1$, получаем равенство

$$(\partial_k \phi, \phi) + (\phi, \partial_k \phi) = (\phi, \partial_k \phi)^\dagger + (\phi, \partial_k \phi) = 0.$$

Отсюда вытекает вещественность компонент (14.34) и, следовательно, фазы Берри.

Теперь рассмотрим множество замкнутых кривых $\gamma = \lambda(t) \in \Omega(\mathbb{M}, \lambda_0)$ на многообразии параметров $\mathbb{M}$ с началом и концом в некоторой фиксированной точке $\lambda_0 \in \mathbb{M}$. Тогда для полного изменения фазы волновой функции получаем ответ

$$\Theta = \Theta_{\text{в}} - \int_0^{t_0} dt E(\lambda(t)),$$

где

$$\Theta_{\text{в}} := \oint_{\gamma} d\lambda^k A_k. \quad (14.35)$$

Динамическая часть фазы волновой функции расходитсся при $t_0 \rightarrow \infty$. Однако в экспериментах наблюдается разность фаз двух векторов состояний с одинаковой динамической фазой, которая определяется фазой Берри. Поэтому рассмотрим фазу Берри подробнее.

Заметим, что выражение для фазы Берри не зависит от параметризации кривой γ . Это значит, что переход к адиабатическому пределу в уравнении Шредингера влияет на фазу Берри только через компоненты локальной формы связности (14.34).

В таком виде можно дать геометрическую интерпретацию фазе Берри $\Theta_{\text{в}}$, которая определяется первым слагаемым в полученном выражении (14.33). А именно, мы имеем главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(1))$, базой которого является многообразие параметров $\lambda \in \mathbb{M}$, а структурной группой – группа $\mathbb{U}(1)$ (фаза вектора состояния $e^{i\Theta}$). Вектор гильбертова пространства $\phi \in \mathbb{H}$ представляет собой локальное сечение ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{H}, \mathbb{U}(1), \mathbb{P})$, типичным слоем которого является гильбертово пространство $\mathbb{H}$.

Рассмотрим изменение локального сечения ассоциированного расслоения, которое вызвано умножением вектора гильбертова пространства на фазовый множитель (вертикальный автоморфизм),

$$\phi' = e^{ia} \phi,$$

где $a = a(\lambda) \in \mathcal{C}^2(\mathbb{M})$ – произвольная дважды дифференцируемая функция. Тогда компоненты (14.34) преобразуются по правилу

$$A'_k = A_k - \partial_k a.$$

Сравнивая это правило с преобразованием компонент локальной формы связности, заключаем, что поля $A_k(\lambda)$ можно интерпретировать, как компоненты локальной формы связности для группы $\mathbb{U}(1)$. Другими словами, $A_k(\lambda)$ – это калибровочное поле для одномерной унитарной группы $\mathbb{U}(1)$. Если база ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{H}, \mathbb{U}(1), \mathbb{P})$ покрыта некоторым семейством карт, $\mathbb{M} = \cup_j \mathbb{U}_j$, то множество сечений, заданных в каждой координатной окрестности $\mathbb{U}_j$, определяет семейство

локальных форм связности на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(1))$. Семейство локальных форм связности $d\lambda^k A_k$ определяет единственную с точностью до изоморфизма связность на $\mathbb{P}$ (теорема 13.1.1).

Вспомним выражение для элемента группы голономии в виде упорядоченной P -экспоненты (13.63). В рассматриваемом случае группа $\mathbb{U}(1)$ абелева и P -экспонента совпадает с обычной экспонентой. Поэтому фаза Берри (14.35) определяет элемент $e^{i\Theta_B}$ группы голономии $\Phi(\lambda_0, e) \subset \mathbb{U}(1)$ главного расслоения в точке $(\lambda_0, e) \in \mathbb{P}$, соответствующей нулевому сечению $\mathbb{M} \ni \lambda \mapsto (\lambda, e) \in \mathbb{P}$, где $\lambda_0 := \lambda(0)$ и e – единица структурной группы $\mathbb{U}(1)$. Сечение является нулевым, поскольку в начальный момент времени фаза Берри равна нулю, $\Theta_B|_{t=0} = 0$. Локальная форма связности $d\lambda^k A_k$ также соответствует нулевому сечению.

Если база $\mathbb{M}$ односвязна, то выражение для фазы Берри (14.35) можно переписать в виде поверхностного интеграла от компонент локальной формы кривизны. Используя формулу Стокса, получаем следующее выражение

$$\Theta_B = \frac{1}{2} \iint_S d\lambda^k \wedge d\lambda^l F_{kl}, \quad (14.36)$$

где S – поверхность в $\mathbb{M}$ с границей $\gamma \in \Omega(\mathbb{M}, \lambda_0)$ и $F_{kl} = \partial_k A_l - \partial_l A_k$ – компоненты локальной формы кривизны (напряженности калибровочного поля). Если база $\mathbb{M}$ не является односвязной, то выражение для фазы Берри в виде поверхностного интеграла (14.36) имеет место только для тех замкнутых путей, которые стягиваются в точку.

14.2.2 Частица со спином $1/2$ в магнитном поле

В качестве примера вычислим фазу Берри для частицы со спином $1/2$, находящейся во внешнем однородном магнитном поле. В нерелятивистской квантовой механике частица со спином $1/2$ описывается двухкомпонентной волновой функцией

$$\psi = (\psi_+, \psi_-).$$

Будем считать, что она находится в евклидовом пространстве $\mathbb{R}^3$ с заданным однородным магнитным полем. Пусть напряженность магнитного поля $H^k(t)$, $k = 1, 2, 3$, не зависит от точки пространства, но меняется со временем t некоторым заданным образом. Кроме того, для простоты, пренебрежем кинетической энергией частицы и будем считать, что другие поля отсутствуют. В этом случае гильбертово пространство $\mathbb{H}$ двумерно, и гамильтониан частицы состоит из одного слагаемого – взаимодействия магнитного момента частицы с внешним магнитным полем (см., например, [91, 92]),

$$H = -\mu H^k \sigma_k^T,$$

где σ_k^T – транспонированные матрицы Паули (поскольку в наших обозначениях они действуют справа) и μ – магнетон (размерная постоянная, равная отношению магнитного момента частицы к ее спину). Чтобы привести гамильтониан к виду, который был рассмотрен ранее, введем новые переменные $\lambda^k := -\mu H^k$. Тогда гамильтониан примет вид

$$H = \lambda^k \sigma_k^T = \begin{pmatrix} \lambda^3 & \lambda^+ \\ \lambda^- & -\lambda^3 \end{pmatrix}, \quad (14.37)$$

где $\lambda^\pm := \lambda^1 \pm i\lambda^2$.

Собственные значения гамильтониана (14.37) находятся из уравнения

$$\det(H - E\mathbb{1}) = 0,$$

которое имеет два вещественных решения

$$E_{\pm} = \pm|\lambda|, \quad (14.38)$$

где

$$|\lambda| := \sqrt{(\lambda^1)^2 + (\lambda^2)^2 + (\lambda^3)^2}$$

– длина вектора $\lambda = \{\lambda^k\} \in \mathbb{R}^3$. Нетрудно проверить, что уравнение на собственные функции,

$$\phi_{\pm} H = E_{\pm} \phi_{\pm},$$

имеет два решения

$$\phi_{\pm} = \frac{1}{\sqrt{2|\lambda|}} \left(\pm \frac{\lambda^-}{\sqrt{|\lambda| \mp \lambda^3}}, \sqrt{|\lambda| \mp \lambda^3} \right). \quad (14.39)$$

В полученном выражении множитель выбран таким образом, что собственные функции нормированы на единицу,

$$(\phi_{\pm}, \phi_{\pm}) = 1.$$

Таким образом, гамильтониан (14.37) частицы со спином $1/2$, находящейся во внешнем однородном магнитном поле, имеет два невырожденных собственных состояния (14.39), соответствующих уровням энергии (14.38).

Для дальнейших вычислений в пространстве параметров $\lambda \in \mathbb{R}^3$ удобно ввести сферические координаты $|\lambda|, \theta, \varphi$:

$$\begin{aligned} \lambda^1 &= |\lambda| \sin \theta \cos \varphi, \\ \lambda^2 &= |\lambda| \sin \theta \sin \varphi, \\ \lambda^3 &= |\lambda| \cos \theta. \end{aligned}$$

Тогда собственные функции примут вид

$$\phi_+ = \left(\cos \frac{\theta}{2} e^{-i\varphi}, \sin \frac{\theta}{2} \right), \quad \phi_- = \left(-\sin \frac{\theta}{2} e^{-i\varphi}, \cos \frac{\theta}{2} \right).$$

Допустим, что экспериментатор, наблюдающий за частицей, достаточно гладко меняет однородное магнитное со временем. То есть параметры $\lambda^k(t)$, от которых зависит гамильтониан, достаточно гладко зависят от времени. Предположим также, что в начальный момент времени $t = 0$ частица находилась в состоянии ϕ_+ . Соответствующее решение уравнения Шредингера (14.1) в адиабатическом приближении имеет вид

$$\psi = e^{i\Theta} \phi_+,$$

где фаза Θ удовлетворяет уравнению (14.20). Нетрудно вычислить компоненты локальной формы связности $A_k = i(\phi_+, \partial_k \phi_+)$ для собственного состояния ϕ_+ :

$$A_{|\lambda|} = 0, \quad A_{\theta} = 0, \quad A_{\varphi} = \cos^2 \frac{\theta}{2}.$$

Соответствующая локальная форма кривизны имеет только две отличные от нуля компоненты:

$$F_{\theta\varphi} = -F_{\varphi\theta} = -\frac{1}{2} \sin \theta.$$

Теперь вычислим фазу Берри для замкнутой кривой в пространстве параметров $\gamma = \{\lambda^k(t)\} \in \mathbb{M}$,

$$\begin{aligned} \Theta_{\text{в}} &= \oint_{\gamma} d\lambda^k A_k = \frac{1}{2} \iint_S d\lambda^k \wedge d\lambda^l F_{kl} = \\ &= \iint_S d\theta \wedge d\varphi F_{\theta\varphi} = -\frac{1}{2} \iint_S d\theta \wedge d\varphi \sin \theta = -\frac{1}{2} \Omega(\gamma), \end{aligned} \quad (14.40)$$

где S – поверхность в $\mathbb{R}^3$ с границей γ и $\Omega(\gamma)$ – телесный угол, который занимает контур γ , если смотреть из начала координат.

Если в начальный момент времени частица находилась в состоянии ϕ_- , то вычисления проводятся аналогично. В этом случае

$$A_{|\lambda|} = 0, \quad A_{\theta} = 0, \quad A_{\varphi} = \sin^2 \frac{\theta}{2},$$

и компоненты локальной формы кривизны отличаются знаком:

$$F_{\theta\varphi} = -F_{\varphi\theta} = \frac{1}{2} \sin \theta.$$

Следовательно, фаза Берри также отличается только знаком.

Таким образом, если в начальный момент времени частица находилась в одном из состояний $\phi_{\pm}$, то после изменения однородного магнитного поля вдоль замкнутой кривой γ , ее волновая функция изменится на фазовый множитель, геометрическая часть которого равна

$$\Theta_{\text{в}\pm} = \mp \frac{1}{2} \Omega(\gamma), \quad (14.41)$$

где $\Omega(\gamma)$ – телесный угол, под которым виден замкнутый контур γ из начала координат. Этот результат имеет место в адиабатическом приближении, когда параметры $\lambda(t)$ медленно меняются со временем.

Выражение для фазы Берри (14.41) было подтверждено экспериментально [93] для рассеяния поляризованных нейтронов в спиральном магнитном поле.

В рассмотренном примере однородное магнитное поле может иметь произвольное направление и величину. Следовательно, база $\mathbb{M}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(1))$ совпадает с евклидовым пространством, $\mathbb{M} = \mathbb{R}^3$. Поэтому, согласно теореме 12.1.1, главное расслоение $\mathbb{P}$ тривиально, $\mathbb{P} \approx \mathbb{R}^3 \times \mathbb{U}(1)$. В случае фазы Берри связность на этом расслоении определяется сечением ассоциированного расслоения, например, ϕ_+ , которое находится путем решения уравнения Шредингера. Это сечение (14.39), как нетрудно проверить, имеет особенность на положительной полуоси $\lambda^3 \geq 0$. Компоненты локальной формы связности относительно декартовой системы координат имеют вид

$$\begin{aligned} A_1 &= \frac{\partial \varphi}{\partial \lambda^1} A_{\varphi} = -\frac{\sin \varphi \cos \frac{\theta}{2}}{2|\lambda| \sin \frac{\theta}{2}}, \\ A_2 &= \frac{\partial \varphi}{\partial \lambda^2} A_{\varphi} = \frac{\cos \varphi \cos \frac{\theta}{2}}{2|\lambda| \sin \frac{\theta}{2}}, \\ A_3 &= \frac{\partial \varphi}{\partial \lambda^3} A_{\varphi} = 0. \end{aligned} \quad (14.42)$$

Здесь мы вынуждены перейти в декартову систему координат, поскольку сферическая система координат сингулярна на оси λ^3 и непригодна для исследования особенностей, которые здесь расположены. Как видим, компоненты локальной формы связности имеют особенность на положительной полуоси $\lambda^3 \geq 0$ как и вектор ϕ_+ . Теперь вычислим компоненты локальной формы тензора кривизны. У нее отличны от нуля все компоненты:

$$\begin{aligned} F_{12} &= -F_{21} = -\frac{\cos \theta}{2|\lambda|^2}, \\ F_{13} &= -F_{31} = \frac{\sin \theta \sin \varphi}{2|\lambda|^2}, \\ F_{23} &= -F_{32} = -\frac{\sin \theta \cos \varphi}{2|\lambda|^2}. \end{aligned}$$

Наконец, вычислим квадрат тензора кривизны, который является инвариантом,

$$F^2 = 2(F_{12}^2 + F_{13}^2 + F_{23}^2) = \frac{1}{2|\lambda|^4}.$$

Таким образом, форма кривизны сингулярна только в начале координат.

Вернемся к нашему главному расслоению $\mathbb{R}^3 \times \mathbb{U}(1)$. Локальная форма связности (14.42) на нем неопределена, т.к. имеет особенность на полуоси $\lambda^3 \geq 0$, которую мы обозначим $\{\lambda_+^3\}$. Поэтому, чтобы построить главное расслоение с заданной связностью, мы вынуждены удалить прообраз $\pi^{-1}(\{\lambda_+^3\})$, где $\pi : \mathbb{R}^3 \times \mathbb{U}(1) \rightarrow \mathbb{R}^3$ – естественная проекция. В результате получаем тривиальное главное расслоение $(\mathbb{R}^3 \setminus \{\lambda_+^3\}) \times \mathbb{U}(1)$, которое является подрасслоением исходного расслоения. На этом главном расслоении локальная форма (14.42) бесконечно дифференцируема.

Можно рассуждать по-другому. Поскольку магнитное поле является внешним, то мы вправе предположить, что оно меняется, например, в полупространстве $\mathbb{R}_+^3$, определяемом условием $\lambda_1 > 0$. Поскольку полупространство $\mathbb{R}_+^3$ диффеоморфно всему евклидову пространству $\mathbb{R}^3$, то соответствующее главное расслоение тривиально: $\mathbb{P} \approx \mathbb{R}_+^3 \times \mathbb{U}(1)$. В этом случае никаких вопросов в определении связности вообще не возникает, т.к. локальная форма связности (14.42) гладкая. При этом выражение для фазы Берри (14.41) останется прежним.

Таким образом, фаза Берри является не топологическим понятием, а геометрическим, т.к. топология главного расслоения тривиальна. Она обязана своим происхождением нетривиальной связности, которая определяется сечениями ассоциированного расслоения.

14.2.3 Неабелев случай: вырожденное состояние

Понятие фазы Берри было обобщено на случай, когда уровни энергии гамильтониана $H(\lambda)$ вырождены [85]. В этом случае при решении уравнения Шредингера возникает главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(r))$ со структурной группой $\mathbb{U}(r)$, где r – количество независимых собственных функций, соответствующих вырожденному уровню энергии E . Опишем эту конструкцию подробно.

Предположим, что гамильтониан квантовой системы зависит от точки некоторого многообразия $\lambda(t) \in \mathbb{M}$, как и ранее. Пусть E – вырожденное собственное значение гамильтониана H , которому соответствуют r независимых собственных функций $\phi^a \in \mathbb{H}$, $a = 1, \dots, r$,

$$\phi^a H = E \phi^a$$

для всех моментов времени. Мы предполагаем, что $E(\lambda)$ и $\phi^a(\lambda)$ являются достаточно гладкими функциями от точки многообразия λ , и число собственных функций r не меняется со временем.

Собственные функции можно выбрать ортонормированными,

$$(\phi^a, \phi_b) = \delta_b^a,$$

где δ_b^a – символ Кронекера. Будем искать решение задачи Коши ψ^a для уравнения Шредингера (14.1) с начальным условием

$$\psi^a(0) = \psi_0^a := \phi^a(\lambda(0)).$$

То есть в начальный момент времени система находится в одном из собственных состояний ϕ^a . В адиабатическом приближении решение ψ^a для всех моментов времени является собственной функцией гамильтониана $H(\lambda)$, соответствующей значению энергии $E(\lambda)$. Поэтому его можно разложить по собственным функциям вырожденного состояния

$$\psi^a = \phi^b U_b^{-1a}, \quad (14.43)$$

где $U(\lambda) \in \mathbb{U}(r)$ – некоторая унитарная матрица, которая достаточно гладко зависит от точки $\lambda \in \mathbb{M}$.

Унитарность матрицы U обусловлена следующим обстоятельством. Рассмотрим решения ψ^a для всех значений индекса $a = 1, \dots, r$. Дифференцируя скалярное произведение (ψ^a, ψ_b) по времени и используя уравнение Шредингера, получаем уравнение

$$\frac{\partial}{\partial t}(\psi^a, \psi_b) = -i(\psi^a H, \psi_b) + i(\psi^a, H \psi_b) = 0.$$

Последнее равенство следует из самосопряженности гамильтониана. Отсюда следует, что, если в начальный момент времени векторы $\psi_0^a := \phi^a(\lambda(0))$ ортонормированы, то соответствующие решения уравнения Шредингера останутся таковыми и во все последующие моменты времени. Поэтому матрица U в разложении (14.43) унитарна.

Для искомого решения (14.43) уравнение Шредингера сводится к уравнению

$$i\phi^c \dot{U}_c^{-1b} + i\dot{\phi}^c U_c^{-1b} = \phi^c H U_c^{-1b}.$$

Возьмем скалярное произведение левой и правой части с ϕ_a . В результате получаем уравнение на унитарную матрицу

$$\dot{U}_a^{-1b} = \dot{\lambda}^k A_{ka}{}^c U_c^{-1b} - iE U_a^{-1b}, \quad (14.44)$$

где введено обозначение

$$A_{ka}{}^c := -(\partial_k \phi^c, \phi_a). \quad (14.45)$$

Из условия ортонормированности собственных функций ϕ^a следует антиэрмитовость компонент $A_{ka}{}^b$ для всех $k = 1, \dots, n$, если индексы a, b рассматриваются, как матричные. Действительно, дифференцируя условие ортонормированности $(\phi^b, \phi_a) = \delta_a^b$, получаем равенство

$$(\partial_k \phi^b, \phi_a) + (\phi^b, \partial_k \phi_a) = (\phi^a, \partial_k \phi_b)^\dagger + (\phi^b, \partial_k \phi_a) = 0.$$

То есть матрицы A_k антиэрмитовы и поэтому принадлежат алгебре Ли $\mathfrak{u}(r)$. Следовательно, матрицы A_k определяют 1-формы в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ со значениями в алгебре Ли, как и компоненты локальной формы связности.

Начальное условие для унитарной матрицы имеет вид

$$U_a^{-1b}|_{t=0} = \delta_a^b.$$

Решение задачи Коши для уравнения (14.44) можно записать в виде упорядоченного Р-произведения (см. раздел 13.5)

$$\begin{aligned} U^{-1}(t) &= \text{P exp} \left(\int_0^t ds \dot{\lambda}^k(s) A_k(s) - i \int_0^t ds E(\lambda(s)) \right) = \\ &= \text{P exp} \left(\int_{\lambda(0)}^{\lambda(t)} d\lambda^k A_k \right) \times \exp \left(-i \int_0^t ds E(\lambda(s)) \right), \end{aligned} \quad (14.46)$$

где мы, для краткости, опустили матричные индексы.

Первый сомножитель является обобщением фазы Берри на случай вырожденных состояний, а второй сомножитель – это динамическая фаза. Динамическая фаза имеет тот же вид, что и в случае невырожденного состояния.

Первый сомножитель в решении (14.46) представляет собой унитарную матрицу Вилчека–Зи

$$U_{\text{wz}}^{-1} := \text{P exp} \left(\int_{\lambda(0)}^{\lambda(t)} d\lambda^k A_k \right), \quad (14.47)$$

которой можно дать следующую геометрическую интерпретацию. Мы имеем главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(r))$ со структурной группой $\mathbb{U}(r)$ (преобразование (14.43)). Набор собственных функций ϕ^a представляет собой сечение ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{H}^r, \mathbb{U}(r), \mathbb{P})$, типичным слоем которого является тензорное произведение гильбертовых пространств

$$\mathbb{H}^r := \underbrace{\mathbb{H} \otimes \dots \otimes \mathbb{H}}_r.$$

При вертикальном автоморфизме, который задан унитарной матрицей $U(\lambda) \in \mathbb{U}(r)$,

$$\phi'^a = \phi^b U_b^{-1a}, \quad \phi'_a = U_a^b \phi_b,$$

поля (14.45) преобразуются по правилу

$$A'_k = U A_k U^{-1} + \partial_k U U^{-1}, \quad (14.48)$$

где мы опустили матричные индексы. Отсюда следует, что поля A_k можно интерпретировать, как компоненты локальной формы связности или поля Янга–Миллса. Совокупность этих компонент, заданная на координатном покрытии базы $\mathbb{M}$, однозначно задает связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(r))$.

Если путь замкнут, $\gamma \in \Omega(\mathbb{M}, \lambda_0)$, то унитарная матрица Вилчека–Зи (14.47) представляет собой элемент группы голономии $U_{\text{wz}}^{-1} \in \Phi(\lambda_0, e)$, в точке $(\lambda_0, e) \in \mathbb{P}$, соответствующей нулевому сечению $\mathbb{M} \ni \lambda \mapsto (\lambda, e) \in \mathbb{P}$, где $\lambda_0 := \lambda(0)$ и e – единица структурной группы $\mathbb{U}(r)$.

Таким образом, в случае вырожденного уровня энергии гамильтониана возникает главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(r))$. В рассматриваемом случае базой $\mathbb{M}$ является многообразие параметров $\lambda \in \mathbb{M}$, от точки которого зависит гамильтониан. Мы предполагаем, что это многообразие конечномерно. Структурной группой является унитарная группа $\mathbb{U}(r)$, которая также конечномерна. Связность на главном расслоении

определяется сечениями ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{H}^r, \mathbb{U}(r), \mathbb{P})$. В общем случае типичным слоем ассоциированного расслоения может быть бесконечномерное гильбертово пространство $\mathbb{H}^r$. В настоящей монографии мы не рассматриваем бесконечномерных многообразий и расслоений, чтобы избежать возникающих при этом трудностей [8]. Однако в данном случае все, что нужно, это формула преобразования компонент локальной формы связности (14.48), которую легко проверить в каждом конкретном случае. Если база $\mathbb{M}$ не покрывается одной картой, то состояние квантовой системы задается семейством локальных сечений на координатном покрытии базы. Оно определяет семейство локальных форм связности (14.45). В свою очередь семейство локальных форм связности однозначно с точностью до изоморфизма определяет связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(r))$.

Опять мы видим, что главные и ассоциированные расслоения могут быть тривиальными или нет, это зависит от рассматриваемой задачи. Связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(r))$ может быть нетривиальна и приводит к нетривиальной матрице Вилчека–Зи (14.47), которая описывает параллельный перенос слоев вдоль пути на базе $\lambda(t) \in \mathbb{M}$, даже для тривиальных расслоений. Это говорит о ее геометрическом, а не топологическом происхождении. Для замкнутых путей $\gamma \in \Omega(\mathbb{M}, \lambda_0)$ с началом и концом в точке $\lambda_0 \in \mathbb{M}$ матрица Вилчека–Зи определяет элемент группы голономии $U_{wz} \in \Phi(\lambda_0, e) \subset \mathbb{U}(r)$.

При рассмотрении фазы Берри и матрицы Вилчека–Зи мы, для простоты, предположили, что гильбертово пространство квантовой системы конечномерно. Это предположение можно существенно ослабить. Полученные формулы справедливы для всех уровней, для которых справедлива адиабатическая теорема. То есть это должен быть изолированный уровень, энергия которого отделена от остального спектра.

14.3 Эффект Ааронова–Бома

Другой пример возникновения нетривиальной связности на тривиальном главном расслоении $\mathbb{P}(\mathbb{R}^4, \pi, \mathbb{U}(1)) \approx \mathbb{R}^4 \times \mathbb{U}(1)$ в нерелятивистской квантовой механике дает эффект Ааронова–Бома [84]. В этом случае, в отличие от фазы Берри, в качестве базы $\mathbb{M}$ главного расслоения выступает не пространство параметров, а пространство-время $\mathbb{R}^4$, в котором частица движется. Эффект Ааронова–Бома не связан с адиабатической теоремой, и обсуждается только с геометрической точки зрения.

Рассмотрим уравнение Шредингера (14.1), в котором гамильтониан описывает движение точечной частицы массы m в трехмерном евклидовом пространстве $\mathbb{R}^3$ с декартовыми координатами x^μ , $\mu = 1, 2, 3$,

$$H_0 = -\frac{\eta^{\mu\nu} p_\mu p_\nu}{2m} + U = -\frac{\hbar^2}{2m} \Delta + U, \quad (14.49)$$

где $p_\mu := i\hbar\partial_\mu$ – оператор импульса частицы, $\eta_{\mu\nu} := \text{diag}(-, -, -)$ – отрицательно определенная пространственная метрика, $\Delta := \partial_1^2 + \partial_2^2 + \partial_3^2$ – оператор Лапласа и $U(x)$ – потенциальная энергия частицы.

Четырехмерный оператор импульса имеет вид $p_\alpha = i\hbar\partial_\alpha$, $\alpha = 0, 1, 2, 3$. При этом нулевая компонента 4-импульса $p_0 = i\hbar\partial_0 = i\hbar\partial_t$ имеет физический смысл оператора энергии частицы.

Если частица взаимодействует с электромагнитным полем, то это взаимодействие описывается с помощью минимальной подстановки для всех четырех компонент импульса

$$p_\alpha \mapsto i\hbar\partial_\alpha - \frac{e}{c}A_\alpha, \quad (14.50)$$

где e – заряд частицы, c – скорость света и A_α – потенциал электромагнитного поля (компоненты локальной формы $\mathbb{U}(1)$ связности). При этом нулевая компонента, разделенная на скорость света, A_0/c , имеет физический смысл потенциала электрического поля, а пространственные компоненты A_μ – ковекторного потенциала магнитного поля. Таким образом, точечная частица, находящаяся в электромагнитном поле, описывается уравнением Шредингера

$$i\hbar \frac{\partial \psi}{\partial t} = \left[\frac{\hbar^2}{2m} \eta^{\mu\nu} \left(\partial_\mu + i \frac{e}{\hbar c} A_\mu \right) \left(\partial_\nu + i \frac{e}{\hbar c} A_\nu \right) + \frac{e}{c} A_0 \right] \psi + U\psi. \quad (14.51)$$

Здесь мы вернулись к стандартным обозначениям квантовой механики, когда гамильтониан действует на вектор гильбертова пространства слева, поскольку он содержит производные.

В дальнейшем, для простоты, положим $\hbar = 1$ и $c = 1$.

С геометрической точки зрения минимальная подстановка (14.50) с точностью до постоянных совпадает с заменой частной производной на ковариантную:

$$\partial_\alpha \mapsto \nabla_\alpha := \partial_\alpha + ieA_\alpha.$$

Посмотрим на уравнение Шредингера (14.51) с геометрической точки зрения. Оно решается во всем пространстве, $\psi = \psi(t, x)$, поэтому базой расслоения является четырехмерное евклидово пространство, $(t, x) \in \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4$. При этом $\mathbb{R}^4$ рассматривается просто как четырехмерное многообразие без какой либо четырехмерной метрики. При желании метрику можно ввести, однако ее наличие никак не влияет на структуру главного расслоения и связности. Метрика $\eta_{\mu\nu}$ определена только на пространственных сечениях $t = \text{const}$, поскольку она входит в уравнение Шредингера. Волновая функция $\psi(t, x)$ представляет собой сечение ассоциированного расслоения $\mathbb{E}(\mathbb{R}^4, \pi_{\mathbb{E}}, \mathbb{C}, \mathbb{U}(1), \mathbb{P})$, типичным слоем которого является комплексная плоскость $\mathbb{C}$ и которое ассоциировано с некоторым главным расслоением $\mathbb{P}(\mathbb{R}^4, \pi, \mathbb{U}(1))$. Это главное расслоение всегда тривиально, $\mathbb{P} \simeq \mathbb{R}^4 \times \mathbb{U}(1)$, т.к. базой является четырехмерное евклидово пространство. На этом главном расслоении задана локальная форма $\mathbb{U}(1)$ связности, которая определяется электромагнитным потенциалом $A_\alpha(t, x)$. В нерелятивистской квантовой механике рассматривается не все множество сечений ассоциированного расслоения, а лишь подмножество, состоящее из тех дифференцируемых функций $\psi(t, x)$, которые в каждый момент времени t принадлежат гильбертову пространству квадратично интегрируемых функций $\mathbb{H} = \mathcal{L}_2(\mathbb{R}^3)$ на пространственных сечениях $\mathbb{R}^3$.

Рассмотрим два случая.

14.3.1 Электрический потенциал

Предположим, что магнитный потенциал равен нулю, $A_\mu = 0$, $\mu = 1, 2, 3$. Запишем уравнение Шредингера в виде

$$i\dot{\psi} = (H_0 + eA_0)\psi, \quad (14.52)$$

где H_0 – гамильтониан системы в отсутствии электромагнитного поля (14.49). Предположим также, что электрический потенциал зависит только от времени, $A_0 = A_0(t)$. Будем искать решение уравнения Шредингера (14.52) в виде $\psi = e^{-i\Theta} \phi$, где ϕ – решение свободного уравнения Шредингера,

$$i\dot{\phi} = H_0 \phi, \quad (14.53)$$

и $\Theta = \Theta(t)$ – некоторая фаза, не зависящая от точки пространства. Подстановка $\psi = e^{-i\Theta}\phi$ в исходное уравнение Шредингера (14.52) приводит к уравнению на фазу

$$\dot{\Theta} = eA_0,$$

где мы сократили общий фазовый множитель $e^{-i\Theta}$ и ϕ . Это можно сделать, т.к. уравнение Шредингера должно выполняться для всех t и x . Решение полученного уравнения имеет вид

$$\Theta(t) = \Theta_0 + e \int_0^t ds A_0(s), \quad (14.54)$$

где Θ_0 – значение фазы волновой функции в начальный момент времени.

Ааронов и Бом предложили следующий эксперимент, схема которого показана на рис.14.1. Пучок электронов делится на два пучка, которые пропускаются через две металлические трубки, на которые подается различный потенциал. Затем пучки собираются, и на экране наблюдается интерференционная картина. Электрический потенциал, который подается на трубки, зависит от времени. Предполагается, что он равен нулю, пока оба пучка не окажутся в своих трубках. Затем он возрастает до некоторых значений, которые отличаются внутри каждой трубки, и снова падает до нуля перед выходом пучков из трубок. Таким образом, пучки находятся в поле A_0 только внутри трубок. Интерференционная картина зависит от разности фаз электронов в пучках, которую можно приближенно оценить следующим образом.

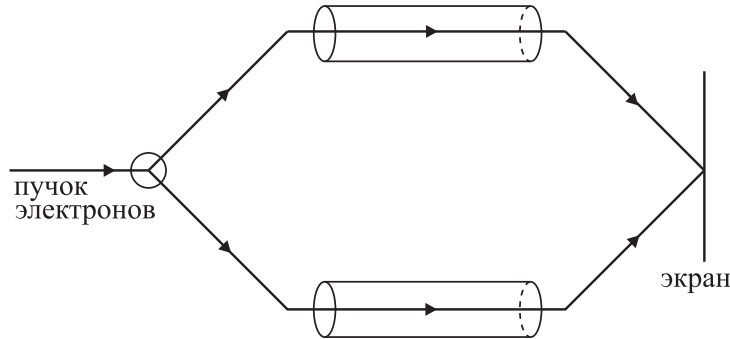


Рис. 14.1: Пучок электронов делится на два пучка, которые пропускаются через две металлические трубки, имеющие разные потенциалы. Затем пучки вновь собираются и наблюдается интерференционная картина, которая зависит от разности фаз электронов в разных пучках.

Предположим, что электрон описывается волновой функцией $\psi(t, x)$, которая удовлетворяет уравнению Шредингера (14.52) во всем пространстве-времени $\mathbb{R}^4$. Мы считаем, что в каждый момент времени носитель волновой функции отличен от нуля в небольшой окрестности пространства вблизи траектории частицы. Только в этом случае вообще можно говорить о траектории частицы. В частности, при прохождении электрона сквозь металлическую трубку предполагается, что носитель волновой функции целиком лежит внутри трубки. Математически это можно описать, выбрав в уравнении Шредингера (14.52) соответствующий потенциал. Этот гипотетический потенциал не меняет пространство-время, т.е. базу главного расслоения, а только обеспечивает движение электронов по заданной траектории. Изменение фазы электрона в верхнем пучке оценим следующим образом. Поскольку потенциал электрического поля однороден внутри трубки и носитель волновой функции целиком

содержится внутри трубки, то можно считать, что фаза волновой функции определяется интегралом (14.54). Обозначим моменты времени, соответствующие расщеплению пучка и достижению экрана, соответственно, t_1 и t_2 . Тогда фаза волновой функции электрона в верхнем пучке при достижении экрана изменится на величину, задаваемую интегралом

$$\Theta_1 = e \int_{t_1}^{t_2} dt A_0^{(1)}(t),$$

где $A_0^{(1)}(t)$ – потенциал электрического поля в момент времени t , т.е. в той точке пространства, где в момент времени t находится электрон из верхнего пучка. Аналогично, изменение фазы волновой функции электрона из нижнего пучка равно

$$\Theta_2 = e \int_{t_1}^{t_2} dt A_0^{(2)}(t),$$

где $A_0^{(2)}$ – потенциал электрического поля вдоль нижней траектории. Ясно, что разность фаз электронов в верхнем и нижнем пучке, $\Theta_{AB} = \Theta_2 - \Theta_1$, можно записать в виде интеграла

$$\Theta_{AB} = e \oint_{\gamma} dt A_0(t). \quad (14.55)$$

вдоль замкнутого контура γ в пространстве-времени, когда сначала проходится нижняя половина контура, изображенного на рис.14.1, а затем – верхняя половина в обратную сторону. На рис.14.1 показана проекция контура γ на пространственную плоскость.

Вернемся к геометрической интерпретации. Разность фаз электронов дается интегралом (14.55), который однозначно определяется контуром γ и заданным на нем потенциалом A_0 . Электрический потенциал A_0 представляет собой временную компоненту локальной формы $\mathbb{U}(1)$ связности на главном расслоении $\mathbb{P}(\mathbb{R}^4, \pi, \mathbb{U}(1))$. Поэтому разность фаз (14.55) определяет элемент группы голономии $e^{i\Theta_{AB}} \in \Phi((t_0, x_0), e) \subset \mathbb{U}(1)$ в точке $(t_0 = 0, x_0)$, где x_0 – точка пространства, в которой пучок расщепляется, а e – единица группы.

В заключение данного раздела рассмотрим преобразование компоненты локальной формы $\mathbb{U}(1)$ связности при изменении сечения. Из уравнения Шредингера (14.52) следует, что при вертикальном автоморфизме

$$\psi' = e^{ia} \psi,$$

где $a = a(t)$ – дифференцируемая функция времени, компонента локальной формы $\mathbb{U}(1)$ связности преобразуется по правилу

$$eA'_0 = eA_0 + \dot{a},$$

как и подобает компонентам локальной формы $\mathbb{U}(1)$ связности.

Таким образом, в основе эффекта Ааронова–Бома, так же как и для фазы Берри, лежит не топология, а нетривиальная геометрия, т.е. связность с нетривиальной группой голономии. При этом топология пространства может быть как тривиальной, так и нетривиальной.

14.3.2 Магнитный потенциал

Рассмотрим теперь случай, когда потенциал электрического поля равен нулю, $A_0 = 0$. Предположим, что ковекторный потенциал магнитного поля зависит только от пространственных координат x^μ и не зависит от времени t (статическое поле). Тогда уравнение Шредингера примет вид

$$\begin{aligned} i\hbar \nabla \psi &= \frac{1}{2m} \eta^{\mu\nu} (\partial_\mu + ieA_\mu)(\partial_\nu + ieA_\nu) \psi + U\psi = \\ &= \frac{1}{2m} \eta^{\mu\nu} (\partial_\mu^2 \psi + 2ieA_\mu \partial_\nu \psi + ie\partial_\mu A_\nu \psi - e^2 A_\mu A_\nu \psi) + U\psi. \end{aligned} \quad (14.56)$$

Пусть ϕ – решение уравнения Шредингера в отсутствии потенциала магнитного поля (14.53). Тогда нетрудно проверить, что функция

$$\psi = e^{-i\Theta} \phi,$$

где фаза Θ удовлетворяет уравнению

$$\partial_\mu \Theta = eA_\mu \quad (14.57)$$

является решением исходного уравнения Шредингера (14.56).

Ааронов и Бом предложили эксперимент для определения фазы Θ , схема которого показана на рис.14.2. В этом эксперименте пучок электронов делится на два пучка, которые огибают бесконечно длинный соленоид с постоянным магнитным потоком Φ , который перпендикулярен плоскости рисунка, с разных сторон. Затем пучки собираются вместе и на экране наблюдается интерференционная картина, которая зависит от разности фаз электронов в разных пучках.

Для оценки разности фаз электронов сделаем те же предположения, что и в случае электрического поля. А именно, будем считать, что уравнение Шредингера без магнитного потенциала имеет решение с носителем, который сосредоточен в малой окрестности траектории электрона. Мы предполагаем, что это можно осуществить путем введения в уравнение (14.53) соответствующего потенциала. Этот потенциал не меняет топологию пространства-времени, а только обеспечивает движение электронов вдоль заданной траектории. Тогда для фазы решения уравнения Шредингера с магнитным потенциалом справедливы уравнения (14.57). Поскольку магнитное поле вне соленоида равно нулю, $\partial_\mu A_\nu - \partial_\nu A_\mu = 0$, то условия интегрируемости для системы уравнений (14.57) выполнены. Поэтому разность фаз можно представить в виде контурного интеграла

$$\Theta_{AB} = e \oint_\gamma dx^\mu A_\mu, \quad (14.58)$$

где γ – замкнутый контур в четырехмерном пространстве времени, который охватывает соленоид. Отметим, что слагаемое $dx^0 A_0$ в подынтегральном выражении равно нулю, т.к. $A_0 = 0$ по предположению. Этот интеграл не зависит от выбора контура, охватывающего соленоид, поскольку магнитное поле вне соленоида равно нулю, $\partial_\mu A_\nu - \partial_\nu A_\mu = 0$.

Фазу Ааронова–Бома, используя формулу Стокса, можно записать в виде поверхностного интеграла

$$\Theta_{AB} = \frac{1}{2} e \iint_S dx^\mu \wedge dx^\nu F_{\mu\nu} = e\Phi, \quad (14.59)$$

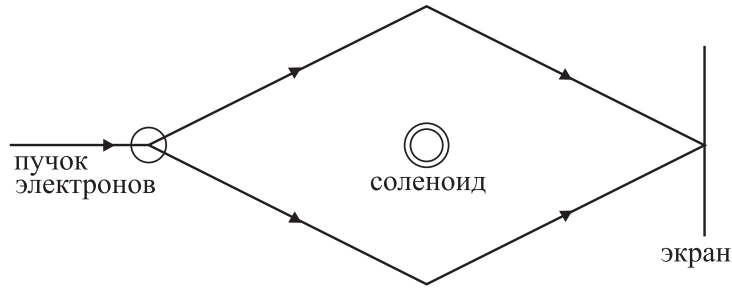


Рис. 14.2: Пучок электронов делится на два пучка, которые огибают тонкий соленоид с разных сторон. Затем пучки вновь собираются и наблюдается интерференционная картина, которая зависит от разности фаз электронов в разных пучках.

где $F_{\mu\nu}$ – напряженность магнитного поля (компоненты локальной 2-формы кривизны) и Φ – полный поток магнитного поля через соленоид. Заметим, что для применения формулы Стокса, необходимо считать, что магнитный потенциал определен в пространстве $\mathbb{R}^3$ всюду, включая сам соленоид.

Геометрическая интерпретация рассмотренного эффекта Ааронова–Бома состоит в следующем. Мы имеем то же самое главное расслоение, что и в случае электрического потенциала, $\mathbb{P}(\mathbb{R}^4, \pi, \mathcal{U}(1))$, базой которого является четырехмерное евклидово пространство, $(t, x) \in \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4$, в котором движутся пучки электронов, а структурной группой – унитарная группа $\mathcal{U}(1)$ (фазовый множитель $e^{i\theta}$ волновой функции). Однако связность на нем другая: отличны от нуля только пространственные компоненты локальной формы связности A_μ , $\mu = 1, 2, 3$. Разность фаз Ааронова–Бома (14.58) однозначно определяется контуром γ (тем же, что и в случае электрического потенциала) и значениями компонент связности A_μ на нем. При записи контурного интеграла в виде поверхностного (14.59) предполагается, что связность задана на всем пространстве-времени $\mathbb{R}^4$. Тем самым, мы рассматриваем соленоид конечного радиуса, чтобы избежать сингулярностей.

Таким образом, главное расслоение тривиально, а фаза Ааронова–Бома Θ_{AB} , зависящая от связности и контура, однозначно определяет элемент группы голономии.

Волновая функция электрона, как и в случае электрического потенциала, является сечением ассоциированного расслоения $\mathbb{E}(\mathbb{R}^4, \pi_E, \mathbb{C}, \mathcal{U}(1), \mathbb{P})$, базой которого является евклидово пространство $\mathbb{R}^4$, а структурной группой – унитарная группа $\mathcal{U}(1)$. При вертикальном автоморфизме

$$\psi' = e^{ia}\psi,$$

где $a = a(x)$ – дифференцируемая функция пространственных координат x^μ , $\mu = 1, 2, 3$, потенциал магнитного поля преобразуется по правилу

$$eA'_\mu = eA_\mu - \partial_\mu a.$$

Это следует из уравнения Шредингера (14.56). Таким образом, компоненты потенциала магнитного поля действительно ведут себя, как компоненты локальной формы связности.

Поскольку разность фаз электронов (14.58) определяется значениями компонент локальной формы связности только вблизи контура интегрирования, то базу тривиального главного расслоения $\mathbb{P}(\mathbb{R}^4, \pi, \mathcal{U}(1))$ можно сузить, не меняя ответа. Например, можно удалить область пространства-времени, лежащую внутри контура γ и

содержащую соленоид. Тогда база расслоения перестанет быть односвязной. По этой причине эффект Ааронова–Бома часто называют топологическим. Как было показано выше, это совершенно необязательно. Достаточно считать, что магнитное поле отлично от нуля в ограниченной области на плоскости рис.14.2 внутри контура интегрирования. Если считать, что базой расслоения является евклидово пространство $\mathbb{R}^4$ с выколотым соленоидом, то формулу Стокса применить нельзя, т.к. она применима только для стягиваемых контуров. Таким образом, эффект Ааронова–Бома, вызванный магнитным потенциалом, так же как и фазу Берри, следует рассматривать как геометрический, а не топологический.

Эффект Ааронова–Бома как с электрическим, так и с магнитным потенциалом привлек большое внимание физиков по следующей причине. Согласно современным представлениям в калибровочных моделях наблюдаемыми величинами являются только калибровочно инвариантные функции. С этой точки зрения потенциал электромагнитного поля A_α , $\alpha = 0, 1, 2, 3$, сам по себе ненаблюдаем, т.к. не является калибровочно инвариантным. В рассмотренных примерах электрического и магнитного полей пучки электронов не подвергаются действию электромагнитных сил, поскольку напряженности электрического и магнитного поля в областях, через которые пролетают электроны, равны нулю. Поэтому, казалось бы, разность фаз в пучках электронов должна быть равна нулю. Однако из уравнения Шредингера следует, что это не так. Следует отметить, что наблюдаемым является не сам потенциал электромагнитного поля, а интеграл от него по замкнутому контуру, который определяет элемент группы голономии $\mathbb{U}(1)$ связности, который является инвариантным объектом.

Вскоре после публикации статьи, эффект Ааронова–Бома был подтвержден экспериментально. Эффект, вызванный магнитным потенциалом наблюдался в экспериментах [94, 95, 96].

Глава 15

Векторные поля Киллинга

В разделе 9.3 были рассмотрены геометрические структуры на многообразии, которые инвариантны относительно действия некоторой группы преобразований. Вопрос ставился так. Пусть задана группа преобразований $(\mathbb{M}, \mathbb{G})$, и требуется найти такие структуры, например, тензорные поля, которые инвариантны относительно этих преобразований. Обратная задача нахождения группы преобразований, которую допускает заданная геометрическая структура на многообразии $\mathbb{M}$ также очень важна.

Изучение преобразований, которые сохраняют метрику пространства-времени играет исключительно важную роль в математической физике. Достаточно сказать, что с такими преобразованиями связаны наиболее важные законы сохранения. В настоящей главе мы рассмотрим (псевдо)риманово многообразие $(\mathbb{M}, g)$ и найдем условия, при которых метрика инвариантна относительно действия группы преобразований $(\mathbb{M}, \mathbb{G})$. Дадим определение векторных полей Киллинга, которые являются генераторами локальных симметрий метрики, а также изучим некоторые из их свойств. Будет доказана теорема о том, что, если (псевдо)риманово многообразие обладает максимально возможной группой симметрии, то это – пространство постоянной кривизны.

15.1 Изометрии и инфинитезимальные изометрии

Рассмотрим n -мерное (псевдо)риманово многообразие $(\mathbb{M}, g)$ с метрикой $g(x) = g_{\alpha\beta}(x)dx^\alpha \otimes dx^\beta$, $\alpha, \beta = 0, 1, \dots, n-1$, и соответствующей связностью Леви-Чивиты Γ .

Определение. Диффеоморфизм

$$\iota : \mathbb{M} \ni x \mapsto x' = \iota(x) \in \mathbb{M}$$

называется *изометрией* или *движением* многообразия $\mathbb{M}$, если он сохраняет метрику,

$$g(x) = \iota^* g(x'), \quad (15.1)$$

где ι^* – возврат отображения ι . □

В настоящей главе, для простоты, мы не будем использовать знак тильды для обозначения компонент связности Леви-Чивиты, т.к. аффинная связность общего вида с кручением и неметричностью использоваться не будет.

Из условия инвариантности метрики (15.1) следует инвариантность скалярного произведения векторов. Пусть $X, Y \in \mathbb{T}_x(\mathbb{M})$ – два произвольных вектора из касательного пространства в точке $x \in \mathbb{M}$. Тогда справедливы равенства:

$$g(X, Y)|_x = (i^*g)(X, Y)|_x = g(i_*X, i_*Y)|_{i(x)},$$

которое эквивалентно определению (15.1). Первое равенство следует из определения (15.1), а второе вытекает из определения возврата отображения (2.78).

Поскольку изометрия сохраняет метрику, то она сохраняет также связность Леви-Чивиты, соответствующий тензор кривизны, экстремали и, вообще, все геометрические объекты, которые определяются только метрикой.

Запишем отображение (15.1) в координатах. Пусть обе точки x и x' принадлежат одной координатной окрестности и имеют, соответственно, координаты x^α и x'^α . Тогда изометрия i в координатах запишется в виде условия

$$g_{\alpha\beta}(x) = \frac{\partial x'^\gamma}{\partial x^\alpha} \frac{\partial x'^\delta}{\partial x^\beta} g_{\gamma\delta}(x'), \quad (15.2)$$

связывающего компоненты метрики в различных точках многообразия. Это условие по виду совпадает с правилом преобразования компонент метрики при преобразовании координат (4.4). Разница заключается в следующем. При преобразовании координат мы считаем, что одной и той же точке $x \in \mathbb{M}$ соответствует два набора координат $\{x^\alpha\}$ и $\{x'^\alpha := x'^\alpha\}$ в двух различных системах координат. При рассмотрении изометрий x и x' – это две различные точки одного и того же многообразия $\mathbb{M}$, и равенство (15.2) связывает значения компонент метрики в этих точках.

Предложение 15.1.1. *Множество всех изометрий данного (псевдо)риманова многообразия $(\mathbb{M}, g)$ является группой, которую обозначим $\mathbb{I}(\mathbb{M}) \ni i$.*

Доказательство. Две последовательных изометрии также являются изометрией. Произведение (последовательное действие изометрий) ассоциативно. Тожественное отображение многообразия $\mathbb{M}$ является изометрией и представляет собой единицу группы. У каждого диффеоморфизма i есть обратной диффеоморфизм i^{-1} , который является обратной изометрией. $\square$

Если метрика на многообразии задана, т.е. определены значения ее компонент во всех точках x , то соотношение (15.2) представляет собой уравнение на функции $x'(x)$, которые определяют изометрию. В общем случае это уравнение не имеет решений и у соответствующего (псевдо)риманова многообразия нет никаких нетривиальных изометрий. В этом случае группа изометрий состоит из одного единичного элемента. Чем шире группа изометрий, тем уже класс соответствующих (псевдо)римановых многообразий.

Пример 15.1.1. Евклидово пространство $\mathbb{R}^n$ с евклидовой метрикой $\delta_{\alpha\beta}$ допускает группу изометрий, которая состоит из преобразований неоднородной группы вращений $\mathbb{IO}(n, \mathbb{R})$, $\dim \mathbb{IO}(n, \mathbb{R}) = \frac{1}{2}n(n+1)$, состоящей из вращений, сдвигов и отражений. $\square$

Группа изометрий $\mathbb{I}(\mathbb{M})$ может быть дискретной или группой Ли.

Определение. Если группа изометрий $\mathbb{I}(\mathbb{M})$ является группой Ли, то имеет смысл говорить об инфинитезимальных преобразованиях (см. раздел 9.2). В этом случае мы говорим об *инфинитезимальных изометриях*

$$x^\alpha \mapsto x'^\alpha = x^\alpha + \epsilon K^\alpha + o(\epsilon), \quad \epsilon \ll 1.$$

Каждая инфинитезимальная изометрия генерируется некоторым достаточно гладким векторным полем $K(x) = K^\alpha(x)\partial_\alpha$, которое называется векторным *полем Киллинга*. $\square$

Замечание. Дискретные изометрии (псевдо)риманова многообразия, например, отражения, не генерируются никакими векторными полями. $\square$

Запишем условие инвариантности метрики относительно инфинитезимальных преобразований из группы изометрий в координатах. В разделе 2.6.5 было показано, что каждое векторное поле генерирует однопараметрическую группу преобразований, которая называется экспоненциальным отображением. Формально условие инвариантности метрики записывается в виде равенства нулю производной Ли вдоль векторного поля Киллинга $K = K^\alpha\partial_\alpha$ от метрики

$$\mathcal{L}_K g = 0. \quad (15.3)$$

Используя явное выражение для производной Ли (2.122), это уравнение в локальной системе координат принимает вид [97]

$$\nabla_\alpha K_\beta + \nabla_\beta K_\alpha = 0, \quad (15.4)$$

где $K_\alpha := K^\beta g_{\beta\alpha}$ – компоненты 1-формы Киллинга, а ковариантная производная

$$\nabla_\alpha K_\beta = \partial_\alpha K_\beta - \Gamma_{\alpha\beta}^\gamma K_\gamma$$

строится по символам Кристоффеля $\Gamma_{\alpha\beta}^\gamma$ (связность Леви–Чивиты).

Определение. Уравнение (15.4) называется *уравнением Киллинга*, а интегральные кривые полей Киллинга называются *траекториями Киллинга*. Если $K = K^\alpha\partial_\alpha$ – векторное поле Киллинга, то ему взаимно однозначно ставится в соответствие 1-форма $K = dx^\alpha K_\alpha$, где $K_\alpha := K^\beta g_{\beta\alpha}$, которая называется *формой Киллинга*, и для которой мы сохранили то же обозначение. $\square$

На любом (псевдо)римановом многообразии $(\mathbb{M}, g)$ уравнения Киллинга (15.3) всегда имеют тривиальное решение $K = 0$. Если уравнения Киллинга имеют только тривиальное решение, то в этом случае нетривиальные непрерывные изометрии отсутствуют.

Траектории Киллинга $\{x^\alpha(t)\} \in \mathbb{M}$, где $t \in \mathbb{R}$, определяются системой обыкновенных дифференциальных уравнений

$$\dot{x}^\alpha = K^\alpha. \quad (15.5)$$

Если траектория Киллинга при $t = 0$ проходит через точку $p = \{p^\alpha\} \in \mathbb{M}$, то при малых t она имеет вид

$$x^\alpha(t) = p^\alpha + tK^\alpha(p) + o(t). \quad (15.6)$$

Если в некоторой точке векторное поле Киллинга равно нулю, то эта точка остается неподвижной, т.е. является стационарной точкой группы изометрий. Поскольку изометрии определены для всего многообразия $\mathbb{M}$ и образуют группу, то векторные поля Киллинга обязаны быть полными, т.е. параметр t должен меняться на всей вещественной прямой $\mathbb{R}$.

Если для (псевдо)риманова многообразия $(\mathbb{M}, g)$ известно векторное поле Киллинга, то оно определяет не только инфинитезимальные изометрии, но и всю однопараметрическую подгруппу изометрий $\mathbb{I}(\mathbb{M})$. Для этого нужно найти интегральные кривые $x(t)$, проходящие, через все точки многообразия $p \in \mathbb{M}$. Если $x(0) = p$, то каждому значению $t \in \mathbb{R}$ соответствует диффеоморфизм

$$\iota : \mathbb{M} \ni p \mapsto x(t) \in \mathbb{M}.$$

Уравнения для векторных полей Киллинга в ковариантной форме (15.4) можно переписать в частных производных,

$$\partial_\alpha K_\beta + \partial_\beta K_\alpha - 2\Gamma_{\alpha\beta}^\gamma K_\gamma = 0.$$

В моделях математической физики часто ставится задача нахождения векторов Киллинга для заданной метрики на многообразии. Для решения этой задачи бывает удобнее использовать контравариантные компоненты векторов Киллинга, для которых уравнение Киллинга принимает вид

$$g_{\alpha\gamma} \partial_\beta K^\gamma + g_{\beta\gamma} \partial_\alpha K^\gamma + K^\gamma \partial_\gamma g_{\alpha\beta} = 0. \quad (15.7)$$

Уравнения Киллинга (15.3), которые в компонентах имеют вид (15.7), линейны и по векторам Киллинга, и по метрике. Отсюда сразу следует, что две метрики, которые отличаются постоянным множителем, имеют один и тот же набор векторов Киллинга. Кроме того, векторное поле Киллинга определено с точностью до умножения на произвольную постоянную, отличную от нуля. В частности, если K – векторное поле Киллинга, то и $-K$ также является полем Киллинга. Если независимых векторных полей Киллинга для заданной метрики несколько, то любая линейная комбинация этих полей также является полем Киллинга. То есть множество всех векторных полей Киллинга образует линейное пространство над полем вещественных чисел, которое является подпространством в множестве векторных полей $\mathcal{X}(\mathbb{M})$. В этом векторном пространстве можно ввести билинейную операцию. Простые вычисления показывают, что коммутатор двух векторных полей Киллинга K_1 и K_2 снова дает поле Киллинга:

$$\mathbb{L}_{[K_1, K_2]} g = \mathbb{L}_{K_1} \circ \mathbb{L}_{K_2} g - \mathbb{L}_{K_2} \circ \mathbb{L}_{K_1} g = 0,$$

Отсюда следует, что векторные поля Киллинга образуют алгебру Ли $\mathfrak{i}(\mathbb{M})$ над полем вещественных чисел, которая является подалгеброй алгебры Ли множества всех векторных полей, $\mathfrak{i}(\mathbb{M}) \subset \mathcal{X}(\mathbb{M})$. Эта алгебра является алгеброй Ли группы Ли изометрий $\mathbb{I}(\mathbb{M})$. В дальнейшем мы увидим, что алгебра Ли векторных полей Киллинга конечномерна.

Замечание. Векторные поля Киллинга не выдерживают умножение на функции. Поэтому они не образуют $\mathcal{C}^\infty(\mathbb{M})$ -модуль в отличие от множества всех векторных полей $\mathcal{X}(\mathbb{M})$. $\square$

Уравнения Киллинга (15.7) представляют собой систему линейных уравнений в частных производных первого порядка на компоненты векторных полей Киллинга. Эта система переопределена: мы имеем $n(n+1)/2$ уравнений на n неизвестных компонент $K^\alpha(x)$ (или $K_\alpha(x)$). Ниже мы увидим, что общее решение уравнений Киллинга при фиксированной метрике не содержит функционального произвола, но может

зависеть от нескольких параметров, число которых совпадает с числом линейно независимых решений. Максимальное число независимых параметров в общем решении $n(n+1)/2$ достигается на пространствах постоянной кривизны (теорема 15.3.2).

В дальнейшем нам понадобится также следующее наблюдение. Допустим, что метрика $g_{\alpha\beta}(x, t)$ зависит от некоторого параметра $t \in \mathbb{R}$, и для каждого значения t уравнения Киллинга выполнены при фиксированном векторном поле Киллинга (которое не зависит от t). Тогда разность метрик для различных значений параметра, $g_{\alpha\beta}(x, t_2) - g_{\alpha\beta}(x, t_1)$, также будет удовлетворять уравнениям Киллинга. Отсюда следует, что производная $\partial_t g_{\alpha\beta}$ является инвариантным тензором второго ранга.

С каждым полем Киллинга как и с произвольным полным векторным полем связана однопараметрическая группа преобразований, которая в данном случае сохраняет метрику.

Предложение 15.1.2. Пусть (псевдо)риманово многообразие (M, g) имеет $N \leq \dim M$ отличных от нуля коммутирующих между собой и линейно независимых векторных полей Киллинга K_i , $i = 1, \dots, N$. Тогда существует такая система координат, в которой все компоненты метрики не зависят от N координат, соответствующих траекториям Киллинга. Обратно. Если в некоторой системе координат компоненты метрики не зависят от N координат, то метрика g допускает локально по крайней мере N коммутирующих между собой ненулевых векторных полей Киллинга.

Доказательство. Мы дадим доказательство предложения только для несветоподобных векторных полей. В разделе 2.6.5 была построена специальная система координат, связанная с произвольным векторным полем, отличным от нуля. Применительно к коммутирующим векторным полям Киллинга K_i это означает, что существует такая система координат $(x^1, \dots, x^n)$, в которой каждое поле Киллинга имеет только одну постоянную компоненту, $K_i = \partial_i$. В этой системе координат уравнение (15.7) для каждого поля Киллинга принимает особенно простой вид

$$\partial_i g_{\alpha\beta} = 0, \quad i = 1, \dots, N. \quad (15.8)$$

Это значит, что все компоненты метрики не зависят от координат x^i . В этой системе координат траектории Киллинга определяются уравнениями

$$\dot{x}^i = 1, \quad \dot{x}^\mu = 0, \quad \mu \neq i.$$

Отсюда следует, что координатные линии x^i являются траекториями Киллинга.

Обратно. Если метрика не зависит от N координат, то выполнены уравнения (15.8). Эти уравнения совпадают с уравнениями Киллинга для векторных полей $K_i := \partial_i$, которые коммутируют. $\square$

Согласно сформулированной теореме, в предельном случае, когда количество коммутирующих полей Киллинга равно размерности многообразия, $N = n$, существует такая система координат, в которой все компоненты инвариантной метрики постоянны.

Пример 15.1.2. В евклидовом пространстве $\mathbb{R}^n$ в декартовой системе координат x^α , $\alpha = 1, \dots, n$, компоненты метрики постоянны, $g_{\alpha\beta} = \delta_{\alpha\beta}$. Эта метрика допускает n коммутирующих между собой векторных полей Киллинга $K_\alpha := \partial_\alpha$, которые соответствуют трансляциям. Все координатные оси являются траекториями Киллинга. $\square$

Если риманово многообразие (M, g) имеет два или более некоммутирующих векторных полей Киллинга, то это отнюдь не означает, что существует такая система координат, в которой компоненты метрики не зависят от двух или более координат.

Пример 15.1.3. Рассмотрим двумерную сферу $S^2 \hookrightarrow \mathbb{R}^3$. Пусть метрика g на сфере индуцирована обычным вложением. Риманово пространство (S^2, g) имеет три векторных поля Киллинга, соответствующих $SO(3)$ вращениям евклидова пространства $\mathbb{R}^3$. Легко понять, что на сфере не существует локальной системы координат, в которой компоненты метрики не зависели бы от двух координат. Действительно, это означает, что в данной системе координат компоненты метрики постоянны, и, следовательно, кривизна равна нулю. Но это невозможно, поскольку кривизна сферы постоянна и отлична от нуля. $\square$

В общей теории относительности мы предполагаем, что пространство-время является псевдоримановым многообразием с метрикой лоренцевой сигнатуры. Используя понятие векторного поля Киллинга, можно дать инвариантное

Определение. Пространство-время или его область называются *стационарным*, если на них существует времениподобное векторное поле Киллинга. $\square$

В статическом пространстве-времени интегральную кривую времениподобного векторного поля можно выбрать в качестве временной координаты. Тогда согласно предложению 15.1.2 в этой системе координат компоненты метрики не будут зависеть от времени, что оправдывает название “стационарное”.

Векторные поля Киллинга определены глобально и удовлетворяют уравнениям Киллинга на всем M . В то же время уравнения Киллинга – это локальный объект в том смысле, что они определены в каждой окрестности и могут иметь нетривиальные решения только на некотором подмногообразии $U \subset M$.

Пример 15.1.4. Рассмотрим гладкую замкнутую двумерную поверхность M , вложенную в трехмерное евклидово пространство $\mathbb{R}^3$, как показано на рис. 15.1. Отличительной особенностью этой поверхности является то, что ее нижняя часть является плоской. Пусть метрика на M индуцирована вложением $M \hookrightarrow \mathbb{R}^3$. Тогда уравнения Киллинга в нижней части поверхности легко интегрируются, как и на евклидовой плоскости. Однако найденные нетривиальные решения не будут в общем случае нетривиальными во всех точках M . Действительно, верхняя часть поверхности может быть искривлена так, что уравнения Киллинга на ней имеют только тривиальное решение. Следовательно, векторные поля Киллинга могут быть нетривиальны только на части многообразия M . Заметим, что в рассматриваемом примере траектории Киллинга не являются полными. Действительно, поля Киллинга ∂_x и ∂_y , соответствующие трансляциям, имеют единичную длину в той части поверхности, где она касается плоскости x, y , а вне этой части равны нулю. Следовательно, они разрывны, и траектории Киллинга имеют конечную длину. $\square$

В разделе ?? рассмотрены конформные векторные поля Киллинга, действие которых приводит к умножению метрики на некоторый множитель.

15.2 Свойства векторных полей Киллинга

Векторные поля Киллинга обладают рядом замечательных свойств. Начнем с простейших.

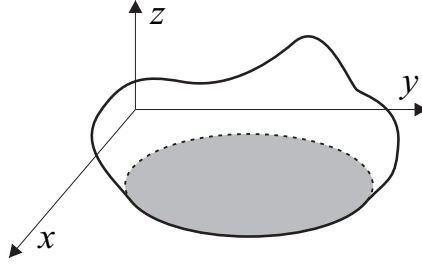


Рис. 15.1: Двумерная поверхность, вложенная в трехмерное евклидово пространство. Нижняя часть поверхности является плоской.

Предложение 15.2.1. *Длина вектора Киллинга остается постоянной вдоль траектории Киллинга:*

$$\mathcal{L}_K K^2 = \nabla_K K^2 = K^\alpha \partial_\alpha K^2 = 0. \quad (15.9)$$

Доказательство. Свернем уравнения Киллинга (15.4) с $K^\alpha K^\beta$:

$$2K^\alpha K^\beta \nabla_\alpha K_\beta = K^\alpha \nabla_\alpha K^2 = K^\alpha \partial_\alpha K^2 = 0. \quad \square$$

Следствие. Если векторные поля Киллинга существуют на лоренцевом многообразии, то они имеют определенную ориентацию: времениподобную, светоподобную или пространственноподобную. $\square$

Метрика g на многообразии $\mathbb{M}$ определяет два типа выделенных кривых: экстремали (или геодезические, если рассматривается связность Леви–Чивиты) и траектории Киллинга, если они существуют. Сравним траектории Киллинга с экстремалими [98].

Предложение 15.2.2. *Пусть $(\mathbb{M}, g)$ – (псевдо)риманово многообразие с векторным полем Киллинга K . Траектории Киллинга являются экстремалими тогда и только тогда, когда их длина постоянна на $\mathbb{M}$, $K^2 = \text{const}$ для всех $x \in \mathbb{M}$.*

Доказательство. Рассмотрим траектории Киллинга $x^\alpha(t)$, которые определяются системой уравнений

$$\dot{x}^\alpha = K^\alpha. \quad (15.10)$$

Длина дуги траектории Киллинга

$$ds^2 = g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta dt^2 = K^2 dt^2$$

постоянна вдоль траектории, т.е. параметр t пропорционален длине и, следовательно, является каноническим. Дифференцируя уравнение (15.10) по каноническому параметру t , получим равенство

$$\ddot{x}^\alpha = \partial_\beta K^\alpha \dot{x}^\beta = (\nabla_\beta K^\alpha - \Gamma_{\beta\gamma}^\alpha K^\gamma) \dot{x}^\beta,$$

которое перепишем в виде

$$\ddot{x}^\alpha = K^\beta \nabla_\beta K^\alpha - \Gamma_{\beta\gamma}^\alpha \dot{x}^\beta \dot{x}^\gamma. \quad (15.11)$$

Уравнения Киллинга позволяют переписать первое слагаемое в правой части в виде

$$K^\beta \nabla_\beta K^\alpha = -\frac{1}{2} g^{\alpha\beta} \partial_\beta K^2.$$

Тогда уравнения (15.11) примут вид

$$\ddot{x}^\alpha = -\frac{1}{2}g^{\alpha\beta}\partial_\beta K^2 - \Gamma_{\beta\gamma}^\alpha \dot{x}^\beta \dot{x}^\gamma.$$

Это уравнение совпадает с уравнением для экстремалей (16.21) тогда и только тогда, когда $K^2 = \text{const}$. $\square$

Доказанное утверждение показывает, что далеко не каждая траектория Киллинга является экстремалью.

Пример 15.2.1. Рассмотрим евклидову плоскость $\mathbb{R}^2$ с евклидовой метрикой. Эта метрика инвариантна относительно трехпараметрической неоднородной группы вращений $\mathbb{IO}(2)$. Обозначим декартовы и полярные координаты на плоскости, соответственно, через x, y и r, φ . Тогда векторные поля Киллинга имеют вид $K_1 = \partial_\varphi$ для вращений и $K_2 = \partial_x, K_3 = \partial_y$ для сдвигов. Квадраты длин векторов Киллинга равны:

$$K_1^2 = r^2, \quad K_2^2 = K_3^2 = 1.$$

Векторы Киллинга K_2 и K_3 имеют постоянную длину на всей плоскости, их траекториями Киллинга являются прямые линии, которые являются экстремальями. Это согласуется с предложением 15.2.2. Траекториями Киллинга для вращений K_1 являются концентрические окружности с центром в начале координат. Длина вектора Киллинга K_1 постоянна на траекториях в соответствии с предложением 15.2.1, однако непостоянна на всей плоскости $\mathbb{R}^2$. Соответствующие траектории Киллинга – окружности – не являются экстремальями. $\square$

Пример 15.2.2. Рассмотрим полупростую группу Ли $\mathbb{G}$, как (псевдо)риманово пространство с формой Киллинга–Картана в качестве метрики (см. раздел 8.5). Это – пространство постоянной кривизны. Левоинвариантные векторные поля генерируют групповые преобразования справа, а правоинвариантные – слева. Групповые преобразования слева и справа сохраняют метрику, и, следовательно, лево- и правоинвариантные векторные поля являются полями Киллинга. Длина этих полей Киллинга равна ± 1 . Поэтому соответствующие траектории Киллинга являются экстремальями. $\square$

Свертывая уравнения Киллинга (15.4) с метрикой, получаем, что дивергенция поля Киллинга равна нулю:

$$\nabla_\alpha K^\alpha = 0. \quad (15.12)$$

Ковариантная производная ∇^β со связностью Леви–Чивиты от уравнения Киллинга (15.4) с учетом уравнения (6.94) для перестановки ковариантных производных и уравнения (15.12) приводит к уравнению

$$\nabla^\beta (\nabla_\beta K_\alpha + \nabla_\alpha K_\beta) = \Delta K_\alpha + (\nabla^\beta \nabla_\alpha - \nabla_\alpha \nabla^\beta) K_\beta = 0,$$

где $\Delta := \nabla^\beta \nabla_\beta$ – оператор Лапласа–Бельтрами на многообразии $\mathbb{M}$, и учтено равенство (15.12). Отсюда вытекает уравнение на компоненты векторов Киллинга

$$\Delta K_\alpha = R_{\alpha\beta} K^\beta, \quad (15.13)$$

Для пространств постоянной кривизны вида (15.25) тензор Риччи выражается через скалярную кривизну (15.24), и уравнение (15.13) принимает вид

$$\Delta K_\alpha = \frac{R}{n} K_\alpha, \quad R = \text{const}.$$

То есть каждая компонента формы Киллинга является собственным вектором оператора Лапласа–Бельтрами.

Предложение 15.2.3. Пусть $X, Y \in \mathcal{X}(\mathbb{M})$ – два произвольных векторных поля на (псевдо)римановом многообразии $(\mathbb{M}, g)$ и K – векторное поле Киллинга. Тогда справедливо равенство

$$g((\mathcal{L}_K - \nabla_K)X, Y) + g(X, (\mathcal{L}_K - \nabla_K)Y) = 0,$$

где $\mathcal{L}_K X = [K, X]$ – производная Ли и $\nabla_K X = K^\alpha (\partial_\alpha X^\beta + \Gamma_{\alpha\gamma}^\beta X^\gamma) \partial_\beta$ – ковариантная производная векторного поля X вдоль поля Киллинга K .

Доказательство. Прямая проверка с учетом явного выражения для символов Кристоффеля (6.24) и уравнения Киллинга (15.4). $\square$

15.3 Однородные и изотропные многообразия

Рассмотрим геодезически полное (псевдо)риманово многообразие $(\mathbb{M}, g)$, метрика которого допускает одно или несколько полных векторных полей Киллинга. Уравнения Киллинга (15.4) накладывают сильные ограничения на векторные поля Киллинга, которые мы сейчас обсудим. Воспользовавшись тождеством для коммутатора ковариантных производных (6.94), получаем равенство

$$\nabla_\alpha \nabla_\beta K_\gamma - \nabla_\beta \nabla_\alpha K_\gamma = -R_{\alpha\beta\gamma}{}^\delta K_\delta. \quad (15.14)$$

Теперь воспользуемся тождеством (6.86) для тензора кривизны и уравнениями Киллинга (15.4). В результате получим тождество для векторных полей Киллинга:

$$\nabla_\alpha \nabla_\beta K_\gamma + \nabla_\beta \nabla_\gamma K_\alpha + \nabla_\gamma \nabla_\alpha K_\beta = 0,$$

где слагаемые отличаются циклической перестановкой индексов. Это равенство позволяет переписать уравнение (15.14) в виде

$$\nabla_\gamma \nabla_\alpha K_\beta = R_{\alpha\beta\gamma}{}^\delta K_\delta. \quad (15.15)$$

Если свернуть данное уравнение по индексам γ, α , то получим в точности уравнение (15.13) из предыдущего раздела.

Полученное равенство (15.15) является следствием уравнений Киллинга, но не эквивалентно им. Тем не менее оно позволяет сделать важные выводы. Предположим, что и (псевдо)риманово многообразие $(\mathbb{M}, g)$, и векторные поля Киллинга вещественно аналитичны, т.е. в окрестности произвольной точки многообразия $p \in \mathbb{M}$ компоненты векторного поля Киллинга разлагаются в ряд Тейлора, который сходится в некоторой окрестности этой точки $\mathbb{U}_p$. Допустим, что в точке $p \in \mathbb{M}$ нам заданы компоненты формы Киллинга $K_\alpha(p)$ и их первых производных $\partial_\beta K_\alpha(p)$. Тогда соотношения (15.15) позволяют вычислить все вторые производные от компонент формы Киллинга $\partial_{\beta\gamma}^2 K_\alpha$ в той же точке p . Теперь возьмем ковариантную производную от равенства (15.15) и получим некоторое соотношение, линейное по третьим производным. Из него можно найти все третьи производные от вектора Киллинга и т.д. до бесконечности. Важно отметить, что все соотношения линейны по компонентам формы Киллинга и их производным. Это значит, что в окрестности $\mathbb{U}_p$ компоненты формы Киллинга имеют вид

$$K_\alpha(x, p) = A_\alpha{}^\beta(x, p) K_\beta(p) + B_\alpha{}^{\beta\gamma}(x, p) [\partial_\beta K_\gamma(p) - \partial_\gamma K_\beta(p)], \quad (15.16)$$

где $A_\alpha^\beta(x, p)$ и $B_\alpha^{\beta\gamma}(x, p)$ – некоторые функции. Антисимметрия последнего слагаемого по индексам β, γ связана с тем, что симметризованная частная производная выражается через компоненты формы Киллинга в силу уравнения Киллинга (15.4). Таким образом, компоненты формы Киллинга в окрестности $\mathbb{U}_p$ являются линейными функциями от компонент формы Киллинга в точке p и ее внешней производной в той же точке.

У формы Киллинга $K_\alpha(x, p)$ второй аргумент p означает, что эта форма имеет определенные свойства в точке $p \in \mathbb{M}$. По предположению, представление (15.16) справедливо для всех точек многообразия $p \in \mathbb{M}$, необходимо только задать значения $K(p)$ и $dK(p)$. Мы предполагаем, что функции $K_\alpha(x, p)$ вещественно аналитичны и по x , и по p .

По предположению, компоненты формы Киллинга разлагаются в ряды Тейлора в окрестности каждой точки $p \in \mathbb{M}$. Обозначим через $\mathbb{U}_p$ окрестность точки p , в которой разложение (15.16) справедливо и обратимо, т.е. аргументы x и p можно поменять местами для некоторых новых матриц A и B . Рассмотрим точку q , которая лежит вне $\mathbb{U}_p$. Для этой точки также справедливо обратимое разложение вида (15.16) в некоторой окрестности $\mathbb{U}_q$. Предположим, что точка q лежит достаточно близко к $\mathbb{U}_p$, так, что окрестности пересекаются, $\mathbb{U}_p \cap \mathbb{U}_q \neq \emptyset$. Тогда для всех точек из пересечения $x \in \mathbb{U}_p \cap \mathbb{U}_q$ справедливо разложение (15.16) по компонентам форм Киллинга $K(p)$ и $K(q)$ и их внешним производным. Отсюда следует, что компоненты формы Киллинга и ее внешней производной в точке q линейно выражаются через компоненты формы Киллинга и ее внешней производной в точке p . Таким образом, разложение (15.16) справедливо также в объединении $\mathbb{U}_p \cup \mathbb{U}_q$. Это построение можно продолжить на все многообразие $\mathbb{M}$. Поэтому разложение (15.16) справедливо для всех точек $x, p \in \mathbb{M}$.

Теперь предположим, что (псевдо)риманово многообразие $(\mathbb{M}, g)$ имеет несколько векторных полей Киллинга K_i , $i = 1, \dots, N$. Тогда для каждого векторного поля Киллинга справедливо разложение (15.16)

$$K_{i\alpha}(x, p) = A_\alpha^\beta(x, p)K_{i\beta}(p) + B_\alpha^{\beta\gamma}(x, p)[\partial_\beta K_{i\gamma}(p) - \partial_\gamma K_{i\beta}(p)]. \quad (15.17)$$

Функции $A_\alpha^\beta(x, p)$ и $B_\alpha^{\beta\gamma}(x, p)$ одинаковы для всех форм Киллинга, потому что определяются соотношениями (15.15), которые линейны по компонентам форм Киллинга и их производным. Они полностью определяются метрикой, тензором кривизны и их ковариантными производными. В полученном разложении точка $p \in \mathbb{M}$ произвольна, но фиксирована, а точка $x \in \mathbb{M}$ пробегает все многообразие.

Соотношение (15.15) представляет собой систему уравнений в частных производных на компоненты формы Киллинга, у которой есть нетривиальные условия разрешимости. Одно из этих условий в ковариантной форме имеет вид

$$[\nabla_\gamma \nabla_\delta] \nabla_\alpha K_\beta = -R_{\gamma\delta\alpha}{}^\epsilon \nabla_\epsilon K_\beta - R_{\gamma\delta\beta}{}^\epsilon \nabla_\alpha K_\epsilon,$$

где квадратные скобки обозначают коммутатор ковариантных производных. Подстановка в левую часть этого уравнения исходного выражения для вторых производных от формы Киллинга (15.15) после несложных алгебраических преобразований приводит к равенству

$$(R_{\alpha\beta\gamma}{}^\epsilon \delta_\delta^\zeta - R_{\alpha\beta\delta}{}^\epsilon \delta_\gamma^\zeta + R_{\gamma\delta\alpha}{}^\epsilon \delta_\beta^\zeta - R_{\gamma\delta\beta}{}^\epsilon \delta_\alpha^\zeta) \nabla_\zeta K_\epsilon = (\nabla_\gamma R_{\alpha\beta\delta}{}^\epsilon - \nabla_\delta R_{\alpha\beta\gamma}{}^\epsilon) K_\epsilon. \quad (15.18)$$

Если кривизна нетривиальна, то это уравнение дает некоторые линейные соотношения между компонентами формы Киллинга K_α и их ковариантными производными

$\nabla_\beta K_\alpha$. Наоборот, если существует некоторая информация в формах Киллинга, то полученное уравнение может определить структуру тензора кривизны. В теореме 15.3.2, которая сформулирована ниже, соотношение (15.18) использовано для доказательства того, что однородное и изотропное многообразие является пространством постоянной кривизны.

Перейдем к определениям.

Определение. (Псевдо)риманово многообразие $(\mathbb{M}, g)$ размерности $\dim \mathbb{M} = n$ называется *однородным в точке* $p \in \mathbb{M}$, если существуют инфинитезимальные изометрии, которые переводят эту точку в любую другую точку из некоторой окрестности $\mathbb{U}_p$ этой точки. Другими словами, метрика должна допускать такие векторные поля Киллинга, которые в точке p имеют все возможные направления. Поскольку векторы Киллинга образуют линейное пространство, то в сопряженном пространстве необходимо и достаточно существования такого набора из n форм Киллинга $K^{(\gamma)} = dx^\alpha K_\alpha^{(\gamma)}(x, p)$, где индекс γ в скобках нумерует формы Киллинга, что выполнены условия:

$$K_\alpha^{(\gamma)}(p, p) = \delta_\alpha^\gamma. \quad (15.19)$$

Если (псевдо)риманово многообразие $(\mathbb{M}, g)$ однородно в каждой своей точке, то оно называется *однородным*. Другими словами, группа изометрий действует на $\mathbb{M}$ транзитивно.

(Псевдо)риманово многообразие $(\mathbb{M}, g)$ называется *изотропным в точке* $p \in \mathbb{M}$, если существуют такие инфинитезимальные изометрии с формами Киллинга $K(x, p)$, которые оставляют эту точку на месте, т.е. $K(p, p) = 0$, и для которых внешняя производная $dK(x, p)$ в точке p принимает любое значение в пространстве 2-форм $\Lambda_2(\mathbb{M})|_p$ в точке p . Для этого необходимо и достаточно существования такого набора из $n(n-1)/2$ форм Киллинга $K^{[\gamma\delta]} = -K^{[\delta\gamma]} = dx^\alpha K_\alpha^{[\gamma\delta]}(x, p)$, где индексы γ, δ нумеруют формы Киллинга, что выполнены условия:

$$\begin{aligned} K_\alpha^{[\gamma\delta]}(p, p) &= 0, \\ \left. \frac{\partial K_\beta^{[\gamma\delta]}(x, p)}{\partial x^\alpha} \right|_{x=p} &= \delta_{\alpha\beta}^{\gamma\delta} - \delta_{\alpha\beta}^{\delta\gamma}. \end{aligned} \quad (15.20)$$

Если (псевдо)риманово многообразие $(\mathbb{M}, g)$ изотропно в каждой своей точке, то оно называется *изотропным*. $\square$

Замечание. В разделе 9.1 было дано определение однородного пространства, как произвольной транзитивной группы преобразований $(\mathbb{M}, \mathbb{G})$. При этом наличие метрики на многообразии $\mathbb{M}$ не предполагалось. Поскольку группа изометрий является группой преобразований, то данное выше определение согласуется с тем, что было дано в разделе 9.1. В обратную сторону утверждение следующее. Пусть дано однородное пространство $(\mathbb{M}, \mathbb{G})$ в смысле группы преобразований. Тогда многообразие $\mathbb{M}$ нетрудно превратить в (псевдо)риманово многообразие с метрикой, инвариантной относительно действия группы $\mathbb{G}$. Для этого достаточно в произвольной точке $p \in \mathbb{M}$ выбрать симметричную невырожденную матрицу и “разнести” ее по всему многообразию при помощи действия группы $\mathbb{G}$. По построению группа $\mathbb{G}$ будет группой изометрий для полученной метрики. $\square$

В силу непрерывности, наборы форм $K^{(\gamma)}$ и $K^{[\gamma\delta]}$ линейно независимы в некоторой окрестности точки p .

Теорема 15.3.1. *Любое изотропное (псевдо)риманово многообразие $(\mathbb{M}, g)$ является также однородным.*

Доказательство. Если многообразие изотропно, то формы Киллинга $K^{[\gamma, \delta]}(x, p)$ и $K^{[\gamma, \delta]}(x, p + dp)$ удовлетворяют условиям (15.20) в близких точках p и $p + dp$, соответственно. Любая их линейная комбинация будет формой Киллинга и, следовательно, произвольная линейная комбинация производных

$$c^\alpha \frac{\partial K_\beta^{[\gamma, \delta]}(x, p)}{\partial p^\alpha} := c^\alpha \lim_{dp^\alpha \rightarrow 0} \frac{K_\beta^{[\gamma, \delta]}(x, p + dp) - K_\beta^{[\gamma, \delta]}(x, p)}{dp^\alpha}$$

также будет формой Киллинга для любого набора постоянных c^α . Вычислим производную по x формы Киллинга $K^{[\gamma, \delta]}$ в точке p . Из первого условия в (15.20) следует равенство

$$\frac{\partial}{\partial p^\alpha} K_\beta^{[\gamma, \delta]}(p, p) = \left. \frac{\partial K_\beta^{[\gamma, \delta]}(x, p)}{\partial x^\alpha} \right|_{x=p} + \left. \frac{\partial K_\beta^{[\gamma, \delta]}(x, p)}{\partial p^\alpha} \right|_{x=p} = 0.$$

Откуда, с учетом второго условия в (15.20), получаем равенство

$$\left. \frac{\partial K_\beta^{[\gamma, \delta]}(x, p)}{\partial p^\alpha} \right|_{x=p} = -\delta_{\alpha\beta}^{\gamma\delta} + \delta_{\alpha\beta}^{\delta\gamma}.$$

Отсюда следует, что из форм Киллинга $K^{[\gamma, \delta]}$ можно построить форму Киллинга, которая в точке p принимает любое заданное значение $dx^\alpha a_\alpha$, где $a_\alpha \in \mathbb{R}$. Для этого достаточно положить

$$K_\alpha := \frac{a_\gamma}{n-1} \frac{\partial K_\alpha^{[\gamma, \delta]}(x, p)}{\partial p^\delta}.$$

Выбрав соответствующим образом постоянные a_γ , получим набор форм Киллинга, который удовлетворяет условиям (15.19). $\square$

Из данной теоремы вытекает, что достаточно говорить изотропное пространство, однако мы предпочитаем название “однородное и изотропное пространство”, т.к. оно отражает важные физические свойства.

Теорема 15.3.2. *Пусть задано связное вещественно аналитическое (псевдо)риманово многообразие $(\mathbb{M}, g)$, метрика которого допускает вещественно аналитические векторные поля Киллинга. Тогда алгебра Ли $\mathfrak{i}(\mathbb{M})$ инфинитезимальных изометрий имеет размерность не более, чем $n(n+1)/2$, где $n := \dim \mathbb{M}$. Максимальная размерность $\dim \mathfrak{i}(\mathbb{M}) = n(n+1)/2$ достигается тогда и только тогда, когда многообразие $\mathbb{M}$ является однородным и изотропным. В этом случае оно представляет собой пространство постоянной кривизны.*

Доказательство. Размерность алгебры Ли $\mathfrak{i}(\mathbb{M})$ равна максимальному числу линейно независимых векторных полей Киллинга на многообразии $\mathbb{M}$. Из равенства (15.17) следует, что число независимых векторных полей Киллинга N не может превышать числа независимых компонент формы $\{K_\alpha(p)\}$ и ее внешней производной $\{\partial_\beta K_\alpha(p) - \partial_\alpha K_\beta(p)\}$ в фиксированной точке $p \in \mathbb{M}$. Число независимых компонент любой 1-формы в фиксированной точке не превосходит n , а число независимых компонент внешней производной не может превышать $n(n-1)/2$. Поэтому справедливо следующее ограничение на размерность алгебры Ли векторных полей Киллинга:

$$\dim \mathfrak{i}(\mathbb{M}) \leq n + \frac{1}{2}n(n-1) = \frac{1}{2}n(n+1).$$

Это доказывает первое утверждение теоремы.

В этом месте важна вещественная аналитичность метрики, т.к. она была использована при получении представления (15.17).

Связность многообразия $\mathbb{M}$ необходима для того, чтобы число независимых векторных полей Киллинга было определено. В противном случае, если многообразие $\mathbb{M}$ имеет несколько компонент связности, число независимых векторных полей Киллинга может отличаться для каждой компоненты связности.

Однородные и изотропные многообразия имеют максимальное число $n(n+1)/2$ векторных полей Киллинга и, в силу разложения (15.17), определяют все возможные векторы Киллинга на многообразии $\mathbb{M}$. Следовательно, если некоторое многообразие имеет максимальное число независимых полей Киллинга, то оно с необходимостью должно быть однородным и изотропным.

Теперь докажем, что любое однородное и изотропное пространство является пространством постоянной кривизны. Если пространство однородно и изотропно, то для каждой точки $x \in \mathbb{M}$ найдутся такие формы Киллинга, для которых $K_\alpha(x) = 0$, а $\nabla_\beta K_\alpha(x)$ является произвольной антисимметричной матрицей. Отсюда следует, что антисимметризованный коэффициент при $\nabla_\zeta K_\epsilon$ в уравнении (15.18) должен быть равен нулю, что приводит к равенству

$$R_{\alpha\beta\gamma}{}^\epsilon \delta_\delta^\zeta - R_{\alpha\beta\delta}{}^\epsilon \delta_\gamma^\zeta + R_{\gamma\delta\alpha}{}^\epsilon \delta_\beta^\zeta - R_{\gamma\delta\beta}{}^\epsilon \delta_\alpha^\zeta = R_{\alpha\beta\gamma}{}^\zeta \delta_\delta^\epsilon - R_{\alpha\beta\delta}{}^\zeta \delta_\gamma^\epsilon + R_{\gamma\delta\alpha}{}^\zeta \delta_\beta^\epsilon - R_{\gamma\delta\beta}{}^\zeta \delta_\alpha^\epsilon. \quad (15.21)$$

Если пространство однородно и изотропно, то для произвольной точки $x \in \mathbb{M}$ существуют также такие формы Киллинга, которые принимают в этой точке произвольные значения. Следовательно, из уравнений (15.18) и (15.21) вытекает равенство

$$\nabla_\gamma R_{\alpha\beta\delta}{}^\epsilon = \nabla_\delta R_{\alpha\beta\gamma}{}^\epsilon. \quad (15.22)$$

Свернем уравнение (15.21) по индексам δ, ζ и опустим верхний индекс. В результате получим выражение тензора кривизны через тензор Риччи и метрику:

$$(n-1)R_{\alpha\beta\gamma\delta} = R_{\beta\delta}g_{\alpha\gamma} - R_{\alpha\delta}g_{\beta\gamma}. \quad (15.23)$$

Поскольку правая часть этой формулы должна быть антисимметрична по индексам δ, γ , то возникает дополнительное ограничение

$$R_{\beta\delta}g_{\alpha\gamma} - R_{\alpha\delta}g_{\beta\gamma} = -R_{\beta\gamma}g_{\alpha\delta} + R_{\alpha\gamma}g_{\beta\delta}.$$

Свертка полученного равенства по индексам β, γ дает связь между тензором Риччи и скалярной кривизной:

$$R_{\alpha\delta} = \frac{1}{n} R g_{\alpha\delta}. \quad (15.24)$$

Подстановка этого выражения в (15.23) приводит к следующему выражению для полного тензора кривизны

$$R_{\alpha\beta\gamma\delta} = \frac{R}{n(n-1)} (g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}). \quad (15.25)$$

Теперь осталось доказать, что скалярная кривизна R однородного и изотропного пространства постоянна. Для этой цели используем свернутые тождества Бианки (6.120)

$$2\nabla_\beta R_\alpha{}^\beta - \nabla_\alpha R = 0.$$

Подставляя в это тождество выражение для тензора Риччи (15.24), получаем условие

$$\left(\frac{2}{n} - 1\right) \partial_\alpha R = 0.$$

При $n \geq 3$ отсюда следует, что $R = \text{const}$.

Случай $n = 2$ требует особого рассмотрения. Свертка равенства (15.22) по индексам β, ϵ приводит к равенству

$$\nabla_\gamma R_{\alpha\delta} - \nabla_\delta R_{\alpha\gamma} = 0.$$

дальнейшая свертка с $g^{\alpha\delta}$ с учетом уравнения (15.24) приводит к условию $\partial_\gamma R = 0$, т.е. $R = \text{const}$ и при $n = 2$.

Таким образом, скалярная кривизна в выражении для полного тензора кривизны (15.25) равна константе, $R = \text{const}$, и максимально симметричное (псевдо)риманово многообразие является пространством постоянной кривизны. $\square$

Замечание. Если тензор кривизны имеет вид (15.25), где $R = \text{const}$, то соответствующее многообразие является пространством постоянной кривизны, т.к. ковариантная производная от метрики в римановой геометрии равна нулю, $\nabla_\epsilon R_{\alpha\beta\gamma\delta} = 0$. Обратное утверждение, вообще говоря, неверно. У пространства постоянной кривизны тензор кривизны не обязательно имеет вид (15.25). Примером является полупростая группа Ли (см. раздел 8.5). Отсюда следует, что не всякое пространство постоянной кривизны является максимально симметричным. $\square$

Пример 15.3.1. Рассмотрим евклидово пространство $\mathbb{R}^n$, на котором задана метрика нулевой кривизны, т.е. $R_{\alpha\beta\gamma\delta} = 0$. Ясно, что это пространство постоянной нулевой кривизны. Тогда в $\mathbb{R}^n$ существует такая система координат x^α , $\alpha = 1, \dots, n$ в которой все компоненты метрики постоянны. В этой системе координат символы Кристоффеля равны нулю и уравнение для векторов Киллинга (15.15) принимает простой вид:

$$\partial_{\beta\gamma}^2 K_\alpha = 0.$$

Общее решение этого уравнения линейно по координатам

$$K_\alpha(x) = a_\alpha + b_{\alpha\beta} x^\beta,$$

где a_α и $b_{\alpha\beta}$ – некоторые постоянные. Из уравнения Киллинга (15.4) следует, что это выражение задает форму Киллинга тогда и только тогда, когда матрица $b_{\alpha\beta}$ антисимметрична, т.е. $b_{\alpha\beta} = -b_{\beta\alpha}$. Следовательно, можно задать $n(n+1)/2$ линейно независимых форм Киллинга:

$$\begin{aligned} K_\alpha^{(\gamma)}(x) &= \delta_\alpha^\gamma, \\ K_\alpha^{[\gamma\delta]}(x) &= \delta_\alpha^\delta x^\gamma - \delta_\alpha^\gamma x^\delta. \end{aligned}$$

Тогда произвольная форма Киллинга выражается в виде линейной комбинации

$$K_\alpha = a_\gamma K_\alpha^{(\gamma)} + \frac{1}{2} b_{\delta\gamma} K_\alpha^{[\gamma\delta]}.$$

В полученном выражении n векторов Киллинга $K^{(\gamma)}$ генерируют трансляции в $\mathbb{R}^n$ вдоль осей координат, а $n(n-1)/2$ векторов $K^{[\gamma\delta]}$ – вращения вокруг начала координат. Таким образом, метрика пространства нулевой кривизны допускает максимальное число $n(n+1)/2$ векторов Киллинга и поэтому является однородным и изотропным пространством.

Известно, что линейным преобразованием координат x^α метрику можно преобразовать к диагональному виду, когда на диагонали будут стоять ± 1 , в зависимости от сигнатуры исходной метрики. Если метрика риманова (положительно определена), то после преобразования координат, она примет стандартный вид $g_{\alpha\beta} = \delta_{\alpha\beta}$. Эта метрика инвариантна относительно неоднородной группы вращений $\mathbb{IO}(n)$. $\square$

Выше мы доказали, что любое однородное и изотропное пространство является пространством постоянной кривизны. Верно также обратное утверждение. А именно, в разделе 18.3 будет доказано, что пространства постоянной кривизны вида (15.25) являются однородными и изотропными.

Замечание. При вейлевском преобразовании метрики $g_{\alpha\beta} \mapsto g'_{\alpha\beta} = kg_{\alpha\beta}$, $k > 0$, скалярная кривизна тоже умножается на постоянный множитель: $R \mapsto R' = R/k$. Поэтому пространства постоянной кривизны с помощью преобразования Вейля можно свести к трем случаям $R = 0, \pm 1$. $\square$

15.4 Симметричные тензоры на пространстве постоянной кривизны

В разделе 15.3 мы выяснили, что однородные и изотропные n -мерные многообразия с необходимостью являются пространствами постоянной кривизны, которое имеют максимальное число $n(n+1)/2$ линейно независимых векторных полей Киллинга. Более того, если под пространством постоянной кривизны понимать (псевдо)риманово многообразие с метрикой, удовлетворяющей условию (15.25), где скалярная кривизна R постоянна, то пространство постоянной кривизны определяется, по существу, единственным образом сигнатурой метрики и знаком скалярной кривизны. Такие пространства часто встречаются в приложениях, причем помимо метрики на таких многообразиях, как правило, задаются дополнительные тензорные поля, например, поля материи. Для того, чтобы вся модель была максимально симметричной необходимо потребовать симметрию не только от метрики, но и от всех остальных полей. В настоящем разделе мы получим условия, которые налагают требования однородности и изотропии на простейшие тензорные поля, заданные на пространстве постоянной кривизны.

Пусть на n -мерном пространстве постоянной кривизны $\mathbb{S}$ помимо метрики $g_{\alpha\beta}$ задано произвольное тензорное поле

$$T = dx^\alpha \otimes \dots \otimes dx^\beta T_{\alpha\dots\beta}.$$

Для определенности мы рассмотрим ковариантные тензорные поля. Пусть задана изометрия $\iota : x \mapsto x'$. Тогда условие симметрии тензорного поля относительно действия данной изометрии имеет тот же вид, что и для метрики (15.1):

$$T(x) = \iota^* T(x'),$$

где ι^* – возврат отображения. В компонентах это условие принимает вид

$$T_{\alpha\dots\beta}(x) = \frac{\partial x'^\gamma}{\partial x^\alpha} \dots \frac{\partial x'^\delta}{\partial x^\beta} T_{\gamma\dots\delta}(x'). \quad (15.26)$$

Пусть инфинитезимальные изометрии генерируются векторными полями Киллинга $K = K^\alpha \partial_\alpha$. Тогда условие симметрии (15.26) запишется в виде равенства нулю производной Ли:

$$\mathcal{L}_K T = 0. \quad (15.27)$$

Такое же условие инвариантности должно выполняться и для произвольных тензорных полей, содержащих как ковариантные, так и контравариантные индексы.

Теперь рассмотрим простейшие случаи, которые часто встречаются в приложениях.

Пример 15.4.1. Пусть на пространстве постоянной кривизны $\mathbb{S}$ задано дифференцируемое скалярное поле $\varphi(x) \in C^1(\mathbb{S})$ (функция). Тогда равенство нулю производной Ли примет вид

$$K^\alpha(x)\partial_\alpha\varphi(x) = 0.$$

Поскольку для пространства постоянной кривизны векторное поле Киллинга можно выбрать таким образом, что компоненты $K^\alpha(x)$ будут принимать произвольные значения в любой точке $x \in \mathbb{S}$, то отсюда вытекает условие постоянства скалярного поля, $\varphi = \text{const}$, на всем $\mathbb{S}$. Таким образом, однородное и изотропное скалярное поле на пространстве постоянной кривизны $\mathbb{S}$ – это постоянная: $\varphi(x) = \text{const}$ для всех $x \in \mathbb{S}$. $\square$

Пример 15.4.2. В качестве второго примера выберем дифференцируемое ковекторное поле $A = dx^\alpha A_\alpha$. Производная Ли для него была вычислена в разделе 2.14, и условие инвариантности (15.27) принимает вид

$$K^\beta\partial_\beta A_\alpha + \partial_\alpha K^\beta A_\beta = 0.$$

Выберем векторное поле Киллинга таким образом, что $K^\beta(x) = 0$ в произвольной, но фиксированной точке $x \in \mathbb{S}$. Кроме этого, векторное поле Киллинга можно выбрать так, что частная производная $\partial_\beta K_\alpha$ будет антисимметрична и произвольна. Поскольку в выбранной точке $\partial_\alpha K^\beta = \nabla_\alpha K^\beta$, то справедливы равенства:

$$\partial_\alpha K^\beta A_\beta = \partial_\alpha K_\beta A^\beta = \partial_\gamma K_\beta (\delta_\alpha^\gamma A^\beta).$$

Данное построение можно провести в произвольной точке многообразия $\mathbb{S}$, и, следовательно,

$$\delta_\alpha^\gamma A^\beta = \delta_\alpha^\beta A^\gamma.$$

После свертки по индексам α и γ возникает соотношение

$$nA^\beta = A^\beta.$$

Поэтому, исключая тривиальный случай $n = 1$, после опускания индекса получаем равенство $A_\alpha = 0$. Следовательно, если ковекторное поле однородно и изотропно, то оно тождественно равно нулю.

Это же относится и к векторному полю $X = X^\alpha\partial_\alpha$: однородное и изотропное векторное поле на пространстве постоянной кривизны $\mathbb{S}$ тождественно равно нулю. $\square$

Пример 15.4.3. В качестве третьего примера рассмотрим дифференцируемый ковариантный тензор второго ранга с компонентами $T_{\alpha\beta}$. Мы не предполагаем наличия какой либо симметрии по индексам α, β . Производная Ли от тензора второго ранга имеет вид

$$\mathcal{L}_K T_{\alpha\beta} = K^\gamma\partial_\gamma T_{\alpha\beta} + \partial_\alpha K^\gamma T_{\gamma\beta} + \partial_\beta K^\gamma T_{\alpha\gamma}.$$

Как и в предыдущем примере выберем векторное поле Киллинга таким образом, чтобы в точке $x \in \mathbb{S}$ было выполнено равенство $K^\gamma(x) = 0$ и частная производная

$\partial_\alpha K_\beta$ была антисимметрична и произвольна. Тогда из равенства нулю производной Ли вытекает равенство

$$\delta_\alpha^\delta T^\gamma_\beta + \delta_\beta^\delta T_\alpha^\gamma = \delta_\alpha^\gamma T^\delta_\beta + \delta_\beta^\gamma T_\alpha^\delta.$$

После свертки по индексам α, δ и опускания γ получаем соотношение

$$(n-1)T_{\gamma\beta} + T_{\beta\gamma} = g_{\beta\gamma}T, \quad T := T_\alpha^\alpha.$$

Теперь поменяем местами индексы β и γ и вычтем полученное равенство:

$$(n-2)(T_{\gamma\beta} - T_{\beta\gamma}) = 0.$$

Отсюда следует, что при $n \neq 2$ инвариантный тензор второго ранга должен быть симметричен. С учетом симметрии получаем выражение для инвариантного тензора второго ранга:

$$T_{\alpha\beta} = \frac{T}{n}g_{\alpha\beta}.$$

Поскольку след тензора T – скаляр, то из его инвариантности вытекает, что он должен быть равен постоянной, как в первом примере. Таким образом, однородное и изотропное ковариантное тензорное поле второго ранга на пространстве постоянной кривизны имеет вид

$$T_{\alpha\beta} = Cg_{\alpha\beta}, \quad C = \text{const.} \quad (15.28)$$

Эта формула справедлива для $n \geq 3$ и для симметричной части тензора при $n = 2$.

В двумерном случае инвариантный тензор может иметь антисимметричную часть, пропорциональную $\varepsilon_{\alpha\beta} = -\varepsilon_{\beta\alpha}$ – полностью антисимметричному тензору второго ранга:

$$T_{[\alpha\beta]} = -T_{[\beta\alpha]} = C\varepsilon_{\alpha\beta},$$

если мы не учитываем пространственные отражения. При пространственных отражениях полностью антисимметричный тензор второго ранга меняет знак $\varepsilon_{\alpha\beta} \mapsto -\varepsilon_{\alpha\beta}$. Поэтому с учетом пространственных отражений наиболее общий вид однородного и изотропного тензора второго ранга при $n = 2$ такой же, как и в более высоких размерностях (15.28).

Аналогичное построение можно провести для инвариантного контравариантного тензора второго ранга и тензора со смешанными индексами:

$$T^{\alpha\beta} = Cg^{\alpha\beta}, \quad T^\alpha_\beta = C\delta^\alpha_\beta.$$

Полученные формулы для симметричных тензоров будут использованы при построении космологических моделей, где роль $T_{\alpha\beta}$ будет играть тензор энергии-импульса полей материи. $\square$

15.5 Пространства с максимально симметричными подпространствами

Во многих важных с физической точки зрения случаях, например, в космологии, (псевдо)риманово многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, представляет собой топологическое произведение двух многообразий, $\mathbb{M} = \mathbb{R} \times \mathbb{S}$, где $\mathbb{R}$ – вещественная прямая, которую

мы в дальнейшем отождествим со временем, и $\mathbb{S}$ – пространство постоянной кривизны. При этом каждой точке $t \in \mathbb{R}$ соответствует подмногообразие $\mathbb{S} \subset \mathbb{M}$. Поскольку $\mathbb{S}$ – пространство постоянной кривизны вида (15.25), то оно однородно и изотропно (теорема 15.3.2). Соответствующая группа изометрий на $\mathbb{S}$ генерируется $n(n-1)/2$ векторными полями Киллинга, где $n := \dim \mathbb{M}$. В настоящем разделе мы найдем наиболее общий вид метрики на $\mathbb{M}$, инвариантной относительно группы симметрий, которая порождается действием группы изометрий на $\mathbb{S}$.

Обозначим координаты на пространстве постоянной кривизны $\mathbb{S}$ через x^μ , $\mu = 1, \dots, n-1$. Пусть компоненты инвариантной метрики на $\mathbb{S}$ будут $\mathring{g}_{\mu\nu}(x)$. По построению эта метрика однородна и изотропна. Она инвариантна относительно группы изометрий, генерируемых векторными полями Киллинга $K_i = K_i^\mu(x)\partial_\mu$, $i = 1, \dots, n(n-1)/2$.

Предположим, что на $\mathbb{M} = \mathbb{R} \times \mathbb{S}$ задана достаточно гладкая метрика лоренцевой сигнатуры такая, что координата t является временем, т.е. $g_{00} > 0$, и все сечения постоянного времени $t = \text{const}$ пространственноподобны. Кроме этого, предположим, что сужение метрики на $\mathbb{S}$ при каждом значении $t \in \mathbb{R}$ совпадает с $\mathring{g}_{\mu\nu}$. Ясно, что такая метрика имеет вид

$$g_{\alpha\beta} = \begin{pmatrix} g_{00} & g_{0\nu} \\ g_{\mu 0} & h_{\mu\nu} \end{pmatrix}, \quad (15.29)$$

где $g_{00}(t, x)$ и $g_{0\mu}(t, x) = g_{\mu 0}(t, x)$ – произвольные функции от t и x , а $h_{\mu\nu}(t, x)$ – метрика постоянной кривизны на $\mathbb{S}$, которая зависит от t как от параметра. Все компоненты метрики предполагаются достаточно гладкими и по t , и по x . Поскольку метрика имеет лоренцеву сигнатуру, то согласно предложению 4.2.2 матрица

$$h_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}}$$

отрицательно определена. Кроме этого, по-предположению, матрица $h_{\mu\nu}$ также отрицательно определена.

Продолжим действие группы изометрий из $\mathbb{S}$ на все $\mathbb{M}$ следующим образом. Будем считать, что компоненты векторных полей Киллинга $K_i^\mu(t, x)$ зависят от t как от параметра. Определим действие группы инфинитезимальных изометрий на $\mathbb{M}$ следующими равенствами:

$$\begin{aligned} t &\mapsto t' = t, \\ x^\mu &\mapsto x'^\mu = x^\mu + \epsilon K^\mu, \quad \epsilon \ll 1, \end{aligned} \quad (15.30)$$

где K – произвольный вектор Киллинга из алгебры Ли, порожденной векторами K_i . То есть преобразования не сдвигают точки вещественной прямой $t \in \mathbb{R} \subset \mathbb{M}$. Это означает, что векторные поля Киллинга продолжаются на все $\mathbb{M}$ таким образом, что у них не возникает дополнительной компоненты: $K^0 = 0$. Нетривиальность продолжения сводится лишь к тому, что пространственные компоненты векторов Киллинга теперь могут зависеть от t как от параметра. Следовательно, алгебра Ли продолженных на $\mathbb{M}$ векторов Киллинга остается прежней.

Пример 15.5.1. В четырехмерном случае векторные поля Киллинга, продолженные на $\mathbb{M} = \mathbb{R} \times \mathbb{S}$, порождают группу преобразований $(\mathbb{M}, \mathbb{G})$, где

$$\mathbb{G} = \begin{cases} \text{SO}(4), & \mathbb{S} = \mathbb{S}^3 \quad \text{– сфера,} \\ \text{ISO}(3), & \mathbb{S} = \mathbb{R}^3 \quad \text{– евклидово пространство,} \\ \text{SO}(3, 1), & \mathbb{S} = \mathbb{H}^3 \quad \text{– двуполостный гиперboloид.} \end{cases}$$

Этот случай важен в космологии. $\square$

Теорема 15.5.1. *Если метрика (15.29) на $M = \mathbb{R} \times \mathbb{S}$ инвариантна относительно преобразований (15.30), то в окрестности произвольной точки существует такая система координат, в которой метрика имеет блочно диагональный вид*

$$ds^2 = dt^2 + h_{\mu\nu} dx^\mu dx^\nu, \quad (15.31)$$

где $h_{\mu\nu}(t, x)$ – метрика постоянной кривизны на $\mathbb{S}$ при всех $t \in \mathbb{R}$. В этой системе координат компоненты векторных полей Киллинга не зависят от времени: $K^\mu = K^\mu(x)$.

Доказательство. Пусть x^μ – координаты на $\mathbb{S}$. Зафиксируем одну из гиперповерхностей $t = \text{const}$. Касательный к ней вектор имеет только пространственные компоненты: $X = X^\mu \partial_\mu$. Ортогональный к ней вектор $n^\alpha \partial_\alpha$ должен удовлетворять равенству

$$n^0 X^\nu N_\nu + n^\mu X^\nu g_{\mu\nu} = 0,$$

где использована АДМ параметризация метрики, см. раздел 24.2. Поскольку данное равенство должно быть выполнено для всех X , то оно определяет пространственные компоненты нормальных векторов,

$$n^\mu = -n^0 N^\mu.$$

Следовательно, квадрат ортогонального вектора положителен:

$$(n, n) = (N^2 + N^\rho N_\rho)(n^0)^2 - 2(n^0)^2 N^\mu N_\mu + (n^0)^2 N^\mu N^\nu g_{\mu\nu} = N^2 (n^0)^2 > 0.$$

Поэтому вектор, ортогональный к пространственноподобной гиперповерхности, является времениподобным.

Выпустим из каждой точки гиперповерхности геодезическую (экстремаль), которая является времениподобной по построению. Выберем в качестве временной координаты длину геодезической s . Не ограничивая общности, можно считать, что фиксированная гиперповерхность соответствует значению $s = 0$. Тогда в некоторой окрестности поверхности $\mathbb{S}$ можно выбрать систему координат $\{x^0 := s, x^\mu\}$. Согласно предложению 23.13.3 в построенной таким образом системе координат метрика имеет блочно диагональный вид в некоторой окрестности фиксированной поверхности:

$$g_{00} = 1, \quad g_{0\mu} = g_{\mu 0} = 0, \quad h_{\mu\nu} = h_{\mu\nu}(s, x).$$

На исходной гиперповерхности $s = 0$ нулевая компонента вектора Киллинга равна нулю, $K^0(0, x) = 0$, по построению. Из $(0, 0)$ компоненты уравнения Киллинга, которое удобнее использовать в форме (15.7), следует равенство $\partial_s K^0(s, x) = 0$. Это дифференциальное уравнение с начальным условием $K^0(0, x) = 0$ имеет единственное решение $K^0(s, x) = 0$ для всех значений координаты s , где определена система координат. Следовательно, все гиперповерхности, определяемые уравнением $s = \text{const}$, будут инвариантными многообразиями, т.е. пространствами постоянной кривизны по крайней мере в некоторой окрестности исходной гиперповерхности.

Если метрика является блочно диагональной (15.31), то $(0, \mu)$ компоненты уравнения Киллинга (15.7) принимают вид $\partial_s K^\mu = 0$. Отсюда следует, что компоненты векторных полей Киллинга не зависят от времени.

Пространственные (μ, ν) компоненты уравнения Киллинга удовлетворяются, поскольку K – векторы Киллинга на $\mathbb{S}$.

Возвращаясь к обозначению $s \mapsto t$, получаем метрику (15.31). $\square$

Если система координат выбрана таким образом, что инвариантная метрика на $\mathbb{M}$ имеет вид (15.31), то координатные линии t , проходящие через каждую точку $x \in \mathbb{S}$ являются геодезическими. Это следует из построения данной системы координат. Данное утверждение уже было доказано с помощью прямой проверки, предложение 23.13.4.

Если метрика имеет блочно диагональный вид (15.31) и $K = K^\mu \partial_\mu$, то уравнения Киллинга (15.7) расщепляются на временные и пространственные компоненты:

$$(\alpha, \beta) = (0, 0) : \quad 0 = 0, \quad (15.32)$$

$$(\alpha, \beta) = (0, \mu) : \quad h_{\mu\nu} \partial_0 K^\nu = 0, \quad (15.33)$$

$$(\alpha, \beta) = (\mu, \nu) : \quad h_{\mu\rho} \partial_\nu K^\rho + h_{\nu\rho} \partial_\mu K^\rho + K^\rho \partial_\rho h_{\mu\nu} = 0. \quad (15.34)$$

Теорема 15.5.2. *В условиях теоремы 15.5.1 метрика (15.31) имеет вид*

$$ds^2 = dt^2 + a^2 g_{\mu\nu} dx^\mu dx^\nu, \quad (15.35)$$

где $a(t) > 0$ – произвольная вещественно аналитическая функция (масштабный множитель) и $\overset{\circ}{g}_{\mu\nu}(x)$ – метрика пространства постоянной кривизны, зависящая только от $x \in \mathbb{S}$.

Доказательство. Поскольку $h_{\mu\nu}(t, x)$ – метрика пространства постоянной кривизны на $\mathbb{S}$ для всех $t \in \mathbb{R}$, то уравнения Киллинга (15.34) выполнены. Из теоремы 15.5.1 следует, что векторные поля Киллинга не зависят от времени. Поэтому дифференцирование уравнения (15.34) по времени приводит к равенству

$$\dot{h}_{\mu\rho} \partial_\nu K^\rho + \dot{h}_{\nu\rho} \partial_\mu K^\rho + K^\rho \partial_\rho \dot{h}_{\mu\nu} = 0.$$

Это значит, что производная метрики по времени $\dot{h}_{\mu\nu}$ является однородным и изотропным тензором второго ранга. Тогда из примера 15.4.3 вытекает, что производная метрики по времени пропорциональна самой метрике:

$$\dot{h}_{\mu\nu} = f h_{\mu\nu}, \quad (15.36)$$

где $f(t)$ – произвольная достаточно гладкая функция времени.

Если $f = 0$, то доказывать нечего, и метрика имеет вид (15.35) с $a = \text{const}$.

Пусть $f \neq 0$. Тогда введем новую временную координату $t \mapsto t'$, определяемую дифференциальным уравнением

$$dt' = f(t) dt.$$

Тогда уравнение (15.36) упростится

$$\frac{dh_{\mu\nu}}{dt'} = h_{\mu\nu}.$$

Его общее решение имеет вид

$$h_{\mu\nu}(t', x) = C e^{t'} \overset{\circ}{g}_{\mu\nu}(x), \quad C = \text{const} \neq 0,$$

где $\overset{\circ}{g}_{\mu\nu}(x)$ – метрика на пространстве постоянной кривизны $\mathbb{S}$, которая не зависит от времени. Отсюда вытекает представление (15.35). $\square$

15.6 Лоренц-инвариантные метрики

Нахождение точных решений уравнений Эйнштейна является одной из основных задач общей теории относительности. Поскольку уравнения Эйнштейна очень сложны, то решения ищутся, как правило, в предположении наличия какой либо симметрии. Это упрощает систему уравнений и дает возможность отыскания решений. В настоящем разделе мы покажем, что лоренц-инвариантные метрики в общем случае зависят от двух произвольных функций одной переменной и найдем соответствующее решение вакуумных уравнений Эйнштейна.

Среди точных решений вакуумных уравнений с космологической постоянной решения де Ситтера [99, 100] и анти-де Ситтера были одними из первых космологических решений. Решение де Ситтера описывает пространство-время постоянной кривизны и инвариантно относительно действия группы Лоренца $\mathbb{SO}(1, 4)$ (см. раздел 35.4). Решение анти-де Ситтера также соответствует пространству постоянной кривизны и инвариантно относительно группы $\mathbb{SO}(2, 3)$. Обе группы симметрии содержат подгруппу Лоренца меньшей размерности $\mathbb{SO}(1, 3)$. Ниже мы покажем, что все решения вакуумных уравнений Эйнштейна, инвариантные относительно действия группы Лоренца $\mathbb{SO}(1, 3)$, описывают пространство-время постоянной кривизны и, следовательно, сводятся или к решению де Ситтера, или к решению анти-де Ситтера в зависимости от знака космологической постоянной.

Рассмотрим пространство-время Минковского $\mathbb{R}^{1, n-1}$ произвольной размерности n . В нем задана метрика Лоренца $\eta_{\alpha\beta} := \text{diag}(+ - \dots -)$ в декартовой системе координат x^α , $\alpha = 0, 1, \dots, n-1$. Эта метрика инвариантна относительно преобразований из группы Пуанкаре, в частности, относительно преобразований Лоренца. Соответствующие векторные поля Киллинга имеют вид

$$K_{\epsilon\delta} = \frac{1}{2} (x_\delta \partial_\epsilon - x_\epsilon \partial_\delta) = \frac{1}{2} (x_\delta \delta_\epsilon^\gamma - x_\epsilon \delta_\delta^\gamma) \partial_\gamma, \quad (15.37)$$

где индексы ϵ, δ нумеруют $n(n-1)/2$ векторов Киллинга и $x_\alpha := x^\beta \eta_{\beta\alpha}$. Пусть в пространстве Минковского $\mathbb{R}^{1, n-1}$ задана вторая метрика $g_{\alpha\beta}(x)$. Поставим следующую задачу: найти все метрики $g_{\alpha\beta}$, инвариантные относительно действия группы Лоренца $\mathbb{SO}(1, n-1)$.

Уравнения

$$\nabla_\alpha K_\beta + \nabla_\beta K_\alpha = 0$$

для векторных полей Киллинга (15.37) принимают вид

$$g_{\alpha\epsilon} \eta_{\beta\delta} - g_{\alpha\delta} \eta_{\beta\epsilon} + g_{\beta\epsilon} \eta_{\alpha\delta} - g_{\beta\delta} \eta_{\alpha\epsilon} + x_\delta \partial_\epsilon g_{\alpha\beta} - x_\epsilon \partial_\delta g_{\alpha\beta} = 0. \quad (15.38)$$

При преобразованиях Лоренца компоненты метрики $g_{\alpha\beta}$ ведут себя как компоненты ковариантного тензора второго ранга. Их необходимо построить из метрики Лоренца $\eta_{\alpha\beta}$ и координат точек $x = \{x^\alpha\}$. Единственная возможность для инвариантной метрики – это метрика вида

$$g_{\alpha\beta} = A \eta_{\alpha\beta} + B x_\alpha x_\beta,$$

где $A(x)$ и $B(x)$ – некоторые функции на $\mathbb{R}^{1, n-1}$. Мы не требуем инвариантности метрики $g_{\alpha\beta}$ относительно трансляций. Поэтому функция B в общем случае отлична от нуля. Подстановка этой метрики в уравнение Киллинга (15.38) ограничивает вид функций A и B . Можно доказать, что они могут быть произвольными функциями от одной переменной

$$s := x^\alpha x^\beta \eta_{\alpha\beta},$$

которая инвариантна относительно преобразований Лоренца. Таким образом, метрика, инвариантная относительно преобразований Лоренца, параметризуется двумя произвольными функциями $A(s)$ и $B(s)$. Ее удобно записать в несколько другом виде

$$g_{\alpha\beta} = f(s)\Pi_{\alpha\beta}^T + g(s)\Pi_{\alpha\beta}^L = f\eta_{\alpha\beta} + (g - f)\frac{x_\alpha x_\beta}{s}, \quad (15.39)$$

где Π^T и Π^L – проекционные операторы (с опущенным индексом):

$$\Pi_{\alpha\beta}^T := \eta_{\alpha\beta} - \frac{x_\alpha x_\beta}{s}, \quad \Pi_{\alpha\beta}^L := \frac{x_\alpha x_\beta}{s},$$

а $f(s) := A(s)$ и $g(s) := A(s) + B(s)s$ – произвольные функции.

Нетрудно вычислить определитель метрики (15.39):

$$\det g_{\alpha\beta} = (-f)^{n-1}g. \quad (15.40)$$

Таким образом, лоренц инвариантная метрика вырождена тогда и только тогда, когда $fg = 0$. Мы будем предполагать, что функции f и g являются достаточно гладкими, $f > 0$ и $g \neq 0$. Кроме этого, предположим также существование предела

$$\lim_{s \rightarrow 0} \frac{f(s) - g(s)}{s},$$

которое необходимо, чтобы метрика была определена при $s = 0$.

Если не оговорено противное, для подъема и опускания индексов будет использоваться метрика Лоренца $\eta_{\alpha\beta}$.

Лоренц инвариантная метрика вида (15.39) при $f = g$ рассматривалась В. А. Фоком [101], §94.

Метрике (15.39) соответствует инвариантный интервал

$$ds^2 = f dx_\alpha dx^\alpha + (g - f) \frac{(x_\alpha dx^\alpha)^2}{s}.$$

Метрический тензор (15.39) имеет одинаковый вид во всех системах координат, связанных между собой преобразованиями Лоренца. Однако его вид меняется при сдвигах $x^\alpha \mapsto x^\alpha + a^\alpha$, поскольку метрика явно зависит от координат, и начало системы отсчета выделено.

Запись метрики (15.39) с помощью проекционных операторов удобна, т.к. обратная метрика имеет простой вид

$$g^{\alpha\beta} = \frac{1}{f}\Pi^{\alpha\beta T} + \frac{1}{g}\Pi^{\alpha\beta L}. \quad (15.41)$$

При $g > 0$ метрике (15.39) можно поставить в соответствие репер

$$e_\alpha^a = \sqrt{f}\delta_\alpha^a + (\sqrt{g} - \sqrt{f})\frac{x_\alpha x^a}{s}. \quad (15.42)$$

Нетрудно проверить следующие свойства проекционных операторов:

$$\begin{aligned} \Pi^{\alpha\beta T} x_\beta &= 0, & \Pi_\alpha^T{}^\alpha &= n - 1, & \partial_\alpha \Pi_{\beta\gamma}^T &= -\frac{\Pi_{\alpha\beta}^T x_\gamma + \Pi_{\alpha\gamma}^T x_\beta}{s}, \\ \Pi^{\alpha\beta L} x_\beta &= x^\alpha, & \Pi_\alpha^L{}^\alpha &= 1, & \partial_\alpha \Pi_{\beta\gamma}^L &= \frac{\Pi_{\alpha\beta}^L x_\gamma + \Pi_{\alpha\gamma}^L x_\beta}{s}, \end{aligned}$$

которые будут использоваться при проведении вычислений.

Несложные вычисления приводят к следующему выражению для символов Кристоффеля, соответствующих метрике (15.39),

$$\Gamma_{\alpha\beta}^{\gamma} = \frac{f'}{f}(x_{\alpha}\Pi_{\beta}^{\tau\gamma} + x_{\beta}\Pi_{\alpha}^{\tau\gamma}) + \frac{g'}{g}(x_{\alpha}\Pi_{\beta}^{\text{L}\gamma} + x_{\beta}\Pi_{\alpha}^{\text{L}\gamma} - x^{\gamma}\Pi_{\alpha\beta}^{\text{L}}) + \frac{g-f-f's}{sg}x^{\gamma}\Pi_{\alpha\beta}^{\tau}, \quad (15.43)$$

где штрих обозначает дифференцирование по аргументу s . Тензор кривизны для метрики (15.39) имеет вид

$$\begin{aligned} R_{\alpha\beta\gamma}^{\delta} = & \Pi_{\alpha\gamma}^{\tau}\Pi_{\beta}^{\tau\delta} \left[\frac{(f+f's)^2}{sf g} - \frac{1}{s} \right] + \\ & + \Pi_{\alpha\gamma}^{\text{L}}\Pi_{\beta}^{\tau\delta} \left[2 \left(\frac{f+f's}{f} \right)' + \left(\frac{f'}{f} - \frac{g'}{g} \right) \frac{f+f's}{f} \right] + \\ & + \Pi_{\alpha}^{\text{L}\delta}\Pi_{\beta\gamma}^{\tau} \left[-2 \left(\frac{f+f's}{g} \right)' + \left(\frac{f'}{f} - \frac{g'}{g} \right) \frac{f+f's}{g} \right] - (\alpha \leftrightarrow \beta), \end{aligned} \quad (15.44)$$

где скобки $(\alpha \leftrightarrow \beta)$ обозначают предыдущие слагаемые с переставленными индексами. Свернув это выражение по индексам β и δ , получим тензор Риччи

$$\begin{aligned} R_{\alpha\beta} = & \Pi_{\alpha\beta}^{\tau} \left[\frac{n-2}{s} \left(\frac{(f+f's)^2}{fg} - 1 \right) + 2 \frac{(f+f's)'}{g} - \left(\frac{f'}{f} + \frac{g'}{g} \right) \frac{f+f's}{g} \right] + \\ & + \Pi_{\alpha\beta}^{\text{L}}(n-1) \left[2 \frac{(f+f's)'}{f} - \left(\frac{f'}{f} + \frac{g'}{g} \right) \frac{f+f's}{f} \right]. \end{aligned} \quad (15.45)$$

Дальнейшая свертка с обратной метрикой (15.41) дает скалярную кривизну

$$R = (n-1) \left[\frac{n-2}{fs} \left(\frac{(f+f's)^2}{fg} - 1 \right) + 4 \frac{(f+f's)'}{fg} - 2 \left(\frac{f'}{f} + \frac{g'}{g} \right) \frac{f+f's}{fg} \right]. \quad (15.46)$$

Пространства постоянной кривизны, определяемые уравнением

$$R_{\alpha\beta\gamma\delta} = -K(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}), \quad (15.47)$$

с некоторой постоянной K , автоматически удовлетворяют вакуумным уравнениям Эйнштейна с космологической постоянной. Решим уравнение (15.47) для лоренц-инвариантной метрики (15.39). Для этого опустим последний индекс у тензора кривизны (15.44) с помощью метрики (15.39)

$$\begin{aligned} R_{\alpha\beta\gamma\delta} = & \Pi_{\alpha\gamma}^{\tau}\Pi_{\beta\delta}^{\tau} \frac{1}{s} \left[\frac{(f+f's)^2}{g} - f \right] + \\ & + (\Pi_{\alpha\gamma}^{\text{L}}\Pi_{\beta\delta}^{\tau} - \Pi_{\alpha\delta}^{\text{L}}\Pi_{\beta\gamma}^{\tau}) \left[2(f+f's)' - \left(\frac{f'}{f} + \frac{g'}{g} \right) (f+f's) \right] - (\alpha \leftrightarrow \beta) \end{aligned} \quad (15.48)$$

и подставим в уравнение (15.47). В результате получим систему дифференциальных уравнений на функции f и g :

$$\frac{(f+f's)^2}{sg} - \frac{f}{s} = -Kf^2, \quad (15.49)$$

$$2(f+f's)' - \left(\frac{f'}{f} + \frac{g'}{g} \right) (f+f's) = -Kfg. \quad (15.50)$$

Из первого уравнения получаем решение для функции g :

$$g = \frac{(f + f's)^2}{f(1 - Kfs)}. \quad (15.51)$$

Поскольку для невырожденности метрики необходимо, чтобы $g \neq 0$, то при $s \neq 0$ функция f должна удовлетворять неравенству

$$f \neq \frac{1}{Ks}, \quad s \neq 0. \quad (15.52)$$

Подстановка выражения (15.51) во второе уравнение (15.50) приводит к тождеству. Таким образом, мы доказали первую часть следующего утверждения.

Теорема 15.6.1. *Лоренц инвариантная метрика*

$$g_{\alpha\beta} = f\Pi_{\alpha\beta}^T + \frac{(f + f's)^2}{f(1 - Kfs)}\Pi_{\alpha\beta}^L, \quad (15.53)$$

где $f(s)$ – произвольная положительная функция, удовлетворяющая условию (15.52), является метрикой пространства постоянной кривизны. Обратно. Метрику пространства постоянной кривизны можно записать в лоренц инвариантном виде (15.53) для некоторой функции $f(s)$.

Доказательство. Нам осталось доказать, что произвольную метрику пространства постоянной кривизны можно привести к лоренц инвариантному виду (15.39). Чтобы ответить на этот вопрос, запишем метрику (15.53) в более известной форме (в стереографических координатах). С этой целью зафиксируем функцию g в исходной метрике (15.39), положив $g = f$. Тогда равенство (15.51) превратится в дифференциальное уравнение

$$f'^2 s + 2f'f + Kf^3 = 0,$$

общее решение которого имеет вид

$$f = \frac{C}{(C + \frac{K}{4}s)^2}, \quad C = \text{const.}$$

Постоянная интегрирования убирается растяжкой координат. Поэтому, без ограничения общности, положим $C = 1$. В результате получим метрику пространства постоянной кривизны

$$g_{\alpha\beta} = \frac{\eta_{\alpha\beta}}{(1 + \frac{K}{4}s)^2}. \quad (15.54)$$

То, что метрику пространства постоянной кривизны можно привести к такому виду – утверждение теоремы 18.3.1. $\square$

Проведенные вычисления просто переносятся на случай метрики в евклидовом пространстве, которая инвариантна относительно $\mathbb{SO}(n)$ вращений. Для этого во всех формулах метрику Лоренца $\eta_{\alpha\beta}$ нужно заменить на евклидову метрику $\delta_{\alpha\beta}$.

Поскольку предел функции $(g - f)/s$ равен

$$\lim_{s \rightarrow 0} \frac{g - f}{s} = 2f' + Kf^2,$$

то выражение для метрики (15.53) определено и при $s = 0$.

Теперь решим вакуумные уравнения Эйнштейна с космологической постоянной Λ

$$R_{\alpha\beta} = \Lambda g_{\alpha\beta}$$

для лоренц инвариантной метрики (15.39). Поскольку число этих уравнений меньше, чем число уравнений в условии постоянства кривизны (15.47), то можно было бы ожидать, что они допускают решения не только с постоянной кривизной. Однако для лоренц-инвариантных метрик классы решений совпадают. Действительно, подстановка тензора Риччи (15.45) в уравнения Эйнштейна приводит к следующей системе уравнений:

$$\frac{n-2}{s} \left[\frac{(f+f's)^2}{fg} - 1 \right] + 2 \frac{(f+f's)'}{g} - \left(\frac{f'}{f} + \frac{g'}{g} \right) \frac{f+f's}{g} = \Lambda f, \quad (15.55)$$

$$(n-1) \left[2 \frac{(f+f's)'}{f} - \left(\frac{f'}{f} + \frac{g'}{g} \right) \frac{f+f's}{f} \right] = \Lambda g. \quad (15.56)$$

Второе уравнение при $\Lambda = -(n-1)K$ совпадает с уравнением (15.50). Линейная комбинация уравнений $(15.55)/f - (15.56)/g$ эквивалентна уравнению (15.49).

Таким образом, мы доказали, что все лоренц инвариантные решения вакуумных уравнений Эйнштейна с космологической постоянной исчерпываются пространствами постоянной кривизны. При $\Lambda > 0$ мы получаем пространство-время де Ситтера, а при $\Lambda < 0$ – анти-де Ситтера. Эти пространства-времена будут рассмотрены, соответственно, в разделах 35.4.2 и 35.4.3.

В заключение раздела покажем, что метрику пространства Минковского $\mathbb{R}^{1,n-1}$ можно записать в новом лоренц-инвариантном виде. Положим $K = 0$ и $f = 1/s^2$. Тогда из формулы (15.51) получаем, что $g = 1/s^2$. Таким образом, метрику (15.39) пространства-времени нулевой кривизны, $K = 0$, можно записать в виде

$$g_{\alpha\beta} = \frac{1}{s^2} \eta_{\alpha\beta}. \quad (15.57)$$

Эта метрика конформно эквивалентна метрике Лоренца и явно инвариантна относительно действия группы Лоренца $\mathbb{O}(1,3)$, но не группы Пуанкаре. Она определена всюду вне светового конуса $s = 0$. Можно проверить прямыми вычислениями, что тензор кривизны для метрики (15.57) равен нулю, $R_{\alpha\beta\gamma\delta} = 0$.

Легко проверить, что метрику произвольного (псевдо)евклидова пространства $\mathbb{R}^{m,n}$ также можно записать в конформно плоском виде (15.57) с соответствующей формой $\eta_{\alpha\beta}$.

15.7 Автомоделные метрики

В настоящем разделе мы рассмотрим еще один пример нахождения метрики по заданному векторному полю Киллинга.

Рассмотрим плоскость $\mathbb{R}^2$ с декартовой системой координат $\{x^\alpha\} = \{u, v\}$ и некоторой метрикой $g_{\alpha\beta}(x)$. Другими словами, рассмотрим координатную окрестность произвольной точки двумерной поверхности и отождествим выбранную систему координат с декартовыми координатами в $\mathbb{R}^2$. Пусть задан набор однопараметрических мультипликативных групп преобразований

$$u \mapsto ku, \quad v \mapsto k^a v, \quad k > 0, \quad (15.58)$$

которые параметризуются некоторым вещественным параметром a . При $a = 1$ это просто масштабные преобразования плоскости. Если параметр преобразования близок к единице, $k = 1 + \epsilon$, $\epsilon \ll 1$, то инфинитезимальные преобразования координат принимают вид

$$\delta u = \epsilon u, \quad \delta v = \epsilon av,$$

что соответствует векторному полю Киллинга

$$K = u\partial_u + av\partial_v \quad \Leftrightarrow \quad \{K^\alpha\} = (u, av). \quad (15.59)$$

Требуется найти наиболее общий вид метрики $g_{\alpha\beta}$, которая допускает векторное поле Киллинга (15.59). Такую метрику будем называть *автомодельной*.

Для решения поставленной задачи необходимо решить уравнения Киллинга. Введем обозначения для компонент метрики:

$$g_{\alpha\beta} = \begin{pmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{pmatrix}, \quad g_{uv} = g_{vu}.$$

Тогда уравнения Киллинга (15.7) примут вид

$$\begin{aligned} 2g_{uu} + u\partial_u g_{uu} + av\partial_v g_{uu} &= 0, \\ (1+a)g_{uv} + u\partial_u g_{uv} + av\partial_v g_{uv} &= 0, \\ 2ag_{vv} + u\partial_u g_{vv} + av\partial_v g_{vv} &= 0. \end{aligned} \quad (15.60)$$

Как видим, система уравнений Киллинга расщепилась на отдельные уравнения для каждой из независимых компонент метрики. Для ее решения перейдем к новым координатам $u, v \mapsto \xi, \eta$, где

$$\xi := \frac{(av)^{1/a}}{u}, \quad (15.61)$$

и $\eta = \eta(u, v)$ – некоторая пока еще неизвестная функция. Тогда для каждой компоненты метрики (относительно старого базиса) справедливо равенство

$$u\partial_u g_{\alpha\beta} + av\partial_v g_{\alpha\beta} = \left(u \frac{\partial \eta}{\partial u} + av \frac{\partial \eta}{\partial v} \right) \partial_\eta g_{\alpha\beta}.$$

Предположим теперь, что функция η является решением дифференциального уравнения

$$u \frac{\partial \eta}{\partial u} + av \frac{\partial \eta}{\partial v} = C = \text{const} \neq 0.$$

Например, можно выбрать

$$\eta = C \ln|u|. \quad (15.62)$$

Тогда уравнения Киллинга сводятся к линейным обыкновенным дифференциальным уравнениям и просто решаются:

$$\begin{aligned} 2g_{uu} + C\partial_\eta g_{uu} &= 0 & \Rightarrow & g_{uu} = h_{uu}(\xi) e^{-2\eta/C}, \\ (1+a)g_{uv} + C\partial_\eta g_{uv} &= 0 & \Rightarrow & g_{uv} = h_{uv}(\xi) e^{-(1+a)\eta/C}, \\ 2ag_{vv} + C\partial_\eta g_{vv} &= 0 & \Rightarrow & g_{vv} = h_{vv}(\xi) e^{-2a\eta/C}. \end{aligned} \quad (15.63)$$

Произвольные функции h_{uu} , h_{uv} и h_{vv} можно рассматривать в качестве компонент некоторой метрики

$$h_{\alpha\beta}(\xi) := \begin{pmatrix} h_{uu}(\xi) & h_{uv}(\xi) \\ h_{vu}(\xi) & h_{vv}(\xi) \end{pmatrix}, \quad h_{uv} = h_{vu}, \quad (15.64)$$

зависящей только от переменной ξ . Для функции (15.62) автомодельная метрика принимает вид

$$g_{\alpha\beta} = \begin{pmatrix} h_{uu}|u|^{-2} & h_{uv}|u|^{-(1+a)} \\ h_{vu}|u|^{-(1+a)} & h_{vv}|u|^{-2a} \end{pmatrix}. \quad (15.65)$$

Таким образом, найден наиболее общий вид автомодельной метрики с точностью до преобразования координат.

Заметим, что переменная ξ (15.61) инвариантна относительно преобразований (15.58). Кроме того, сигнатура автомодельной метрики определяется метрикой (15.64) и может быть произвольной.

Теперь найдем наиболее общий вид функций на $\mathbb{R}^2$, инвариантных относительно преобразований (15.58). Условие инвариантности – это равенство нулю производной Ли от функции $f(u, v) \in C^1(\mathbb{R}^2)$:

$$Kf = 0. \quad (15.66)$$

Любое решение этого уравнения имеет вид $f = f(\xi)$, где переменная ξ была определена формулой (15.61). Это значит, что решения уравнения (15.66) по существу являются функциями одной переменной ξ . То есть они постоянны на орбитах действия группы преобразований, которые параметризуются переменной ξ . Если функция f одновременно удовлетворяет некоторому уравнению в частных производных на поверхности, то его решения, зависящие от переменной ξ , называются *автомодельными*. Для таких решений уравнение в частных производных сводится к обыкновенному дифференциальному уравнению относительно ξ , что существенно упрощает задачу.

Глава 16

Геометрические и экстремали

Пусть на многообразии $\mathbb{M}$ задана аффинная геометрия, т.е. метрика g и аффинная связность Γ . Тогда можно построить два типа выделенных кривых: геометрические и экстремали. Геометрические линии определяются только связностью как линии, касательный вектор к которым остается касательным при параллельном переносе. Экстремали, напротив, определяются только метрикой как линии экстремальной длины. Поскольку метрика и связность являются независимыми геометрическими объектами, то в общем случае геометрические линии и экстремали различны. В частном случае (псевдо)римановой геометрии, когда связностью является связность Леви–Чивиты, геометрические и экстремали совпадают.

В настоящей главе мы рассмотрим оба типа кривых, т.к. они играют важную роль в моделях математической физики. Достаточно сказать, что одним из постулатов общей теории относительности является предположение о том, что свободные точечные частицы, подверженные действию только гравитационных сил, движутся по экстремали. Кроме того, понятие полноты многообразий связано также с экстремалими.

16.1 Геометрические

В аффинной геометрии $(\mathbb{M}, g, \Gamma)$ существует выделенное семейство линий, которые называются геометрическими. Рассмотрим произвольную кривую $\gamma = x(t) = \{x^\alpha(t)\}$, где $-\infty \leq t_1 < t < t_2 \leq \infty$, на многообразии $\mathbb{M}$. Вектор скорости кривой, $u(t) = \{u^\alpha(t) := \dot{x}^\alpha(t)\}$, как всегда, предполагается отличным от нуля.

Определение. Геометрической линией на многообразии $\mathbb{M}$ называется кривая $x(t)$ класса $\mathcal{C}^2([t_1, t_2])$, касательный вектор к которой остается касательным при параллельном переносе вдоль нее. $\square$

Замечание. В определении геометрической линии присутствует только аффинная связность. Поэтому понятие геометрической линии никакого отношения к метрике не имеет, которой может вообще не быть на многообразии. $\square$

Получим уравнения, которым должны удовлетворять координатные функции $x^\alpha(t)$ для того, чтобы кривая γ была геометрической. Выберем произвольный отличный от нуля вектор X_0 , который касается кривой γ в некоторой точке $x(t_0)$, и разнесем его вдоль всей кривой с помощью параллельного переноса. В результате получим векторное поле $X(x(t))$, определенное на кривой γ . Множество векторных полей,

определенных на кривой γ , будем обозначать $\mathcal{X}(\mathbb{M}, \gamma)$. Из определения геодезической следует, что это векторное поле всюду касается γ и, следовательно, пропорционально векторному полю скорости: $X^\alpha(t) = f(t)u^\alpha(t)$, где f – некоторая отличная от нуля функция на γ . Изменим параметризацию кривой $t \mapsto s(t)$. Тогда условие параллельности примет вид

$$X^\alpha = f \frac{ds}{dt} \frac{dx^\alpha}{ds}. \quad (16.1)$$

Выберем новый параметр вдоль геодезической таким образом, чтобы было выполнено уравнение

$$\frac{ds}{dt} = \frac{1}{f},$$

которое всегда имеет решение, поскольку $f \neq 0$. Таким образом, на геодезической линии существует такая параметризация, что вектор скорости u при параллельном переносе остается вектором скорости. В дальнейшем мы предполагаем, что параметр t вдоль геодезической выбран таким образом, что $f = 1$. Если вектор скорости геодезической при параллельном переносе остается касательным, то ковариантная производная от него вдоль геодезической равна нулю (6.31):

$$\nabla_u u = u^\alpha (\partial_\alpha u^\beta + \Gamma_{\alpha\gamma}^\beta u^\gamma) \partial_\beta = 0. \quad (16.2)$$

С формальной точки зрения частная производная $\partial_\alpha u^\beta$ не определена, т.к. векторное поле скорости задано только вдоль кривой γ . Однако, мы можем продолжить это векторное поле достаточно гладким образом в некоторую окрестность кривой γ . При этом результат не зависит от способа продолжения, поскольку дифференцирование в уравнении (16.2) происходит вдоль самой кривой γ .

Поскольку $d/dt = u^\alpha \partial_\alpha$, то это уравнение в компонентах принимает вид

$$\ddot{x}^\alpha = -\Gamma_{\beta\gamma}^\alpha \dot{x}^\beta \dot{x}^\gamma. \quad (16.3)$$

Это уравнение не инвариантно относительно перепараметризации кривой. Однако оно допускает линейную замену параметра

$$t \mapsto at + b, \quad a, b \in \mathbb{R}, \quad a \neq 0. \quad (16.4)$$

Таким образом, мы получили критерий того, что кривая $x(t)$ является геодезической.

Предложение 16.1.1. *Кривая $x(t)$ на многообразии $\mathbb{M}$ с заданной аффинной связностью Γ является геодезической тогда и только тогда, когда существует такая параметризация, что ее координатные функции $x^\alpha(t)$ удовлетворяют системе уравнений (16.3).*

Замечание. В уравнение (16.2) входит частная производная $\partial_\alpha u^\beta$ от вектора скорости. Эта производная неопределена, т.к. векторное поле скорости $u^\beta(x(t))$ задано только вдоль кривой γ . Тем не менее в уравнение входит производная по направлению $u^\alpha \partial_\alpha u^\beta = \ddot{x}^\alpha$, которая имеет смысл. Это значит, что, для определения частной производной $\partial_\alpha u^\beta$, векторное поле скорости можно продолжить в некоторую окрестность кривой γ произвольным дифференцируемым образом, а конечный ответ от такого продолжения не зависит. Подобный трюк будет часто встречаться в дальнейшем. $\square$

Определение. Векторное поле $a(t)$, определенное вдоль произвольной кривой $x(t)$ на многообразии $\mathbb{M}$ с заданной связностью Γ ,

$$a := \nabla_u u = (\ddot{x}^\alpha + \Gamma_{\beta\gamma}^\alpha \dot{x}^\beta \dot{x}^\gamma) \partial_\alpha, \quad (16.5)$$

называется *ускорением* кривой. $\square$

Сравнивая ускорение кривой с уравнением для геодезической, получаем, что кривая будет геодезической тогда и только тогда, когда ее ускорение равно нулю.

Ускорение кривой, так же как и скорость, является векторным полем вдоль кривой, $u, a \in \mathcal{X}(\mathbb{M}, \gamma)$, и зависит от ее параметризации. При выводе уравнений геодезической линии (16.3) была выбрана специальная параметризация кривой γ .

Определение. Параметр t , по которому проводится дифференцирование в системе обыкновенных дифференциальных уравнений (16.3), определяющих геодезическую линию, называется *каноническим* или *аффинным*. $\square$

Поскольку уравнение (16.1) имеет тензорную форму, то канонический параметр не зависит от выбора системы координат на $\mathbb{M}$.

Предложение 16.1.2. Любые два канонических параметра вдоль геодезической связаны между собой линейным преобразованием (16.4).

Доказательство. Рассмотрим вопрос как меняется уравнение для геодезических при произвольной замене канонического параметра. При другой параметризации геодезической $x^\alpha(s)$, где $s = s(t)$, $ds/dt \neq 0$, уравнение (16.3) изменится:

$$\left(\frac{ds}{dt}\right)^2 \frac{d^2 x^\alpha}{ds^2} + \frac{d^2 s}{dt^2} \frac{dx^\alpha}{ds} = - \left(\frac{ds}{dt}\right)^2 \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds}. \quad (16.6)$$

Отсюда следует, что форма уравнений (16.3) не изменится тогда и только тогда, когда $d^2 s/dt^2 = 0$, т.е. замена параметра является аффинной. $\square$

Уравнение (16.6) представляет собой уравнение геодезической $x(s)$ при произвольной параметризации кривой. Действительно, из правила дифференцирования сложной функции вытекает формула

$$\frac{d^2 s}{dt^2} / \left(\frac{ds}{dt}\right)^2 = - \frac{d^2 t}{ds^2} / \frac{dt}{ds}.$$

где

$$\left(\frac{dt}{ds}\right)^2 = g_{\alpha\beta} \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds}.$$

Тогда уравнение для геодезической в произвольной параметризации примет вид

$$\frac{d^2 x^\alpha}{ds^2} = - \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds} + \frac{dx^\alpha}{ds} \frac{d^2 t}{ds^2} / \frac{dt}{ds}. \quad (16.7)$$

Пример 16.1.1. В пространстве Минковского $\mathbb{R}^{1,3}$ точечная частица движется по некоторой мировой линии $x(t)$. Если в качестве параметра вдоль траектории выбрано время $t = x^0$, то скорость и ускорение кривой совпадают со скоростью и ускорением частицы. Если частица свободна, т.е. на нее не действуют никакие силы и,

следовательно, ее ускорение равно нулю, то траекторией частицы будет одна из геодезических линий. Поскольку связность в $\mathbb{R}^{1,3}$ в декартовой системе координат имеет равные нулю компоненты, то уравнения (16.3) сводятся к уравнениям

$$\ddot{x}^\alpha = 0,$$

которые определяют прямые линии. Таким образом, в пространстве Минковского $\mathbb{R}^{1,3}$ прямые и только они являются геодезическими. Это значит, что свободная частица в пространстве Минковского равномерно движется вдоль прямой линии. $\square$

Геодезическая линия в аффинной геометрии обобщает понятие прямой в (псевдо)евклидовом пространстве, сохраняя то свойство, что касательный вектор остается касательным при параллельном переносе.

Пример 16.1.2. В пространствах абсолютного параллелизма, для которых тензор кривизны равен нулю $R_{\alpha\beta\gamma}{}^\delta = 0$, локально существует система координат, в которой симметричная часть аффинной связности равна нулю $\Gamma_{\{\alpha\beta\}}{}^\gamma = 0$ (см., раздел 6.7). Поскольку геодезические линии (16.3) определяются только симметричной частью аффинной связности, то в этой системе координат геодезические являются прямыми линиями. В частности, если на группе Ли задана каноническая связность, т.е. параллельный перенос отождествлен с групповым действием справа, то такая система координат локально существует. Заметим, что в правоинвариантном базисе на группе Ли компоненты связности равны нулю, однако он не является голономным и не определяет ту систему координат, о которой идет речь. $\square$

Для двух параметризаций $x^\alpha(t)$ и $x^\alpha(s)$ одной геодезической справедливо равенство

$$\frac{d^2 s}{dt^2} = \frac{ds}{dt} \frac{dx^\beta}{ds} \partial_\beta \left(\frac{ds}{dt} \right),$$

и уравнение (16.6) переписывается в эквивалентной форме

$$\begin{aligned} \frac{d^2 x^\alpha}{ds^2} &= -\Gamma_{\beta\gamma}{}^\alpha \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds} - \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} \partial_\beta \ln \left| \frac{ds}{dt} \right| = \\ &= -\Gamma_{\beta\gamma}{}^\alpha \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds} - \frac{1}{2} \left(\delta_\beta^\alpha \partial_\gamma \ln \left| \frac{ds}{dt} \right| + \delta_\gamma^\alpha \partial_\beta \ln \left| \frac{ds}{dt} \right| \right) \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds}. \end{aligned} \quad (16.8)$$

Будем говорить, что две геодезические линии $x^\alpha(t)$ и $x^\alpha(s)$ для различных связностей $\Gamma_{1\alpha\beta}{}^\gamma$ и $\Gamma_{2\alpha\beta}{}^\gamma$ на одном и том же многообразии $\mathbb{M}$ имеют одинаковую форму, если множества точек, через которые проходят геодезические линии, совпадают. Тогда из уравнения (16.8) следует критерий совпадения формы геодезических для двух аффинных связностей.

Теорема 16.1.1. Для того, чтобы форма геодезических $x^\alpha(t)$ и $x^\alpha(s)$, где t и s – канонические параметры, для двух аффинных связностей $\Gamma_{1\alpha\beta}{}^\gamma$ и $\Gamma_{2\alpha\beta}{}^\gamma$ на многообразии $\mathbb{M}$ совпадала, необходимо и достаточно, чтобы симметричные части этих связностей были связаны соотношением

$$\Gamma_{2\{\alpha\beta\}}{}^\gamma = \Gamma_{1\{\alpha\beta\}}{}^\gamma + \frac{1}{2} (\delta_\beta^\gamma \partial_\alpha \varphi + \delta_\alpha^\gamma \partial_\beta \varphi), \quad (16.9)$$

где $\varphi = \varphi(x)$ – некоторая функция, определяющая связь канонических параметров для каждой геодезической линии

$$\frac{ds}{dt} = e^\varphi. \quad (16.10)$$

Пусть задан диффеоморфизм многообразий

$$f : \mathbb{M} \rightarrow \mathbb{N}$$

такой, что аффинная связность на $\mathbb{M}$ отображается в аффинную связность на $\mathbb{N}$. Тогда диффеоморфизм f отображает геодезическую линию в геодезическую, поскольку определение геодезической линии инвариантно (не зависит от системы координат).

Уравнения для геодезических (16.3) – это система нелинейных обыкновенных дифференциальных уравнений второго порядка. Поэтому при достаточно гладких компонентах связности для однозначного решения задачи Коши необходимо задать начальную точку $\{x^\alpha(0)\}$ и вектор скорости $\{\dot{x}^\alpha(0)\}$. Геометрически это означает, что через данную точку многообразия в данном направлении проходит одна и только одна геодезическая.

Предложение 16.1.3. *Если аффинная связность Γ на многообразии $\mathbb{M}$ гладкая, то геодезические линии также являются гладкими C^∞ кривыми.*

Доказательство. Продифференцируем уравнение геодезических (16.3) по каноническому параметру. Правая часть полученного равенства зависит от компонент связности и их первых производных, а также от компонент вектора скорости и их первой производной. Поэтому правая часть определена и непрерывна. Следовательно, третья производная от координатных функций $x^\alpha(t)$ существует и непрерывна. Дальнейшее дифференцирование приводит к гладкости геодезических. $\square$

Для геодезической линии можно также поставить краевую задачу: найти геодезическую, соединяющую две фиксированные точки многообразия, которое предполагается линейно связным. Эта задача разрешима в малом, т.е. любые две достаточно близкие точки можно соединить геодезической и притом только одной. Для удаленных точек эта задача может не иметь решения или иметь несколько решений.

Определение. (Псевдо)риманово многообразие $\mathbb{M}$ называется *геодезически выпуклым*, если любые две его точки могут быть соединены геодезической.

Пример 16.1.3. Рассмотрим плоскость с конической сингулярностью и положительным углом дефицита $\mathbb{M} = \mathbb{U} \cup \mathbb{V}$, где $\mathbb{U}$ – область, изометричная евклидовой плоскости $\mathbb{R}^2$ с выколотой полупрямой $\bar{\mathbb{R}}_+ := \{(x, y) \in \mathbb{R}^2 : x \geq 0, y = 0\}$, и $\mathbb{V}$ – клин евклидовой плоскости, который вставлен (см. рис.16.1). Тогда на ней существуют точки, которые нельзя соединить геодезической. Действительно, все геодезические, проходящие через точку p , соединяют ее со всеми точками евклидовой плоскости $\mathbb{R}^2$, до того, как клин был вставлен. Поскольку при диффеоморфизме $(\mathbb{R}^2 \setminus \bar{\mathbb{R}}_+ \text{ на } \mathbb{U})$ геодезическая переходит в геодезическую и их число сохраняется, то точку p нельзя соединить геодезической ни с какой точкой q , лежащей на клине $\mathbb{V}$, который вставлен со стороны, противоположной конической сингулярности. Это многообразие не является геодезически выпуклым. $\square$

Пример 16.1.4. Рассмотрим сферу, вложенную в евклидово пространство, $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. Пусть метрика на сфере индуцирована вложением и на ней задана связность Леви–Чивиты. Тогда любые две точки на сфере $\mathbb{S}^2$, не являющиеся диаметрально противоположными, можно соединить двумя разными геодезическими (две дуги большой окружности, проходящей через эти точки). Диаметрально противоположные точки соединяются бесконечным числом геодезических. Сфера $\mathbb{S}^2$ является геодезически выпуклым многообразием. $\square$

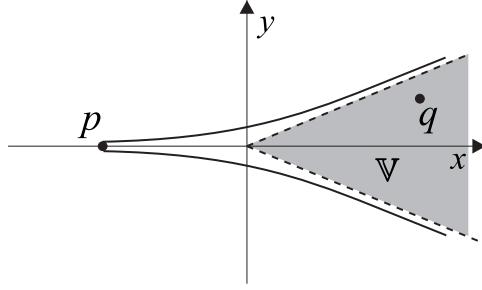


Рис. 16.1: Точки p и q на плоскости с конической сингулярностью и положительным углом дефицита нельзя соединить геодезической.

Введем важное понятие полноты геодезической по каноническому параметру t , В силу однозначного разрешения задачи Коши, через каждую точку многообразия в заданном направлении проходит одна геодезическая. Это значит, что геодезическую линию, если она заканчивается в некоторой точке $q \in \mathbb{M}$, всегда можно продолжить. Действительно, если при конечном значении канонического параметра геодезическая попадает в точку q , то продолжим ее, склеив с геодезической, выходящей из точки q в том же направлении.

Определение. Геодезическая в $\mathbb{M}$ называется *полной*, если ее можно продолжить в обе стороны до бесконечного значения канонического параметра, $t \in (-\infty, \infty)$. $\square$

Замечание. Поскольку канонический параметр определен с точностью до аффинного преобразования и не зависит от выбора системы координат, то данное определение корректно. $\square$

Геодезическую линию можно рассматривать как интегральную кривую векторного поля скорости $u := \{\dot{x}^\alpha\}$. Тогда полнота геодезической означает полноту векторного поля u .

Пусть на многообразии $\mathbb{M}$ задана не только аффинная связность Γ , но и метрика g . Тогда компоненты аффинной связности можно выразить через метрику, кручение и неметричность по формуле (6.17). Хотя в уравнение для геодезических входит только симметричная часть связности, тем не менее оно нетривиально зависит от кручения и неметричности. Действительно, из формулы (6.17) следует, что симметричная часть аффинной связности имеет вид

$$\begin{aligned} \Gamma_{\{\beta\gamma\}}^\alpha &= \frac{1}{2}(\Gamma_{\beta\gamma}^\alpha + \Gamma_{\gamma\beta}^\alpha) = \\ &= \tilde{\Gamma}_{\beta\gamma}^\alpha + \frac{1}{2}(T^\alpha_{\beta\gamma} + T^\alpha_{\gamma\beta}) + \frac{1}{2}(Q_{\beta\gamma}^\alpha + Q_{\gamma\beta}^\alpha - Q^\alpha_{\beta\gamma}), \end{aligned} \quad (16.11)$$

где $\tilde{\Gamma}_{\beta\gamma}^\alpha$ – символы Кристоффеля. Ясно, что две связности, имеющие одинаковую симметричную часть определяют одно и то же семейство геодезических.

Рассмотрим два вектора X и Y , которые параллельно переносятся вдоль геодезической γ .

Предложение 16.1.4. Зависимость скалярного произведения (X, Y) двух векторов, которые параллельно переносятся вдоль γ , от точки геодезической определяется только тензором неметричности:

$$\partial_u(X, Y) = \nabla_u(X^\alpha Y^\beta g_{\alpha\beta}) = -u^\gamma X^\alpha Y^\beta Q_{\gamma\alpha\beta}. \quad (16.12)$$

Отсюда следует, что в римановой геометрии и геометрии Римана–Картана, где $Q = 0$, скалярное произведение двух векторов при параллельном переносе вдоль геодезической сохраняется. В частности, квадрат вектора скорости геодезической постоянен вдоль нее.

Доказательство. Следует из определения тензора неметричности (6.16). $\square$

Замечание. Это утверждение верно и для произвольной кривой (предложение 6.4.2). Скалярное произведение двух векторных полей, полученных в результате параллельного переноса двух векторов вдоль произвольной кривой γ в римановой геометрии и геометрии Римана–Картана, не зависит от точки кривой. $\square$

Следствие. Если неметричность лоренцева многообразия $\mathbb{M}$ равна нулю, то квадрат вектора скорости постоянен вдоль геодезических, и их можно разделить на три класса: времениподобные, пространственноподобные и светоподобные (изотропные). При этом тип геодезической не может меняться от точки к точке. $\square$

В заключение настоящего раздела получим уравнение девиации геодезических. Пусть на многообразии $\mathbb{M}$ задана произвольная кривая $\sigma = y(s)$, $s \in [s_1, s_2]$ и отличное от нуля векторное поле $X(s) \in \mathcal{X}(\mathbb{M}, \sigma)$, определенное вдоль этой кривой и которое нигде не касается кривой σ . Рассмотрим семейство геодезических $\gamma_s = x(t, s)$, проходящих через каждую точку кривой σ в направлении $X(s)$:

$$x(0, s) = y(s), \quad \dot{x}(0, s) = X(s),$$

где точка, как и раньше, обозначает дифференцирование по каноническому параметру t . Если кривая $y(s)$ и векторное поле $X(s)$ достаточно гладкие, то множество точек $x(t, s)$ образует двумерное подмногообразие (поверхность) Σ в $\mathbb{M}$ при достаточно малых t . При этом параметры t, s можно выбрать в качестве координат на Σ . Этим координатам соответствуют векторные поля

$$u := \partial_t = \frac{\partial x^\alpha}{\partial t} \partial_\alpha, \quad Y := \partial_s = \frac{\partial x^\alpha}{\partial s} \partial_\alpha, \quad (16.13)$$

которые коммутируют между собой, $[u, Y] = 0$. После замены в коммутаторе частных производных на ковариантные, получим равенство

$$u^\beta \nabla_\beta Y^\alpha = Y^\beta \nabla_\beta u^\alpha + u^\beta Y^\gamma T_{\beta\gamma}{}^\alpha, \quad (16.14)$$

где $T_{\beta\gamma}{}^\alpha := \Gamma_{\beta\gamma}{}^\alpha - \Gamma_{\gamma\beta}{}^\alpha$ – тензор кручения.

Определение. Векторное поле $Y \in \mathcal{X}(\mathbb{M}, \Sigma)$, определенное равенством (16.13), называется *вектором девиации* геодезических. Векторные поля с компонентами:

$$V^\alpha := u^\beta \nabla_\beta Y^\alpha, \quad (16.15)$$

$$A^\alpha := u^\beta \nabla_\beta V^\alpha, \quad (16.16)$$

называются, соответственно, *скоростью* и *ускорением девиации* близких геодезических. $\square$

Векторное поле девиации определяет расположение близких геодезических на поверхности Σ , а векторные поля скорости и ускорения девиации характеризуют, как меняются расположение геодезических при движении вдоль геодезической.

Предложение 16.1.5. Ускорение девиации геодезических определяется тензором кривизны и кручения:

$$A^\alpha = u^\gamma u^\delta Y^\beta R_{\gamma\beta\delta}{}^\alpha + u^\gamma u^\delta \nabla_\delta (Y^\beta T_{\gamma\beta}{}^\alpha). \quad (16.17)$$

Доказательство. Прямая проверка:

$$\begin{aligned} u^\gamma \nabla_\gamma (u^\beta \nabla_\beta Y^\alpha) &= u^\gamma \nabla_\gamma (Y^\beta \nabla_\beta u^\alpha + u^\beta Y^\delta T_{\beta\delta}{}^\alpha) = \\ &= u^\gamma \nabla_\gamma Y^\beta \nabla_\beta u^\alpha + u^\gamma Y^\beta \nabla_\gamma \nabla_\beta u^\alpha + u^\beta u^\gamma \nabla_\gamma (Y^\delta T_{\beta\delta}{}^\alpha) = \\ &= u^\gamma \nabla_\gamma Y^\beta \nabla_\beta u^\alpha + u^\gamma Y^\beta \nabla_\beta \nabla_\gamma u^\alpha + u^\gamma u^\delta Y^\beta R_{\gamma\beta\delta}{}^\alpha - u^\gamma Y^\beta T_{\gamma\beta}{}^\delta \nabla_\delta u^\alpha + u^\beta u^\gamma \nabla_\gamma (Y^\delta T_{\beta\delta}{}^\alpha) = \\ &= Y^\gamma \nabla_\gamma (u^\beta \nabla_\beta u^\alpha) + u^\gamma u^\delta Y^\beta R_{\gamma\beta\delta}{}^\alpha + u^\beta u^\gamma \nabla_\gamma (Y^\delta T_{\beta\delta}{}^\alpha) = \\ &= u^\gamma u^\delta Y^\beta R_{\gamma\beta\delta}{}^\alpha + u^\beta u^\gamma \nabla_\gamma (Y^\delta T_{\beta\delta}{}^\alpha), \end{aligned}$$

где мы учли равенство (16.14), уравнения геодезических (16.2) и выражение коммутатора ковариантных производных через тензоры кривизны и кручения (6.93). $\square$

В (псевдо)римановой геометрии уравнение (16.17) упрощается

$$A^\alpha = u^\gamma u^\delta Y^\beta \tilde{R}_{\gamma\beta\delta}{}^\alpha. \quad (16.18)$$

В литературе по общей теории относительности оно известно, как *уравнение девиации геодезических*. Это уравнение можно переписать в виде

$$\nabla_u^2 Y^\alpha + u^\delta u^\beta Y^\gamma \tilde{R}_{\delta\beta}{}^\alpha{}_\gamma = 0. \quad (16.19)$$

Оно совпадает с равенством нулю второй вариации действия для экстремалей (16.40), которое будет получено позже в разделе 16.4.

16.2 Экстремали

Другим выделенным типом кривых в аффинной геометрии $(\mathbb{M}, g, \Gamma)$ являются экстремали, которые определяются как линии экстремальной длины. Рассмотрим произвольную достаточно гладкую кривую $\gamma = x(t) \in \mathbb{M}$, $t \in [t_1, t_2]$. Предположим, что квадрат вектора скорости кривой, $u^\alpha := \dot{x}^\alpha(t)$, всюду отличен от нуля, $u^2 := u^\alpha u^\beta g_{\alpha\beta} \neq 0$.

Замечание. Для римановой метрики это условие автоматически выполняется, поскольку вектор скорости предполагается отличным от нуля. Если метрика не является знакоопределенной, то это условие нетривиально. Например, для лоренцевой метрики это условие равносильно тому, что мы рассматриваем либо времени-, либо пространственноподобные кривые. $\square$

Пусть кривая соединяет две точки $p = \{x^\alpha(t_1)\}$ и $q = \{x^\alpha(t_2)\}$. Тогда длина этой кривой задается интегралом

$$S = \int_p^q ds, \quad ds := dt \sqrt{|g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta|} = dt \sqrt{|u^2|}. \quad (16.20)$$

Этот функционал инвариантен относительно общих преобразований координат x^α и произвольной перепараметризации кривой $t \rightarrow \tau(t)$.

Определение. Экстремалью, соединяющей две точки (псевдо)риманова многообразия $p, q \in \mathbb{M}$, если она существует, называется неизотропная кривая γ класса $\mathcal{C}^2([t_1, t_2])$, для которой функционал (16.20) принимает экстремальное значение. $\square$

Замечание. Если метрика на многообразии $\mathbb{M}$ не является знакоопределенной, то существуют изотропные кривые, для которых функционал длины (16.20) равен нулю и данное выше определение экстремалей не проходит. Определение изотропных экстремалей будет дано ниже. $\square$

Экстремали в римановом пространстве обобщают понятие прямой в евклидовом пространстве, сохраняя свойство быть линиями минимальной (экстремальной) длины, соединяющей две точки.

Пример 16.2.1. Не любые две точки линейно связного многообразия $p, q \in \mathbb{M}$, на котором задана риманова метрика, можно соединить кривой наименьшей длины. Рассмотрим кольцо на евклидовой плоскости $\mathbb{R}^2$, рис.16.2. Это – риманово некомпактное связное неодносвязное двумерное многообразие. Если отрезок, соединяющий точки p и q , пересекает вырезанную дырку, то эти точки нельзя соединить кривой наименьшей длины. Действительно, для любой кривой γ , соединяющей точки p и q , всегда найдется кривая γ' меньшей длины. Построение ясно из рисунка. $\square$

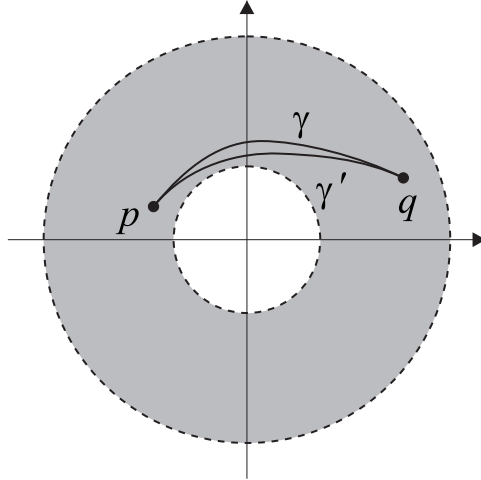


Рис. 16.2: Точки p и q нельзя соединить экстремалью.

Найдем уравнения, которым должны удовлетворять координатные функции $x^\alpha(t)$ для того, чтобы кривая $x(t)$ была экстремалью. Пусть экстремаль соединяет две точки p и q многообразия. Выберем такую карту на многообразии, которая целиком содержит данную экстремаль. Для этого достаточно взять объединение всех достаточно малых шаров, центры которых лежат на экстремали. Пусть в этой карте экстремаль и ее вариация задаются набором функций $x^\alpha(t)$ и $\delta x^\alpha(t)$. Мы предполагаем, что вариации кривой в конечных точках $p, q \in \mathbb{M}$ равны нулю, $\delta x^\alpha(p) = \delta x^\alpha(q) = 0$, и поэтому вкладом граничных членов при интегрировании по частям можно пренебречь. Вариация длины кривой (16.20) с точностью до знака имеет вид

$$\delta S = \int \frac{dt}{2\sqrt{|u^2|}} [2\delta(\dot{x}^\alpha)\dot{x}^\beta g_{\alpha\beta} + \dot{x}^\alpha \dot{x}^\beta \delta g_{\alpha\beta}].$$

Проинтегрируем первое слагаемое по частям и воспользуемся равенством

$$\delta g_{\beta\gamma} = \partial_\alpha g_{\beta\gamma} \delta x^\alpha.$$

Тогда вариация длины кривой принимает вид

$$\delta S = - \int dt \left[\frac{d}{dt} \left(\frac{\dot{x}^\beta g_{\alpha\beta}}{\sqrt{|u^2|}} \right) - \frac{1}{2\sqrt{|u^2|}} \dot{x}^\beta \dot{x}^\gamma \partial_\alpha g_{\beta\gamma} \right] \delta x^\alpha.$$

Поскольку

$$ds = \sqrt{|u|^2} dt, \quad \text{и} \quad \dot{x}^\alpha = \sqrt{|u^2|} \frac{dx^\alpha}{ds},$$

то вариацию длины кривой можно переписать в виде

$$\delta S = - \int ds \left(\frac{d^2 x^\beta}{ds^2} + \tilde{\Gamma}_{\gamma\delta}{}^\beta \frac{dx^\gamma}{ds} \frac{dx^\delta}{ds} \right) g_{\alpha\beta} \delta x^\alpha,$$

где $\tilde{\Gamma}_{\beta\gamma}{}^\alpha$ – символы Кристоффеля (6.24). Фактически, на этом этапе мы воспользовались инвариантностью интеграла (16.20) относительно перепараметризации кривой, выбрав длину кривой в качестве параметра, $t \mapsto s$. Параметр s вдоль экстремали называется *каноническим*. В дальнейшем канонический параметр мы будем обозначать буквой t . Таким образом, мы доказали следующее утверждение.

Теорема 16.2.1. *Для того, чтобы кривая $x(t)$ была экстремалью в канонической параметризации, необходимо и достаточно, чтобы координатные функции $x^\alpha(t)$ удовлетворяли системе обыкновенных дифференциальных уравнений*

$$\ddot{x}^\alpha = -\tilde{\Gamma}_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{x}^\gamma, \quad (16.21)$$

где точка обозначает дифференцирование по каноническому параметру t .

Замечание. Параметр t , по которому проводится дифференцирование в уравнении (16.21) так же, как и для геодезических, называется каноническим или аффинным. Он определен с точностью до аффинного преобразования. Таким образом, канонический параметр в общем случае не равен, а пропорционален длине экстремали. $\square$

В приложениях иногда полезно использовать произвольную параметризацию экстремалей. Обозначим экстремаль в произвольной параметризации через $\{x^\alpha(u)\}$, где u – произвольный параметр. Он связан с каноническим параметром некоторым достаточно гладким и невырожденным преобразованием $t \mapsto u(t)$, $dt/du \neq 0$. Равенство (16.6) представляет собой уравнение экстремалей (если $\Gamma_{\beta\gamma}{}^\alpha$ – символы Кристоффеля), записанные в произвольной параметризации. Запишем их в виде

$$\frac{d^2 x^\alpha}{du^2} = -\tilde{\Gamma}_{\beta\gamma}{}^\alpha \frac{dx^\beta}{du} \frac{dx^\gamma}{du} - \frac{dx^\alpha}{du} \frac{d^2 u}{dt^2} \left(\frac{dt}{du} \right)^2. \quad (16.22)$$

Легко проверить формулу дифференцирования

$$\frac{d^2 u}{dt^2} \left(\frac{dt}{du} \right)^2 = -\frac{d^2 t}{du^2} \frac{du}{dt}.$$

Тогда уравнение экстремалей в произвольной параметризации (16.22) примет вид

$$\frac{d^2 x^\alpha}{du^2} = -\tilde{\Gamma}_{\beta\gamma}{}^\alpha \frac{dx^\beta}{du} \frac{dx^\gamma}{du} + \frac{dx^\alpha}{du} \frac{d^2 t}{du^2} \frac{du}{dt}. \quad (16.23)$$

Если в качестве канонического параметра выбрана длина экстремали $t = s$,

$$ds^2 := g_{\alpha\beta} dx^\alpha dx^\beta,$$

то равенство (16.23) необходимо дополнить уравнением на параметр

$$\left(\frac{ds}{du}\right)^2 = g_{\alpha\beta} \frac{dx^\alpha}{du} \frac{dx^\beta}{du}. \quad (16.24)$$

Отметим, что уравнения для экстремалей (16.22) и (16.23) инвариантны относительно линейного преобразования канонического параметра $t \mapsto kt$, $k \neq 0$. Поэтому выбор длины экстремали в качестве канонического параметра не является существенным. Если же изменить параметризацию экстремали $u \mapsto ku$, то уравнение целиком умножится на постоянную $1/k^2$, т.е., по существу, не изменится.

Пример 16.2.2. Пусть в некоторой области $\mathbb{U} \subset \mathbb{M}$ (псевдо)риманова многообразия существует такая система координат, в которой метрика имеет постоянные компоненты. Тогда в этой области символы Кристоффеля равны нулю и, следовательно, уравнения (16.21) принимают вид $\ddot{x}^\alpha = 0$. Поэтому прямые и только они являются экстремальными в рассматриваемой области $\mathbb{U}$. $\square$

Мы получили критерий того, что неизотропная кривая является экстремалью. При выводе уравнений (16.21) из вариационного принципа для действия (16.20) существенно используется условие $u^2 \neq 0$, которое исключает изотропные (светоподобные) экстремали. Поэтому изотропные экстремали определим не с помощью функционала длины, а непосредственно уравнениями (16.21). Для любой изотропной кривой $u^2 = 0$ и интеграл (16.20) равен нулю. В то же время уравнения (16.21) имеют смысл.

Определение. *Изотропной экстремалью* в канонической параметризации называется изотропная кривая γ класса $\mathcal{C}^2([t_1, t_2])$, которая задана функциями $x^\alpha(t)$, удовлетворяющими системе обыкновенных дифференциальных уравнений (16.21). $\square$

Это определение корректно, т.к. квадрат вектора скорости экстремали постоянен. При этом не всякая изотропная кривая является экстремалью.

Пример 16.2.3. В четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$ любая изотропная кривая имеет вид (1.195). В то же время экстремальными являются прямые линии и только они. Следовательно, не всякая изотропная кривая является экстремалью. Отметим, что в двумерном пространстве-времени любая изотропная кривая является экстремалью, что нетрудно доказать в изотермических координатах (конформной калибровке). $\square$

Если параметр вдоль изотропной экстремали $x^\alpha(u)$ не канонический, то система уравнений меняется на уравнения (16.23). При этом длина касательного вектора остается нулевой.

Определения изотропных и неизотропных экстремалей можно объединить.

Определение. Кривая γ класса $\mathcal{C}^2([t_1, t_2])$ называется *экстремалью*, если она задается функциями $x^\alpha(t)$, где t – канонический параметр, удовлетворяющими системе дифференциальных уравнений (16.21). Другими словами, экстремалью на (псевдо)римановом многообразии $(\mathbb{M}, g)$ называется геодезическая для связности Леви-Чивиты. $\square$

Замечание. Мы начали с другого определения экстремалей, т.к. в последнем определении теряется основное свойство неизотропных экстремалей реализовывать экстремумы функционала длины (16.20). $\square$

Уравнение для экстремалей (16.21) можно переписать в явно ковариантном виде

$$\dot{x}^\beta \widetilde{\nabla}_\beta \dot{x}^\alpha = 0. \quad (16.25)$$

Это просто уравнение геодезических (16.2) для связности Леви–Чивиты. Опустив в этом уравнении индекс α и расписав ковариантную производную, получим альтернативную форму уравнения для экстремалей

$$\frac{\partial}{\partial t}(g_{\alpha\beta}\dot{x}^\beta) - \frac{1}{2}\partial_\alpha g_{\beta\gamma}\dot{x}^\beta\dot{x}^\gamma = 0. \quad (16.26)$$

Уравнения (16.21) определяются только метрикой, поскольку функционал длины не зависит от аффинной связности. Тем самым экстремали не зависят от того заданы ли на многообразии $\mathbb{M}$ тензоры кручения и неметричности или нет. Сравнение уравнений (16.21) с уравнением для геодезических (16.3) показывает, что экстремали являются геодезическими линиями по отношению к параллельному переносу, определяемому символами Кристоффеля. Это означает, что все свойства геодезических справедливы также и для экстремалей. В частности, канонический параметр вдоль экстремали инвариантен относительно преобразования координат. При произвольной параметризации уравнение для экстремалей имеет вид (16.23).

Из предложения 16.1.3 следует, что если метрика g на многообразии является гладкой функцией, то экстремали являются гладкими кривыми класса C^∞ .

Через произвольную точку $x \in \mathbb{M}$ в направлении X^α проходит одна и только одна экстремаль, поскольку она определяется системой обыкновенных дифференциальных уравнений второго порядка (16.21). Это значит что любую экстремаль, которая заканчивается в некоторой точке $q \in \mathbb{M}$ можно продолжить. Действительно, если она заканчивается в точке $q \in \mathbb{M}$ при значении канонического параметра t_2 , то существует единственная экстремаль, проходящая через q и имеющая тот же касательный вектор.

Определение. Экстремаль γ в $\mathbb{M}$ называется *полной*, если ее можно продолжить до бесконечного значения канонического параметра в обе стороны, $t \in (-\infty, \infty)$. $\square$

Замечание. При продолжении экстремали возможны два случая. Во-первых, она может оказаться полной и иметь бесконечную длину. К полным экстремалиям мы относим также и замкнутые экстремали, которые имеют конечную длину. Хотя их длина конечна, но канонический параметр продолжается до бесконечности, что соответствует бесконечному числу проходов вдоль экстремали. Во-вторых, при конечном значении канонического параметра экстремаль может попасть в такую точку многообразия, в которой один из геометрических инвариантов, например, скалярная кривизна, неопределен. Эта точка является сингулярной, и продолжение экстремали через нее не имеет смысла. $\square$

Очевидно, что в (псевдо)римановой геометрии экстремали и геодезические совпадают, поскольку совпадают уравнения (16.3) и (16.21) при $\Gamma_{\alpha\beta}^\gamma = \widetilde{\Gamma}_{\alpha\beta}^\gamma$. В общем случае экстремали и геодезические не совпадают и их поведение должно исследоваться отдельно. Критерий совпадения экстремалей и геодезических дает следующая

Теорема 16.2.2. В аффинной геометрии $(\mathbb{M}, g, \Gamma)$ экстремали и геодезические совпадают тогда и только тогда, когда кручение и неметричность удовлетворяют соотношению

$$\frac{1}{2}(T_{\alpha\beta\gamma} - T_{\gamma\alpha\beta}) = Q_{\alpha\beta\gamma}. \quad (16.27)$$

Доказательство. Экстремали и геодезические совпадают если и только если симметричная часть компонент связности равна символам Кристоффеля, $\Gamma_{\{\beta\gamma\}}^\alpha = \tilde{\Gamma}_{\beta\gamma}^\alpha$. Тогда из (16.11) следует соотношение между тензором кручения и неметричностью:

$$T_{\alpha\beta\gamma} - T_{\gamma\alpha\beta} = -Q_{\beta\gamma\alpha} - Q_{\gamma\alpha\beta} + Q_{\alpha\beta\gamma}. \quad (16.28)$$

Сделав циклическую перестановку индексов $\alpha\beta\gamma \mapsto \beta\gamma\alpha$ и сложив полученное уравнение с (16.28) получим (16.27). Обратно. Подстановка тензора неметричности (16.27) в уравнение (16.28), как легко проверить, приводит к тождеству. $\square$

Следствие. В геометрии Римана–Картана, где $Q = 0$, экстремали и геодезические совпадают тогда и только тогда, когда тензор кручения с опущенным индексом антисимметричен по всем трем индексам, $T_{\alpha\beta\gamma} = T_{[\alpha\beta\gamma]}$.

16.3 Интегрирование уравнений для экстремалей и геодезических

Уравнения для экстремалей и геодезических в ряде случаев имеют первые интегралы, наличие которых существенно упрощает их исследование. Начнем с универсального закона сохранения. Из сравнения уравнений (16.21) и (16.3) следует, что экстремаль является геодезической линией для связности, определяемой символами Кристоффеля. Поскольку при параллельном переносе и в римановой геометрии, и в геометрии Римана–Картана длины векторов не меняется, то отсюда сразу следует, что длина вектора скорости $u = \{\dot{u}^\alpha\}$ вдоль экстремалей и геодезических постоянна.

Предложение 16.3.1. В геометрии Римана–Картана и (псевдо)римановой геометрии для уравнений геодезических (16.3) и экстремалей (16.21) существует первый интеграл

$$C_0 = u^2 := \dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta} = \text{const}, \quad (16.29)$$

квадратичный по первым производным (скоростям).

Доказательство. Следствие предложения 16.1.4. Можно доказать и формально, продифференцировав уравнение (16.29) по каноническому параметру и воспользовавшись уравнением для геодезических или экстремалей. $\square$

Рассмотрим как меняется квадрат длины касательного вектора $C_0(t) = \dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}$ вдоль геодезической линии в аффинной геометрии общего вида. Дифференцируя это соотношение по каноническому параметру и используя уравнение для геодезических, получим

$$\frac{du^2}{dt} = \dot{x}^\alpha \dot{x}^\beta \dot{x}^\gamma Q_{\alpha\beta\gamma}. \quad (16.30)$$

Отсюда следует, что $C_0 = \text{const}$ при нулевом тензоре неметричности. В геометрии Римана–Картана–Вейля уравнение (16.30) принимает вид

$$\frac{du^2}{dt} = \dot{x}^\alpha W_\alpha C_0(t). \quad (16.31)$$

Первый интеграл (16.29) имеет кинематический характер и существует для любой экстремали и геодезической в (псевдо)римановой геометрии и геометрии Римана–Картана. Если метрика имеет лоренцеву сигнатуру, то экстремали и геодезические можно разделить на три класса: времениподобные, $C_0 > 0$, изотропные или светоподобные, $C_0 = 0$, и пространственноподобные, $C_0 < 0$. Поскольку канонический параметр определен с точностью до аффинных преобразований, то для времениподобных и пространственноподобных экстремалей его всегда можно выбрать таким образом, что $C_0 = \pm 1$. В этом случае для времениподобных экстремалей канонический параметр называется *собственным временем*, а для пространственноподобных – *длиной* экстремали.

Замечание. Отметим, что, если некоторая кривая, не обязательно экстремаль или геодезическая, имеет определенный тип, то вдоль нее параметр всегда можно выбрать таким образом, что будет выполнено условие (16.29). Для изотропных кривых равенство (16.29), очевидно, выполняется. Допустим, что кривая имеет определенный тип в некоторой области, т.е. $C_0(t) \neq 0$. Тогда, вводя новый параметр $s(t)$, получим

$$C_0(t) = \left(\frac{ds}{dt} \right)^2 \dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta},$$

где точка обозначает дифференцирование по s . Уравнение

$$\frac{ds}{dt} = \sqrt{|C_0|}$$

всегда имеет решение. Поэтому условие (16.29) будет выполнено относительно нового параметра s . $\square$

Существование других первых интегралов связано с инфинитезимальными симметриями метрики, которые определяются векторными полями Киллинга.

Предложение 16.3.2. *Если метрика на многообразии имеет один или несколько векторов Киллинга $K_i = \{K_i^\alpha\}$, $i = 1, \dots, m$, то для каждого вектора Киллинга имеется свой интеграл движения и для экстремалей, и для геодезических*

$$C_i = K_i^\alpha \dot{x}^\beta g_{\alpha\beta} = \text{const}, \quad i = 1, \dots, m, \quad (16.32)$$

который линеен по компонентам скорости.

Доказательство. Дифференцируем соотношения (16.32) по каноническому параметру и используем уравнения (16.3) или (16.21). $\square$

16.4 Вторая вариация уравнений для экстремалей

Допустим, что метрика положительно определена, т.е. многообразие риманово. Экстремали на римановом многообразии (M, g) определяются действием (16.20). Тогда любая экстремаль удовлетворяет уравнениям Эйлера–Лагранжа (16.21). Это свойство является необходимым условием для того, чтобы экстремаль была линией наименьшей или наибольшей длины. Для того, чтобы найти достаточное условие реализации минимума или максимума функционалов (16.20) или (16.44) необходимо исследовать вторую вариацию функционала.

Напомним общее определение второй вариации. Пусть задана достаточно гладкая кривая $\gamma = \{x^\alpha(t)\} \in \mathbb{M}$, где $t \in [t_1, t_2]$, соединяющая две фиксированные точки многообразия $p = \gamma(t_1)$ и $q = \gamma(t_2)$. Рассмотрим функционал

$$S[\gamma] = \int_{\gamma} dt L(x, \dot{x}). \quad (16.33)$$

Допустим, что заданы два произвольных векторных поля $X, Y \in \mathcal{X}(\mathbb{M}, \gamma)$, которые определены на кривой γ и равны нулю в точках p и q . Обозначим через λ и μ два вещественных параметра.

Определение. Назовем *первой вариацией* функционала (16.33) частную производную

$$\left. \frac{\partial}{\partial \lambda} S[\gamma + \lambda X] \right|_{\lambda=0} = \int_{\gamma} dt \frac{\delta S}{\delta x^\alpha} X^\alpha, \quad (16.34)$$

Частная производная

$$G_\gamma[X, Y] = G_\gamma[Y, X] = \left. \frac{\partial^2}{\partial \lambda \partial \mu} S[\gamma + \lambda X + \mu Y] \right|_{\lambda=0, \mu=0} \quad (16.35)$$

называется *второй вариацией* функционала (16.33). $\square$

Поскольку векторное поле X произвольно, то равенство нулю первой вариации эквивалентно уравнениям Эйлера–Лагранжа, которые мы записываем в виде

$$\frac{\delta S}{\delta x^\alpha} = \frac{\partial L}{\partial x^\alpha} - \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \dot{x}^\alpha} \right) = 0. \quad (16.36)$$

Предложение 16.4.1. Если кривая $\gamma = \{x^\alpha(t)\}$ удовлетворяет уравнениям Эйлера–Лагранжа (16.36), то вторая вариация функционала (16.33) имеет вид

$$G_\gamma[X, Y] = - \int_{t_1}^{t_2} dt (J_{\alpha\beta} X^\beta) Y^\alpha, \quad (16.37)$$

где

$$J_{\alpha\beta} X^\beta = \frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{x}^\alpha \partial \dot{x}^\beta} \dot{X}^\beta + \frac{\partial^2 L}{\partial \dot{x}^\alpha \partial x^\beta} X^\beta \right) - \frac{\partial^2 L}{\partial x^\alpha \partial \dot{x}^\beta} \dot{X}^\beta - \frac{\partial^2 L}{\partial x^\alpha \partial x^\beta} X^\beta. \quad (16.38)$$

Доказательство. Прямые вычисления с учетом выражения для первой вариации:

$$\begin{aligned} G_\gamma[X, Y] &= \left. \frac{\partial}{\partial \mu} \int dt \left(\frac{\partial L}{\partial x^\beta} X^\beta + \frac{\partial L}{\partial \dot{x}^\beta} \dot{X}^\beta \right) \right|_{\mu=0} = \\ &= \int dt \left(\frac{\partial^2 L}{\partial x^\alpha \partial x^\beta} X^\beta Y^\alpha + \frac{\partial^2 L}{\partial \dot{x}^\alpha \partial x^\beta} X^\beta \dot{Y}^\alpha + \frac{\partial^2 L}{\partial x^\alpha \partial \dot{x}^\beta} \dot{X}^\beta Y^\alpha + \frac{\partial^2 L}{\partial \dot{x}^\alpha \partial \dot{x}^\beta} \dot{X}^\beta \dot{Y}^\alpha \right). \end{aligned}$$

После интегрирования по частям двух слагаемых, пропорциональных $\dot{Y}^\alpha$, получаем выражение (16.38). $\square$

Определение. Линейный оператор $J = \{J^\alpha_\beta := g^{\alpha\gamma} J_{\gamma\beta}\}$,

$$J : \mathcal{X}(\mathbb{M}, \gamma) \rightarrow \mathcal{X}(\mathbb{M}, \gamma),$$

который действует на векторные поля X , определенные вдоль кривой γ , по правилу (16.38), называется *оператором Якоби*. $\square$

Теперь применим формулу (16.37) для вычисления второй вариации действия для экстремалей. Во-первых, отметим, что действие для экстремалей в виде (16.20) не является единственным. Рассмотрим действие

$$S = \int_{\gamma} dt \frac{1}{2} g_{\alpha\beta} \dot{x}^{\alpha} \dot{x}^{\beta}, \quad (16.39)$$

Предложение 16.4.2. Действие (16.39) имеет тот же набор экстремалей, что и действие (16.20).

Доказательство. Вариация действия (16.39) по $x^{\alpha}(t)$ приводит к уравнениям (16.21). $\square$

В отличие от действия (16.20) функционал (16.39) не инвариантен относительно перепараметризации кривой. Это значит, что параметр t в действии (16.39) является каноническим.

Пусть $\gamma = \{x^{\alpha}(t)\}$, $t \in [t_1, t_2]$, – экстремаль и $u := \{\dot{x}^{\alpha}\}$ – вектор скорости экстремали.

Теорема 16.4.1. Вторая вариация действия для экстремалей (16.39) имеет вид

$$G_{\gamma}(X, Y) = - \int_{t_1}^{t_2} dt \left(\nabla_u^2 Y^{\alpha} + u^{\delta} u^{\beta} Y^{\gamma} \tilde{R}_{\gamma\delta\beta}{}^{\alpha} \right) X_{\alpha}, \quad (16.40)$$

где

$$\begin{aligned} \nabla_u Y^{\alpha} &:= u^{\beta} \left(\partial_{\beta} Y^{\alpha} + \tilde{\Gamma}_{\beta\gamma}{}^{\alpha} Y^{\gamma} \right) = \dot{Y}^{\alpha} + u^{\beta} \tilde{\Gamma}_{\beta\gamma}{}^{\alpha} Y^{\gamma}, \\ \nabla_u^2 Y^{\alpha} &:= u^{\delta} \nabla_{\delta} \left(\dot{Y}^{\alpha} + u^{\beta} \tilde{\Gamma}_{\beta\gamma}{}^{\alpha} Y^{\gamma} \right) = \\ &= \ddot{Y}^{\alpha} + 2u^{\beta} \tilde{\Gamma}_{\beta\gamma}{}^{\alpha} \dot{Y}^{\gamma} + u^{\delta} u^{\epsilon} Y^{\gamma} \left(\partial_{\delta} \tilde{\Gamma}_{\epsilon\gamma}{}^{\alpha} - \tilde{\Gamma}_{\delta\epsilon}{}^{\beta} \tilde{\Gamma}_{\gamma\beta}{}^{\alpha} + \tilde{\Gamma}_{\delta\gamma}{}^{\beta} \tilde{\Gamma}_{\epsilon\beta}{}^{\alpha} \right) \end{aligned} \quad (16.41)$$

– ковариантные производные вдоль вектора скорости и $\tilde{R}_{\gamma\delta\beta}{}^{\alpha}$ – тензор кривизны.

Доказательство. Первая вариация действия (16.39) имеет вид

$$\left. \frac{\partial S[\gamma + \lambda X]}{\partial \lambda} \right|_{\lambda=0} = - \int_{t_1}^{t_2} dt \left(\ddot{x}^{\alpha} + \tilde{\Gamma}_{\beta\gamma}{}^{\alpha} \dot{x}^{\beta} \dot{x}^{\gamma} \right) X_{\alpha}.$$

Поэтому

$$\begin{aligned} G_{\gamma}(X, Y) &= \\ &= - \frac{\partial}{\partial \mu} \int_{t_1}^{t_2} dt \left(\ddot{x}^{\alpha} + \mu \ddot{Y}^{\alpha} + \tilde{\Gamma}_{\beta\gamma}{}^{\alpha} \dot{x}^{\beta} \dot{x}^{\gamma} + 2\mu \tilde{\Gamma}_{\beta\gamma}{}^{\alpha} \dot{x}^{\beta} \dot{Y}^{\gamma} + \mu \partial_{\epsilon} \tilde{\Gamma}_{\gamma\delta}{}^{\alpha} \dot{x}^{\beta} \dot{x}^{\gamma} Y^{\epsilon} \right) X_{\alpha} = \\ &= - \int_{t_1}^{t_2} dt \left(\ddot{Y}^{\alpha} + 2\tilde{\Gamma}_{\beta\gamma}{}^{\alpha} \dot{x}^{\beta} \dot{Y}^{\gamma} + \partial_{\epsilon} \tilde{\Gamma}_{\gamma\delta}{}^{\alpha} \dot{x}^{\beta} \dot{x}^{\gamma} Y^{\epsilon} \right), \end{aligned}$$

где мы оставили только линейные по μ слагаемые и учли, что вклад производной $\partial X_{\alpha} / \partial \mu$ равен нулю при выполнении уравнений для экстремалей $\nabla_u \dot{x}^{\alpha} = 0$. Исключив из полученного выражения производные $\ddot{Y}^{\alpha}$ и $\dot{Y}^{\alpha}$ с помощью второго из соотношений (16.41), получим равенство (16.40). $\square$

Ясно, что билинейная форма (16.40) симметрична, $G_\gamma(X, Y) = G_\gamma(Y, X)$, т.к. ковариантное дифференцирование можно перекинуть на вектор X , а тензор кривизны симметричен относительно перестановки пар индексов: $\tilde{R}_{\alpha\beta\gamma\delta} = \tilde{R}_{\gamma\delta\alpha\beta}$ (6.87).

Поскольку метрика $g_{\alpha\beta}$ риманова, то условие минимальности экстремали γ , соединяющей точки p и q , состоит в том, что квадратичная форма $G_\gamma(X, X)$ положительно определена для всех векторных полей, определенных на экстремали γ и обращающихся в нуль на ее концах.

Определение. Векторное поле X , которое определено на экстремали γ , соединяющей точки p и q , называется *якобиевым*, если оно удовлетворяет уравнению Якоби

$$J^\alpha_\beta X^\beta = 0, \quad (16.42)$$

где J – оператор Якоби (16.38), и обращается в нуль на концах p и q . Точки p и q называются *сопряженными* вдоль экстремали γ , если существует ненулевое якобиево поле X вдоль γ . $\square$

Для действия (16.39) уравнение Якоби принимает вид

$$\nabla_u^2 X^\alpha + u^\delta u^\beta X^\gamma \tilde{R}_{\gamma\delta\beta}{}^\alpha = 0. \quad (16.43)$$

Это уравнение совпадает с уравнением девиации геодезических (16.19).

Предложение 16.4.3. *Билинейная форма $G_\gamma(X, Y)$ невырождена тогда и только тогда, когда концевые точки p и q экстремали γ не сопряжены вдоль γ .*

Доказательство. Пусть X и Y произвольные векторные поля вдоль γ , которые обращаются в нуль в концевых точках. Напомним, что билинейная форма $G_\gamma(X, Y)$ называется невырожденной, если не существует такого векторного поля X , что $G_\gamma(X, Y) = 0$ для всех Y . Если поле X якобиево, то $G_\gamma(X, Y) = 0$ при любом Y .

Обратно. Допустим, что для вектора X выполнено уравнение $G_\gamma(X, Y) = 0$ для всех Y . Положим $Y = f(t)JX$, где функция $f(t)$ неотрицательна и обращается в нуль на концах экстремали. Тогда из выражения для второй вариации (16.37) имеем равенство

$$G_\gamma(X, Y) = - \int_{t_1}^{t_2} dt f g(JX, JX) = 0.$$

Отсюда следует, что $JX = 0$. Поэтому концы сопряжены. $\square$

Теперь сформулируем необходимое условие минимальности экстремали.

Теорема 16.4.2. *Если экстремаль γ , соединяющая точки p и q , содержит внутри себя пару сопряженных точек p' и q' , то она не является минимальной.*

Доказательство. См., например, [7], глава 5, §36, теорема 2. $\square$

Теорема 16.4.3. *Для достаточно малых отрезков $l = [t_1, t_2]$ экстремали задают минимум функционала (16.39). Поэтому каждая экстремаль является кратчайшей линией в классе дважды дифференцируемых кривых, соединяющих достаточно близкие точки p и q .*

Доказательство. Экстремаль γ реализует минимум функционала (16.39), если форма $G_\gamma(X, X)$ положительна для всех ненулевых векторных полей $X \in \mathcal{X}(\mathbb{M}, \gamma)$, которые обращаются в нуль на концах p и q . Из формулы для второй вариации (16.40) следует, что

$$\begin{aligned} G_\gamma(X, X) &= - \int_p^q dt \left[(\nabla_u^2 X, X) - u^\alpha X^\beta u^\gamma X^\delta \tilde{R}_{\alpha\beta\gamma\delta} \right] = \\ &= \int_p^q dt \left[(\nabla_u X, \nabla_u X) + u^\alpha X^\beta u^\gamma X^\delta \tilde{R}_{\alpha\beta\gamma\delta} \right], \end{aligned}$$

где мы проинтегрировали первое слагаемое по частям с учетом равенств $X(p) = X(q) = 0$. Можно доказать, что для достаточно коротких экстремалей справедлива оценка

$$\left| \int_p^q dt u^\alpha X^\beta u^\gamma X^\delta \tilde{R}_{\alpha\beta\gamma\delta} \right| = O(l) \int_p^q dt (\nabla_u X, \nabla_u X)$$

при $l \rightarrow 0$. Поскольку квадратичная форма $(\nabla_u X, \nabla_u X)$ положительна, то для достаточно коротких экстремалей положительна и форма $G_\gamma(X, X)$. $\square$

16.5 Уравнение Гамильтона–Якоби для экстремалей

Уравнения для экстремалей $\gamma = x(t)$ вытекают из вариационного принципа для действия (16.20). Важным обстоятельством является то, что уравнения для экстремалей являются уравнениями Эйлера–Лагранжа также и для другого действия (16.39). А именно, рассмотрим действие

$$S_m = \int_a^b dt L_m, \quad L_m = -\frac{1}{2} m g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta, \quad (16.44)$$

где точка обозначает дифференцирование по каноническому параметру t и $m = \text{const}$ – постоянная, имеющая физический смысл массы точечной частицы. Это действие совпадает с (16.39) при $m = -1$. Действие (16.44) приводит к уравнениям для экстремалей, в которых переменная t уже является каноническим параметром. Это согласуется с тем обстоятельством, что рассмотренное действие инвариантно относительно общих преобразований координат и сдвигов параметра t . Для сравнения напомним, что исходное действие для экстремалей (16.20) инвариантно также относительно произвольных преобразований параметра t вдоль экстремали.

Действие (16.44) имеет простой физический смысл. Предположим, что метрика имеет лоренцеву сигнатуру и зафиксируем временную калибровку (23.174)

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix}, \quad \text{sign } g_{\mu\nu} = (- \dots -). \quad (16.45)$$

Символы Кристоффеля для этой метрики имеют вид (23.180). Предположим также, что пространственная часть метрики $g_{\mu\nu}$ не зависит от времени x^0 . Тогда уравнения для экстремалей расщепляются:

$$\ddot{x}^0 = 0, \quad (16.46)$$

$$\ddot{x}^\mu = -\tilde{\Gamma}_{\nu\rho}^{\mu} \dot{x}^\nu \dot{x}^\rho. \quad (16.47)$$

Из первого уравнения следует, что, не ограничивая общности, канонический параметр t можно отождествить с временем $x^0 = ct$, где $c = \text{const}$ – скорость света. Тогда лагранжиан (16.44) имеет прямой физический смысл – с точностью до аддитивной постоянной это кинетическая энергия точечной частицы, которая движется в римановом пространстве со статической метрикой $g_{\mu\nu}(x)$. Несмотря на то, что потенциальная энергия частицы равна нулю, ее траекториями уже не будут прямые линии, если метрика нетривиально зависит от точки пространства.

Вернемся к исходному действию (16.44) до фиксирования временной калибровки. Переформулируем эту лагранжеву систему на гамильтоновом языке, рассматривая канонический параметр t в качестве параметра эволюции. Под временем мы подразумеваем координату x^0 и, соответственно, предполагаем, что $g_{00} > 0$. Кроме этого мы предполагаем, что все сечения постоянного времени $x^0 = \text{const}$ пространственно-подобны. Импульс, сопряженный координатам x^α , и гамильтониан системы равны

$$\begin{aligned} p_\alpha &= \frac{\partial L_m}{\partial \dot{x}^\alpha} = -mg_{\alpha\beta}\dot{x}^\beta, \\ H_m &= -\frac{1}{2m}g^{\alpha\beta}p_\alpha p_\beta. \end{aligned} \quad (16.48)$$

Для действия (16.44) связи на канонические переменные отсутствуют, т.к. метрика рассматривается как внешнее поле и варьирование по ней не проводится. Соответствующие уравнения Гамильтона (уравнения движения) имеют вид

$$\dot{x}^\alpha = [x^\alpha, H_m] = -\frac{1}{m}g^{\alpha\beta}p_\beta, \quad (16.49)$$

$$\dot{p}_\alpha = [p_\alpha, H_m] = \frac{1}{2m}\partial_\alpha g^{\beta\gamma}p_\beta p_\gamma. \quad (16.50)$$

Дифференцируя первое из этих уравнений по каноническому параметру и исключая импульсы p_α и производные $\dot{p}_\alpha$ с помощью уравнений движения, нетрудно проверить, что система уравнений (16.49), (16.50) эквивалентна системе уравнений для экстремалей (16.21). Тем самым мы переписали уравнения для экстремалей в виде канонической системы уравнений движения.

Ранее было доказано, что длина касательного вектора к экстремали постоянна (16.29). В гамильтоновой форме это утверждение имеет вид

$$C_0 = \frac{1}{m^2}g^{\alpha\beta}p_\alpha p_\beta = \text{const}. \quad (16.51)$$

По своей физической сути это есть закон сохранения энергии точечной частицы. В данном случае только кинетической, т.к. потенциальная энергия тождественно равна нулю.

В дальнейшем нам понадобятся гамильтоновы уравнения для нулевых экстремалей, где в качестве параметра эволюции выбрано время x^0 , а не канонический параметр t . Они получаются следующим образом. Для нулевых экстремалей интеграл движения (16.51) принимает вид

$$g^{00}p_0^2 + 2g^{0\mu}p_0 p_\mu + g^{\mu\nu}p_\mu p_\nu = 0.$$

Это квадратное уравнение решается относительно p_0 :

$$p_0 = N^\mu p_\mu \pm N\hat{p}, \quad \hat{p} := \sqrt{-\hat{g}^{\mu\nu}p_\mu p_\nu}, \quad (16.52)$$

где $N = 1/\sqrt{g^{00}} \neq 0$ и $N^\mu = \hat{g}^{\mu\nu}g_{0\nu}$ – функции хода и сдвига, используемые в АДМ параметризации метрики (24.5), а $\hat{g}^{\mu\nu}$ – метрика, обратная к пространственной метрике $g_{\mu\nu}$. Если частица движется, то $p_\mu \neq 0$ и $\hat{p} \neq 0$. Тогда из уравнения (16.49) находим производную координаты x^0 по каноническому параметру

$$\dot{x}^0 = \mp \frac{\hat{p}}{mN}.$$

Отсюда следует, что выбор знака в (16.52) соответствует выбору взаимной ориентации канонического параметра t и времени x^0 . После этого канонические уравнения (16.49), (16.50) можно записать в виде системы уравнений движения только для пространственных координат и импульсов:

$$\begin{aligned} \partial_0 x^\mu &= \frac{\dot{x}^\mu}{\dot{x}^0} = -N^\mu \pm \frac{N\hat{g}^{\mu\nu}p_\nu}{\hat{p}}, \\ \partial_0 p_\mu &= \frac{\dot{p}_\mu}{\dot{x}^0} = \partial_\mu N^\nu p_\nu \pm \partial_\mu N \hat{p} \mp \frac{N\partial_\mu \hat{g}^{\nu\rho}p_\nu p_\rho}{2\hat{p}}. \end{aligned} \quad (16.53)$$

Таким образом, из системы канонических уравнений (16.49), (16.50) для нулевых экстремалей мы исключили в явном виде канонический параметр t и нулевую компоненту импульса p_0 .

Предложение 16.5.1. Система уравнений для пространственных компонент канонически сопряженных переменных x^μ и p_μ (16.53) является гамильтоновой. При этом эволюция системы уравнений рассматривается по отношению к времени x^0 , а гамильтонианом является выражение для p_0 (16.52).

Доказательство. Прямое сравнение уравнений

$$\begin{aligned} \partial_0 x^\mu &= [x^\mu, p_0], \\ \partial_0 p_\mu &= [p_\mu, p_0]. \end{aligned} \quad \square$$

Число гамильтоновых уравнений движения, определяющих экстремаль, сократилось с $2n$ в (16.49), (16.50) до $2n - 2$ в (16.53). Это достигнуто за счет использования интеграла движения (16.51) и выбора специального параметра эволюции x^0 .

Продолжим анализ гамильтоновой формы уравнений для экстремалей. Функция действия $S_m(x, t)$ (21.38) удовлетворяет уравнению Гамильтона–Якоби (21.40)

$$\frac{\partial S_m}{\partial t} - \frac{1}{2m} g^{\alpha\beta} \frac{\partial S_m}{\partial x^\alpha} \frac{\partial S_m}{\partial x^\beta} = 0. \quad (16.54)$$

Поскольку гамильтониан (16.48) не зависит от параметра t явно, то функция действия имеет вид

$$S_m(x^\alpha, t) = \frac{mC_0}{2}t + W_m(x^\alpha), \quad C_0 = \text{const},$$

где укороченная функция действия W_m удовлетворяет укороченному уравнению Гамильтона–Якоби

$$g^{\alpha\beta} \frac{\partial W_m}{\partial x^\alpha} \frac{\partial W_m}{\partial x^\beta} = m^2 C_0. \quad (16.55)$$

Поскольку экстремали не зависят от массы пробной частицы, то, не ограничивая общности, можно положить $m = 1$ (включить в постоянную C_0). Так как $p_\alpha = \partial W_m / \partial x^\alpha$, то постоянная C_0 равна длине касательного вектора к экстремали

$$g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta = g^{\alpha\beta} p_\alpha p_\beta = C_0.$$

Для экстремалей в пространстве-времени знак постоянной C_0 определяет тип экстремали: постоянные $C_0 > 0$, $C_0 = 0$ и $C_0 < 0$ соответствуют времениподобным, светоподобным и пространственноподобным экстремалиям. Поскольку канонический параметр определен с точностью до линейных преобразований, то можно считать, что $C_0 = 1$, $C_0 = 0$ или $C_0 = -1$. Отметим, что для экстремалей нулевой длины $C_0 = 0$, и укороченное действие совпадает с полным $W_m = S_m$.

Для решения уравнения Гамильтона–Якоби часто применяют метод разделения переменных, который описан далее в разделе 21.13. Если уравнение Гамильтона–Якоби для частицы в (псевдо)римановой геометрии допускает полное разделение переменных, то такие пространства называются *штеккелевыми* [102, 103, 104, 105, 106]. К штеккелевым метрикам относятся метрика Шварцшильда (глава 34), Керра [107], де Ситтера (раздел 35.4.2), Казнера (раздел 35.5.3) и многие другие (практически все известные) метрики, полученные в результате решения уравнений Эйнштейна. Проблеме классификации штеккелевых пространств посвящены монографии [108, 109].

16.6 Волновое уравнение

Пусть задано многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, с метрикой лоренцевой сигнатуры, $\text{sign } g_{\alpha\beta} = (+ - \dots -)$. Рассмотрим волновое уравнение для скалярного поля $\varphi(x) \in \mathcal{C}^2(\mathbb{M})$:

$$g^{\alpha\beta} \tilde{\nabla}_\alpha \tilde{\nabla}_\beta \varphi = g^{\alpha\beta} \partial_\alpha \partial_\beta \varphi - g^{\alpha\beta} \tilde{\Gamma}_{\alpha\beta}^\gamma \partial_\gamma \varphi = 0, \quad (16.56)$$

или, эквивалентно,

$$\frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} g^{\alpha\beta} \partial_\beta \varphi) = 0, \quad g := \det g_{\alpha\beta},$$

где мы воспользовались тождеством (6.70). Это – линейное дифференциальное уравнение в частных производных второго порядка с переменными коэффициентами гиперболического типа, т.к. метрика имеет лоренцеву сигнатуру. Важным понятием в теории дифференциальных уравнений является характеристика (или характеристическая поверхность), которая для дифференциальных уравнений второго порядка определяется квадратичной формой $g^{\alpha\beta}$.

Определение. *Характеристикой* гиперболического дифференциального уравнения второго порядка (16.56) называется $\mathcal{C}^1$ гиперповерхность в многообразии $\mathbb{M}$, которая задается уравнением

$$W(x) = 0, \quad (16.57)$$

где функция $W \in \mathcal{C}^1(\mathbb{M})$ на поверхности $W = 0$ удовлетворяет условию

$$g^{\alpha\beta} \partial_\alpha W \partial_\beta W|_{W=0} = 0. \quad (16.58)$$

При этом требуется, чтобы по крайней мере одна из частных производных $\partial_\alpha W$ была отлична от нуля на гиперповерхности (16.57). $\square$

Замечание. Отметим, что в определении характеристики важна гиперболичность, т.к. при положительно или отрицательно определенной метрике уравнение (16.58) не имеет вещественных решений. $\square$

Уравнение (16.58) для характеристики совпадает с укороченным уравнением Гамильтона–Якоби для экстремалей (16.55) при $C_0 = 0$. Это значит, что характеристика соответствует укороченной функции действия для экстремалей нулевой длины. Напомним, что для экстремалей нулевой длины укороченная и полная функции действия совпадают. Однако условие (16.58) является более слабым, т.к. мы требуем выполнения (16.58) только на характеристике, а не во всем пространстве-времени.

Из определения характеристик следует, что они являются изотропными гиперповерхностями, которые будут рассмотрены в разделе 23.13.3. Метрика, индуцированная на таких поверхностях, по определению, вырождена, и все нормальные векторы изотропны (теорема 23.13.7).

Характеристики обладают следующим важным свойством. Допустим, что каждая гиперповерхность $W(x) - y^0 = 0$, где $-\epsilon < y^0 < \epsilon$, $\epsilon > 0$, есть характеристика уравнения (16.56). Другими словами, мы требуем, чтобы уравнение характеристик (16.58) выполнялось не только на самой характеристике, но и в некоторой ее окрестности. Поскольку на каждой характеристике по крайней мере одна из частных производных $\partial_\alpha W$ отлична от нуля, то это семейство заполняет некоторую достаточно малую область, через каждую точку которой проходит одна и только одна характеристика. Тогда можно перейти в новую систему координат $x^\alpha \rightarrow y^\alpha(x)$, где $y^0 = W$. При этом обратная метрика преобразуется по тензорному закону

$$g^{\alpha\beta}(y) = \frac{\partial y^\alpha}{\partial x^\gamma} \frac{\partial y^\beta}{\partial x^\delta} g^{\gamma\delta}(x).$$

Отсюда следует, что в новой системе координат $g^{00} = 0$. Это обстоятельство имеет важное следствие. Рассмотрим задачу Коши для волнового уравнения (16.56) в новой системе координат. Пусть на характеристике заданы начальные условия

$$\varphi|_{y^0=0} = \varphi_0, \quad \partial_0 \varphi|_{y^0=0} = \varphi_1,$$

где φ_0 и φ_1 – достаточно гладкие функции от “пространственных” координат $\{y^\mu\} = (y^1, \dots, y^{n-1})$. По начальным данным на характеристике при $y^0 = 0$ можно вычислить все частные производные по “пространственным” координатам $\partial_\mu \varphi, \partial_{\mu\nu}^2 \varphi, \dots$ и все частные производные с одной производной по “времени” $\partial_{0\mu}^2 \varphi, \partial_{0\mu\nu}^3 \varphi, \dots$. Для определения эволюции скалярного поля необходимо знать вторые производные по “времени” $\partial_{00}^2 \varphi$. Если задача Коши корректно поставлена, то вторые производные по “времени” находятся из волнового уравнения. На характеристике это не так, поскольку $g^{00} = 0$. Вместо определения второй производной $\partial_{00}^2 \varphi$ волновое уравнение накладывает ограничение (связь) на возможный выбор начальных условий:

$$g^{0\mu} \partial_\mu \varphi_1 - g^{\alpha\beta} \tilde{\Gamma}_{\alpha\beta}^0 \varphi_1 - g^{\alpha\beta} \tilde{\Gamma}_{\alpha\beta}^\mu \partial_\mu \varphi_0 = 0.$$

Для нахождения вторых производных $\partial_{00}^2 \varphi$ волновое уравнение нужно продифференцировать по y^0 . Однако после этой процедуры вторые производные не будут определены однозначно. Таким образом, при постановке задачи Коши на характеристике, во-первых, начальные данные нельзя задавать произвольно, и, во-вторых, волновое уравнение не определяет эволюцию поля единственным образом. Тем самым задача Коши на характеристике не допускает корректную постановку.

Выше мы взяли слова “пространственный” и “время” в кавычки, т.к. характеристика является изотропной поверхностью и не может быть пространственноподобной, а координата y^0 светоподобна и не может играть роль времени.

Поскольку вторые производные от неизвестной функции по одну сторону характеристики не определяются уравнением (16.56) и значениями функции φ по другую сторону от характеристики, то они могут иметь разрывы. Это значит, что решения уравнения (16.56) могут иметь разрывы производных, которые распространяются в пространстве вдоль характеристик.

Для корректной постановки задачи Коши для волнового уравнения (16.56) на многообразии с метрикой лоренцевой сигнатуры по x^0 координаты выбираются таким образом, что $g^{00} > 0$, а сечение $x^0 = 0$ является пространственноподобным. Поскольку пространственноподобное сечение не может быть изотропным, то оно не может быть также характеристикой. Оно не может также касаться характеристики, потому что в этом случае один из касательных векторов был бы изотропным, что противоречит отрицательной определенности метрики, индуцированной на гиперповерхности.

Замечание. В разделе 4.2 мы определили время x^0 как такую координату, что касательное векторное поле ∂_0 является времениподобным. Это соответствует условию $g_{00} > 0$. Отметим, что условия $g_{00} > 0$ и $g^{00} > 0$ неэквивалентны. Подробнее эти условия будут обсуждаться в разделе 24.2 при АДМ параметризации метрики. $\square$

Уравнение характеристик (16.58) представляет собой нелинейное дифференциальное уравнение в частных производных первого порядка. В общем случае оно очень сложно и, как правило, имеет решения с особыми точками. Характеристика, например, может быть не гладкой поверхностью.

Относительно частной производной по времени $\partial_0 W$ уравнение характеристик является алгебраическим квадратным уравнением, и имеет не более двух вещественных корней. Допустим, что метрика имеет лоренцеву сигнатуру и $g_{00} > 0$, тогда оно имеет два вещественных корня разных знаков

$$\partial_0 W = N^\mu \partial_\mu W \pm N \sqrt{-\hat{g}^{\mu\nu} \partial_\mu W \partial_\nu W}, \quad (16.59)$$

где $N(x)$ и $N^\mu(x)$ – функции хода и сдвига в АДМ параметризациях метрики (24.5). Здесь мы предполагаем, что все сечения $x^0 = \text{const}$ являются пространственноподобными и, следовательно, метрика $g_{\mu\nu}$ и ее обратная $\hat{g}^{\mu\nu}$ отрицательно определены. Обозначим $p_\mu := \partial_\mu W$. Введем функцию Гамильтона в фазовом пространстве $\{x^\mu, p_\mu\}$ (см. главу 21)

$$H(x^0, x^\mu, p_\mu) = -N^\mu p_\mu \mp N \sqrt{-\hat{g}^{\mu\nu} p_\mu p_\nu}, \quad (16.60)$$

где зависимость от координат x^0 и x^μ входит через метрику $g^{\alpha\beta} = g^{\alpha\beta}(x^0, x^\mu)$, и время x^0 рассматривается в качестве параметра эволюции.

Определение. Траектории $x^\mu(x^0)$ в конфигурационном пространстве, соответствующем гамильтониану (16.60), называются *бихарактеристиками* волнового уравнения (16.56). $\square$

Предложение 16.6.1. *Бихарактеристики волнового уравнения (16.56) совпадают с образами нулевых экстремалей для метрики $g_{\alpha\beta}$.*

Доказательство. Следствие предложения 16.5.1. $\square$

Единственное отличие бихарактеристик и нулевых экстремалей сводится к тому, что экстремали определяются в параметрическом виде $x^\alpha(t)$, а бихарактеристики задаются в виде явной зависимости координат $x^\mu(x^0)$.

Гамильтониан (16.60) совпадает с временной компонентой импульса (16.52). Функция $W(x^0, x^\mu)$, определяющая характеристику, является действием для бихарактеристик.

Замечание. Характеристики и бихарактеристики имеют физическую интерпретацию. Из теории дифференциальных уравнений известно, что если в момент времени x^0 в точке $\mathbf{x}$ произошло некоторое возмущение решения φ волнового уравнения (16.56), то оно будет распространяться в виде волны. При этом сечения $x^0 = \text{const}$ соответствующей характеристической поверхности (характеристического коноида с вершиной в точке $\{x^0, \mathbf{x}\}$) являются фронтом волны в момент времени x^0 , а бихарактеристики – это лучи, вдоль которых распространяется волна. $\square$

Предложение 16.6.2. *Если бихарактеристика касается характеристики в некоторой точке, то она целиком лежит на этой характеристике.*

Доказательство. Для экстремалей лагранжиан можно выбрать в виде

$$L = \frac{1}{2} g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta.$$

Для нулевых экстремалей он равен нулю, $L = 0$. Следовательно, функция действия $S(x, t)$ и укороченная функция действия $W = S + Et$ равны нулю, т.к. для нулевых экстремалей $E = 0$. Тем самым для бихарактеристик $W = 0$ и они целиком лежат на какой-либо из характеристик. $\square$

16.7 Приближение эйконала

Посмотрим на волновое уравнение (16.56) с другой точки зрения. Любое решение волнового уравнения (16.56) можно представить в виде¹

$$\varphi = A e^{iW/\hbar}, \quad (16.61)$$

где $A(x) \neq 0$ и $W(x)$ – амплитуда и фаза волны. Тогда волновое уравнение примет вид

$$g^{\alpha\beta} \frac{\partial_{\alpha\beta}^2 A}{A} + \frac{1}{\hbar} 2i g^{\alpha\beta} \frac{\partial_\alpha A}{A} \partial_\beta W + \frac{1}{\hbar} i g^{\alpha\beta} \partial_{\alpha\beta}^2 W - \\ - \frac{1}{\hbar^2} g^{\alpha\beta} \partial_\alpha W \partial_\beta W - g^{\alpha\beta} \Gamma_{\alpha\beta}^\gamma \left(\frac{\partial_\gamma A}{A} + \frac{1}{\hbar} i \partial_\gamma W \right) = 0.$$

В формальном пределе $\hbar \rightarrow 0$ волновое уравнение сводится к уравнению на фазу

$$g^{\alpha\beta} \partial_\alpha W \partial_\beta W = 0. \quad (16.62)$$

Оно совпадает с уравнением для характеристик (16.58), однако должно выполняться во всем пространстве-времени.

Определение. Для волновых решений предел $\hbar \rightarrow 0$ означает, что относительное изменение амплитуды мало, по сравнению с изменением фазы. Этот предел называют *приближением эйконала*. Фазу $W(x)$ называют *эйконалом*, а уравнение (16.62) – *уравнением эйконала*. В электродинамике это приближение называют также *приближением геометрической оптики*. Для плоских волн оно соответствует высоким частотам излучения и маленьким длинам волн. $\square$

¹В этом представлении $\hbar \in \mathbb{R}$ рассматривается, как вещественный параметр. Обозначение продиктовано аналогией с квантовой механикой, где $\hbar$ – постоянная Планка.

Параметр $\hbar$ был введен с единственной целью – определить эйкональное приближение. Поэтому в дальнейшем, для краткости, мы включим его в определение фазы $W/\hbar \mapsto W$.

Пример 16.7.1. Рассмотрим волновое уравнение в трехмерном пространстве-времени Минковского $\mathbb{R}^{1,2}$ в декартовой системе координат $(x^\alpha) = (x^0, x^1, x^2)$:

$$\partial_0^2 \varphi - \partial_1^2 \varphi + \partial_2^2 \varphi = 0.$$

В данном примере трехмерное пространство Минковского выбрано для наглядности. Все следующие формулы естественным образом переносятся на пространство Минковского произвольной размерности $\mathbb{R}^{1,n-1}$.

Уравнение характеристик имеет вид

$$((\partial_0 W)^2 - (\partial_1 W)^2 - (\partial_2 W)^2)|_{W=0} = 0. \quad (16.63)$$

Это уравнение допускает два семейства характеристик (изотропные поверхности в примере 23.13.2).

Первое семейство характеристик

$$W = (x^0 - x_0^0)^2 - (x^1 - x_0^1)^2 - (x^2 - x_0^2)^2 = 0,$$

где $x_0 = (x_0^0, x_0^1, x_0^2)$ – три произвольных вещественных параметра (координаты точки x_0). Нетрудно проверить, что

$$(\partial_0 W)^2 - (\partial_1 W)^2 - (\partial_2 W)^2 = 4W.$$

Это значит, что уравнение характеристик (16.63) выполняется только на характеристиках, а не во всем пространстве-времени. Для каждой точки x_0 характеристики первого семейства состоят из двух конусов (конуса прошлого, $x^0 < x_0^0$, и будущего, $x^0 > x_0^0$) с общей вершиной в точке x_0 .

Второе семейство характеристик

$$W = x^0 + k_\mu x^\mu - C, \quad \mu = 1, 2,$$

параметризуется единичным вектором в пространстве $|\mathbf{k}| := \sqrt{k_1^2 + k_2^2} = 1$ и постоянной C . В этом случае уравнение характеристик выполняется во всем пространстве, а не только на поверхности $W = 0$. Характеристики второго семейства параметризуются двумя параметрами, и представляют собой плоскости, перпендикулярные ковектору $(1, k_1, k_2)$ нулевой длины и пересекающие плоскость $x^0 = 0$ по прямой $k_\mu x^\mu = C$. Эти характеристики касаются характеристик первого семейства (конусов).

Напомним, что нулевыми экстремалими в пространстве Минковского являются прямые и только они с нулевым касательным вектором. Мы видим, что если произвольная нулевая экстремаль (или бихарактеристика) в некоторой точке касается характеристики, то она целиком принадлежит этой характеристике. Кроме того, через каждую регулярную точку характеристики (исключение составляют вершины конусов первого семейства) проходит одна и только одна нулевая экстремаль. Поэтому нулевые экстремали полностью заматают характеристические поверхности. Отметим, что одна и та же нулевая экстремаль может принадлежать разным характеристическим конусам.

Из уравнения для характеристик (16.63) можно найти производную по времени

$$\partial_0 W = \pm \sqrt{(\partial_1 W)^2 + (\partial_2 W)^2}.$$

Отсюда следует выражение для гамильтониана, определяющего бихарактеристики (траектории)

$$H = \pm \sqrt{p_1^2 + p_2^2} = \pm \hat{p}.$$

Уравнения движения для бихарактеристик имеют вид

$$\begin{aligned} \dot{x}^\mu &= \pm \frac{p^\mu}{\hat{p}}, \\ \dot{p}_\mu &= 0, \quad \Rightarrow \quad p_\mu = \text{const}, \end{aligned}$$

где точка обозначает дифференцирование по времени $t = x^0$. Отсюда следует, что импульс бихарактеристик постоянен, а траектории – это прямые $x^\mu = \pm n^\mu t + x_0^\mu$, проходящие через все точки пространства $x_0 \in \mathbb{R}^2$ во всех возможных направлениях $n^\mu := p^\mu / \hat{p}$. Поскольку пространственный вектор n^μ имеет единичную длину, то бихарактеристики в пространстве-времени совпадают с нулевыми экстремалами.

В пространстве Минковского уравнение (16.56) допускает решение в виде плоской волны

$$\varphi = A_0 e^{ik_\alpha x^\alpha} = A_0 e^{i(\omega t - \mathbf{k}\mathbf{x})}, \quad A_0 = \text{const} \neq 0,$$

где $\mathbf{k}\mathbf{x} = -k_\mu x^\mu$. Сравнение этого выражения с (16.61) дает выражения для частоты ω и волнового вектора $\mathbf{k}$ (в рассматриваемом примере мы, для краткости, полагаем $\hbar = 1$)

$$\omega = k_0 = \partial_0 W, \quad \mathbf{k} = \{\partial_1 W, \partial_2 W\}$$

Тогда волновое уравнение (16.58) сводится к соотношению между частотой и волновым вектором

$$g^{\alpha\beta} k_\alpha k_\beta = \omega^2 - \mathbf{k}^2 = 0,$$

где $\mathbf{k}^2 := -g^{\mu\nu} k_\mu k_\nu \geq 0$. Для плоской волны в пространстве Минковского эйконал $W(x) = x^\alpha k_\alpha$ является линейной функцией от координат, а частота и волновой вектор постоянны. $\square$

В рассмотренном примере плоских волн зависимость эйконала от координат была линейной. В общем случае эта зависимость является более сложной. Тогда частота и волновой вектор определяются соотношениями

$$\omega := \partial_0 W, \quad k_\mu := \partial_\mu W, \quad (16.64)$$

и зависят от точки пространства-времени. В этом случае уравнение эйконала (16.62) определяет зависимость частоты от волнового вектора

$$g^{00} \omega^2 + 2g^{0\mu} \omega k_\mu + g^{\mu\nu} k_\mu k_\nu = 0.$$

Эта зависимость называется *дисперсией*, а производная

$$v_g^\mu := \frac{\partial \omega}{\partial k_\mu}$$

называется *групповой скоростью*. Сравнение уравнения, определяющего дисперсию, с формулой (16.59) показывает, что частота ω и волновой вектор k_μ по сути дела совпадают, соответственно, с гамильтонианом и импульсами бихарактеристик волнового уравнения.

16.8 Гармонические координаты

Для исследования волнового уравнения (16.56) на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, с метрикой лоренцевой сигнатуры $g_{\alpha\beta}$ удобно использовать гармонические координаты. Как отмечено в разделе 1.5 функции перехода к новой системе координат являются скалярными полями на $\mathbb{M}$.

Определение. Рассмотрим волновое уравнение (16.56) и допустим, что в некоторой области $\mathcal{U} \subset \mathbb{M}$ оно имеет n функционально независимых решений (это так при достаточно общих предположениях). Пронумеруем эти решения φ^A , $A, \dots = 0, 1, \dots, n-1$. Тогда в области $\mathcal{U}$ решения волнового уравнения задают систему координат $x^A := \varphi^A$, которая называется *гармонической*. $\square$

Замечание. Гармоническая система координат была введена Т. де Дондером [110] и К. Ланцосом [111] и получила физическую интерпретацию в работах В. А. Фока [112, 101]. $\square$

Гармоническая система координат обладает следующим важным свойством.

Предложение 16.8.1. Система координат является гармонической тогда и только тогда, когда выполнены условия

$$\partial_A(\sqrt{|g|}g^{AB}) = 0 \quad \Leftrightarrow \quad \Xi^C := g^{AB}\Gamma_{AB}{}^C = 0, \quad (16.65)$$

где $g := \det g_{AB}$.

Доказательство. Справедливо тождество

$$\partial_\alpha \left[\det \left(\frac{\partial x}{\partial \varphi} \right) \frac{\partial \varphi^A}{\partial x^\beta} \right] \frac{\partial x^\alpha}{\partial \varphi^A} = 0,$$

где $\partial x / \partial \varphi$ – обратная матрица Якоби преобразования координат (1.68). Это тождество доказывается прямым дифференцированием с учетом правила дифференцирования определителя матрицы (37.10). С учетом доказанного равенства и правила преобразования координат справедлива следующая цепочка равенств:

$$\begin{aligned} \partial_A(\sqrt{|g|}g^{AB}) &= \frac{\partial x^\gamma}{\partial \varphi^A} \partial_\gamma \left[\sqrt{|\overset{\circ}{g}|} \det \left(\frac{\partial x}{\partial \varphi} \right) \frac{\partial \varphi^A}{\partial x^\alpha} \frac{\partial \varphi^B}{\partial x^\beta} g^{\alpha\beta} \right] = \\ &= \det \left(\frac{\partial x}{\partial \varphi} \right) \partial_\alpha \left(\sqrt{|\overset{\circ}{g}|} g^{\alpha\beta} \partial_\beta \varphi^B \right) = 0, \end{aligned}$$

где $\overset{\circ}{g} := \det g_{\alpha\beta}$. Таким образом, из гармоничности координат следует первое равенство (16.65) и наоборот. Эквивалентность равенств (16.65) между собой проверяется прямыми вычислениями. $\square$

Условия гармоничности координат можно записать с помощью принципа наименьшего действия. Пусть задано действие для n скалярных полей φ^A :

$$S = \int dx \sqrt{|g|} \frac{1}{2} g^{\alpha\beta} \partial_\alpha \varphi^A \partial_\beta \varphi^A \eta_{AB},$$

где η_{AB} – произвольная невырожденная симметричная матрица (метрика в пространстве-мишени). Тогда соответствующие уравнения Эйлера–Лагранжа примут вид

$$S_{,A} := \frac{\delta S}{\delta \varphi^A} = -\sqrt{|g|} g^{\alpha\beta} \tilde{\nabla}_\alpha \tilde{\nabla}_\beta \varphi^B \eta_{BA} = 0.$$

Эти уравнения эквивалентны волновому уравнению (16.56).

Уравнения на метрику (16.65) называются *условиями гармоничности*. В дальнейшем будем, как обычно, обозначать гармоническую систему координат снова греческими буквами x^α , предполагая, что условия гармоничности (16.65) выполнены.

Предложение 16.8.2. *В гармонической системе координат волновой оператор, действующий на функцию f , принимает вид*

$$\tilde{\square} f := g^{\alpha\beta} \tilde{\nabla}_\alpha \tilde{\nabla}_\beta f = g^{\alpha\beta} \partial_{\alpha\beta}^2 f \quad (16.66)$$

и не содержит первых частных производных $\partial_\alpha f$.

Доказательство. Прямая проверка. □

Пример 16.8.1. В пространстве-времени Минковского $\mathbb{R}^{1,n-1}$ декартова система координат является гармонической. □

Допустим, что на произвольном лоренцевом многообразии $\mathbb{M}$, которое топологически совпадает с $\mathbb{R}^{1,n-1}$, выбрана гармоническая система координат x^α . Тогда произвольная линейная комбинация координат также удовлетворяет волновому уравнению (16.66). Можно доказать, что все гармонические системы координат на $\mathbb{M}$, которые являются асимптотически декартовыми, связаны между собой преобразованиями Лоренца [101]. В. А. Фок придавал гармоническим системам координат в общей теории относительности физический смысл, считая, что на произвольном лоренцевом многообразии $\mathbb{M}$ они выделены и играют ту же роль, что и декартовы координаты в пространстве-времени Минковского.

Предложение 16.8.3. *В переменных Картана (см. раздел 5.4) условие гармоничности (16.65) имеет вид*

$$\tilde{\omega}^a = g^{\alpha\beta} \partial_\alpha e_\beta^a = 0.$$

Доказательство. Прямая проверка. □

16.9 Нормальные, геодезические или римановы координаты

Нормальные координаты, которые называются также геодезическими или римановыми играют большую роль при изучении свойств метрики и связности, а также в приложениях в математической физике. Сначала мы введем эту систему координат с помощью рядов, что является более наглядным и важным для приложений. Затем дадим определение с помощью экспоненциального отображения и сформулируем ряд утверждений, связанных с полнотой римановых многообразий.

Пример 16.9.1. В дальнейшем мы увидим, что декартовы координаты являются нормальными координатами в евклидовом пространстве $\mathbb{R}^n$. □

В определенном смысле нормальные координаты являются обобщением декартовой системы координат на общий случай многообразия с заданной аффинной геометрией.

16.9.1 Нормальные координаты. Локальное рассмотрение.

Пусть на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, задана аффинная геометрия $(\mathbb{M}, g, \Gamma)$. В настоящем разделе мы предполагаем, что все геометрические объекты (в том числе метрика и связность) являются вещественно аналитическими, т.е. их компоненты в окрестности каждой точки $x_0 \in \mathbb{M}$ представимы в виде сходящихся степенных рядов. Мы рассмотрим специальную систему координат в окрестности точки x_0 , которая является аналогом декартовой системы координат в евклидовом пространстве и имеет многочисленные приложения. В литературе эта система координат встречается под разными названиями: геодезические, римановы или нормальные координаты, и обычно строится в (псевдо)римановом пространстве. Мы построим такую систему координат в более общем случае произвольной аффинной геометрии.

Сначала мы рассмотрим геодезические линии вблизи точки x_0 . Пусть на многообразии задана кривая $x(t)$, проходящая через точку x_0 в заданном направлении,

$$x^\alpha(0) =: x_0^\alpha, \quad \dot{x}^\alpha(0) =: \dot{x}_0^\alpha. \quad (16.67)$$

Для того, чтобы кривая была геодезической, функции $x^\alpha(t)$ должны удовлетворять уравнениям для геодезических (16.3) с начальными условиями (16.67). При достаточно малых t будем искать решение уравнений для геодезических в виде степенного ряда

$$x^\alpha = x_0^\alpha + \dot{x}_0^\alpha t + \frac{1}{2} \ddot{x}_0^\alpha t^2 + \frac{1}{3!} \dddot{x}_0^\alpha t^3 + \dots \quad (16.68)$$

Первые два члена этого ряда определяются начальными условиями (16.67), а все последующие – уравнениями (16.3). Тем самым мы предполагаем, что вектор скорости геодезической в точке x_0 совпадает с заданным вектором $\dot{x}_0$. Отсюда вытекает, что канонический параметр вдоль геодезической однозначно фиксирован.

В нулевом порядке по t из уравнений для геодезических следует равенство

$$\ddot{x}_0^\alpha = -\overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha \dot{x}_0^\beta \dot{x}_0^\gamma, \quad \text{где} \quad \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha = \Gamma_{\beta\gamma}{}^\alpha(x_0).$$

Первый порядок уравнений по t определяет кубический член разложения в (16.68):

$$\dddot{x}_0^\alpha = (-\partial_\delta \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha + 2\overset{\circ}{\Gamma}_{\beta\gamma}{}^\epsilon \overset{\circ}{\Gamma}_{\{\delta\epsilon\}}{}^\alpha) \dot{x}_0^\beta \dot{x}_0^\gamma \dot{x}_0^\delta,$$

где, для краткости, мы допускаем некоторую вольность в обозначениях:

$$\partial_\delta \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha := \partial_\delta \Gamma_{\beta\gamma}{}^\alpha|_{x=x_0},$$

и фигурные скобки обозначают симметризацию по индексам.

Следовательно, решение задачи Коши для геодезических в третьем порядке по t имеет вид

$$x^\alpha = x_0^\alpha + \dot{x}_0^\alpha t - \frac{1}{2} \overset{\circ}{\Gamma}_{0\beta\gamma}{}^\alpha \dot{x}_0^\beta \dot{x}_0^\gamma t^2 - \frac{1}{6} \left(\partial_\delta \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha - 2\overset{\circ}{\Gamma}_{\beta\gamma}{}^\epsilon \overset{\circ}{\Gamma}_{\{\delta\epsilon\}}{}^\alpha \right) \dot{x}_0^\beta \dot{x}_0^\gamma \dot{x}_0^\delta t^3 + \dots \quad (16.69)$$

Коэффициенты при более высоких степенях t^k , $k \geq 4$ пропорциональны $(\dot{x}_0)^k t^k$ с коэффициентами, зависящими от аффинной связности и их частных производных вплоть до $k-2$ порядка, вычисленными в точке x_0 . Интервал сходимости ряда (16.69) определяется начальными данными и компонентами аффинной связности. Мы, конечно, предполагаем, что он больше нуля.

Перейдем к определению нормальных координат в окрестности $U_0 \subset M$ точки x_0 . Пусть M – многообразие с заданной аффинной связностью. Предположим, на время, что кручение равно нулю, т.е. $\Gamma_{\beta\gamma}^\alpha = \Gamma_{\gamma\beta}^\alpha$. Перепишем закон преобразования компонент аффинной связности (6.5) в виде

$$\Gamma_{\beta'\gamma'}^{\alpha'} = \partial_{\beta'} x^\beta \partial_{\gamma'} x^\gamma (\Gamma_{\beta\gamma}^\alpha \partial_\alpha x^{\alpha'} - \partial_{\beta\gamma}^2 x^{\alpha'}).$$

Совершим преобразование координат, которое задается квадратичным полиномом с постоянными коэффициентами:

$$x^{\alpha'} := B_\alpha^{\alpha'} (x^\alpha - x_0^\alpha) + \frac{1}{2} \overset{\circ}{\Gamma}_{\beta\gamma}^\alpha B_\alpha^{\alpha'} (x^\beta - x_0^\beta)(x^\gamma - x_0^\gamma), \quad (16.70)$$

где $B_\alpha^{\alpha'}$ – произвольная невырожденная постоянная матрица. Тогда в новой системе координат компоненты аффинной связности будут равны

$$\Gamma_{\beta'\gamma'}^{\alpha'} = \partial_{\beta'} x^\beta \partial_{\gamma'} x^\gamma \left[\Gamma_{\beta\gamma}^\alpha \left(B_\alpha^{\alpha'} + \overset{\circ}{\Gamma}_{\alpha\delta}^\epsilon B_\epsilon^{\alpha'} (x^\delta - x_0^\delta) \right) - \overset{\circ}{\Gamma}_{\beta\gamma}^\alpha B_\alpha^{\alpha'} \right].$$

Отсюда следует, что в точке x_0 компоненты аффинной связности без кручения обращаются в нуль, т.к. обращается в нуль выражение в квадратных скобках.

Замечание. Если связность обладает кручением, то, поскольку кручение является тензором, его нельзя обратить в нуль никаким преобразованием координат даже в одной точке. $\square$

Выше мы доказали, что для произвольной аффинной связности в произвольной заданной точке x_0 всегда можно обратить в нуль симметричную часть связности. Для этого достаточно выбрать соответствующим образом только квадратичные члены в функциях преобразований координат (16.70). Данная система координат существует в окрестности произвольной точки $x_0 \in M$ и определена неоднозначно. Во-первых, матрица $B_\alpha^{\alpha'}$ в (16.70) является невырожденной, а в остальном произвольна. Во-вторых, к правилу преобразования координат (16.70) можно добавить произвольные слагаемые третьей и более высокой степени по $(x - x_0)$. Этот произвол в выборе системы координат можно использовать для дальнейшей специализации системы координат.

Геометрический смысл построенной системы координат состоит в том, что в достаточно малой окрестности точки x_0 свойства многообразия близки к свойствам аффинного пространства, т.к. при параллельном переносе компоненты тензоров в линейном приближении по вектору смещения не меняются. Как и декартовы координаты в евклидовом пространстве, данные координаты в точке x_0 определены, в частности, с точностью до линейных преобразований.

Можно доказать, что симметричную часть компонент аффинной связности можно обратить в нуль не только в фиксированной точке, но и вдоль произвольной кривой γ без самопересечений на многообразии M [113] (см. также [15], глава VII, §91). Соответствующая система координат называется *геодезической вдоль кривой $\gamma \in M$* .

Оставшийся произвол в выборе системы координат можно использовать для дальнейшей специализации геометрических объектов. Воспользуемся свободой добавления высших степеней $(x - x_0)^k$, $k \geq 3$, в закон преобразования координат (16.70) и покажем, что в произвольной точке $x_0 \in M$ можно обратить в нуль не только симметричные части самих компонент аффинной связности, но и все их полностью симметризованные частные производные.

Теорема 16.9.1. *Если компоненты аффинной связности вещественно аналитичны, то в окрестности произвольной точки $x_0 \in \mathbb{M}$ существует такая система координат, что выполнены равенства:*

$$\overset{\circ}{\Gamma}_{\{\gamma_1\gamma_2\}}^\alpha = 0, \quad \partial_{\{\gamma_3\}} \overset{\circ}{\Gamma}_{\gamma_1\gamma_2}^\alpha = 0, \quad \dots \quad \partial_{\{\gamma_3 \dots \gamma_k\}}^{k-2} \overset{\circ}{\Gamma}_{\gamma_1\gamma_2}^\alpha = 0, \quad (16.71)$$

где фигурные скобки обозначают симметризацию по всем индексам, заключенным между ними.

Доказательство. Перепишем закон преобразования компонент аффинной связности (5.32) в виде

$$\Gamma_{\beta'\gamma'}^{\alpha'} A_{\alpha'}^\alpha = A_{\beta'}^\beta A_{\gamma'}^\gamma \Gamma_{\beta\gamma}^\alpha + \partial_{\beta'} A_{\gamma'}^\alpha, \quad (16.72)$$

где матрица $A_{\alpha'}^\alpha(x) = J_{\alpha'}^{-1\alpha} = \partial_{\alpha'} x^\alpha$ – обратная матрица Якоби преобразования координат. Дифференцирование этого соотношения по $x^{\delta'}$ приводит к равенству

$$\begin{aligned} \partial_{\delta'} \Gamma_{\beta'\gamma'}^{\alpha'} A_{\alpha'}^\alpha + \Gamma_{\beta'\gamma'}^{\alpha'} \partial_{\delta'} A_{\alpha'}^\alpha &= \partial_{\delta'} A_{\beta'}^\beta A_{\gamma'}^\gamma \Gamma_{\beta\gamma}^\alpha + \\ &+ A_{\beta'}^\beta \partial_{\delta'} A_{\gamma'}^\gamma \Gamma_{\beta\gamma}^\alpha + A_{\beta'}^\beta A_{\gamma'}^\gamma \partial_{\delta'} \Gamma_{\beta\gamma}^\alpha + \partial_{\delta'\beta'}^2 A_{\gamma'}^\alpha. \end{aligned} \quad (16.73)$$

Последовательное дифференцирование полученного соотношения приводит к равенствам, содержащим старшие производные от компонент аффинной связности и обратной матрицы Якоби $A_{\alpha'}^\alpha$. Рассматривая эти тождества в точке x^0 , мы докажем возможность выбора системы координат, в которой выполнены равенства (16.71), по теории возмущений.

Совершим преобразование координат $x \mapsto y(x)$. Предположим, что вблизи точки $x_0 \in \mathbb{M}$ обратное преобразование задается степенным рядом

$$x^\alpha = x_0^\alpha + \left. \frac{\partial x^\alpha}{\partial y^\gamma} \right|_0 y^\gamma + \frac{1}{2} \left. \frac{\partial^2 x^\alpha}{\partial y^{\gamma_1} \partial y^{\gamma_2}} \right|_0 y^{\gamma_1} y^{\gamma_2} + \frac{1}{3!} \left. \frac{\partial^3 x^\alpha}{\partial y^{\gamma_1} \partial y^{\gamma_2} \partial y^{\gamma_3}} \right|_0 y^{\gamma_1} y^{\gamma_2} y^{\gamma_3} + \dots \quad (16.74)$$

Здесь для новой системы координат мы используем букву y , чтобы избежать штрихов у индексов. Выберем первые три члена разложения в виде

$$\begin{aligned} \left. \frac{\partial x^\alpha}{\partial y^\gamma} \right|_0 &:= \delta_\gamma^\alpha, \\ \left. \frac{\partial^2 x^\alpha}{\partial y^{\gamma_1} \partial y^{\gamma_2}} \right|_0 &:= -\overset{\circ}{\Gamma}_{\{\gamma_1\gamma_2\}}^\alpha, \\ \left. \frac{\partial^3 x^\alpha}{\partial y^{\gamma_1} \partial y^{\gamma_2} \partial y^{\gamma_3}} \right|_0 &:= 2\overset{\circ}{\Gamma}_{\{\gamma_1\gamma_2\}}^\delta \overset{\circ}{\Gamma}_{\delta\gamma_3}^\alpha - \partial_{\{\gamma_3\}} \overset{\circ}{\Gamma}_{\gamma_1\gamma_2}^\alpha. \end{aligned}$$

Тогда из уравнений (16.72), (16.73) следует, что в новой системе координат

$$\overset{\circ}{\Gamma}_{\{\gamma_1\gamma_2\}}^\alpha = 0, \quad \partial_{\{\gamma_3\}} \overset{\circ}{\Gamma}_{\gamma_1\gamma_2}^\alpha = 0.$$

Выбор квадратичного члена ряда (16.74) уже был использован ранее. Выбор кубического члена разложения позволил обратить в нуль симметризованную первую частную производную аффинной связности в той же точке.

В дифференциальные тождества, получаемые последовательным дифференцированием (16.72), максимальная степень производной от обратной матрицы Якоби $A_{\alpha'}^\alpha$ всегда входит линейно и на единицу превышает максимальную степень производной

от компонент аффинной связности. Это значит, что коэффициенты ряда (16.74) всегда можно подобрать таким образом, чтобы обратить в нуль все симметризованные частные производные от коэффициентов аффинной связности (16.71). В этом нет ничего удивительного, т.к. члены ряда (16.74), начиная с квадратичного, находятся во взаимно однозначном соответствии с условиями на аффинную связность (16.71). $\square$

Нетрудно проверить, что ряд (16.74), определяющий преобразование координат, в точности совпадает с рядом (16.69), определяющим геодезические линии, если положить $y^\alpha = \dot{x}_0^\alpha t$. Это означает, что в новой системе координат прямые линии $y^\alpha = X_0^\alpha t$, где $t \in \mathbb{R}$, и X_0^α – произвольные числа, из которых по крайней мере одно отлично от нуля, являются геодезическими линиями на $\mathbb{M}$. В этом месте прослеживается аналогия с декартовой системой координат в евклидовом пространстве: геодезические являются прямыми линиями.

Построенная система координат является нормальной системой координат в некоторой достаточно малой окрестности $\mathbb{U}_0$ точки $x_0 \in \mathbb{M}$, определяется только симметричной частью компонент аффинной связности и никак от метрики не зависит. При этом у нас остался произвол в выборе первых двух членов разложения функций преобразования координат (16.74). Слагаемое нулевого порядка определим так, чтобы точка x_0 отображалась в начало координат $y_0 = (0, \dots, 0)$. Слагаемое первого порядка по y^α можно использовать для фиксирования значения метрики в начале координат. Очевидно, что его всегда можно подобрать таким образом, чтобы метрика в точке x_0 была диагональной

$$\overset{\circ}{g}_{\alpha\beta} = \eta_{\alpha\beta} := \text{diag} \left(\underbrace{+ \dots +}_p \underbrace{- \dots -}_q \right), \quad p + q = n. \quad (16.75)$$

Таким образом, мы однозначно фиксировали все члены разложения функций преобразования координат (16.74). Это доказывает следующее утверждение.

Теорема 16.9.2. Пусть в окрестности произвольной точки $x_0 \in \mathbb{M}$ многообразия, на котором задана аффинная геометрия, метрика и связность заданы вещественно аналитическими функциями. Тогда существует такая система координат y^α , что точка x_0 соответствует началу координат, а также выполнены равенства (16.71) и (16.75). Такая система координат определена однозначно с точностью до действия группы $\mathbb{O}(p, q)$ вращений координат y^α .

Замечание. В доказательстве теоремы нигде не использовалась сигнатура метрики. Поэтому сформулированная теорема справедлива как для римановых, так и для псевдоримановых метрик $\eta_{\alpha\beta}$. $\square$

Определение. Система координат y^α в теореме 16.9.2 называется *нормальной, геодезической* или *римановой*. $\square$

Предложение 16.9.1. Система координат y^α в некоторой окрестности начала координат является нормальной тогда и только тогда, когда выполнены условия:

$$\Gamma_{\beta\gamma}^\alpha(y) y^\beta y^\gamma = 0 \quad \text{и} \quad g_{\alpha\beta}(0) = \eta_{\alpha\beta}. \quad (16.76)$$

Доказательство. В нормальной системе координат геодезическая линия, проходящая через точку x_0 в направлении u_0 , является прямой и задается параметрически в виде

$$y^\alpha(t) = u_0^\alpha t. \quad (16.77)$$

Подстановка этого выражения в уравнение для геодезических линий (16.3) дает

$$u_0^\alpha u_0^\beta \Gamma_{\alpha\beta}{}^\gamma = 0.$$

Умножив это уравнение на t^2 , получим (16.76).

Обратно. Пусть выполнено уравнение (16.76). Разложим компоненты связности в ряд Тейлора вблизи начала координат. Тогда уравнение (16.76) эквивалентно цепочке равенств (16.71). Затем разложим уравнение геодезических и сами геодезические в ряды Тейлора по t . Приравняв нулю коэффициенты при одинаковых степенях t , получаем, что прямые линии и только они являются геодезическими линиями, проходящими через начало координат. $\square$

Замечание. Сравнение первого условия в (16.76) со вторым условием в (16.65) показывает, что нормальные координаты в общем случае не являются гармоническими. $\square$

Нормальная система координат обладает рядом замечательных свойств. В частности, в окрестности начала координат компоненты произвольного тензорного поля можно представить в виде ряда, коэффициенты которого зависят от ковариантных производных этого поля и тензоров кривизны, кручения и их ковариантных производных, вычисленных в точке x_0 . Доказательство этого утверждения проводится конструктивно, путем явного построения соответствующего разложения.

Для доказательства нам понадобится предварительный результат.

Лемма 16.9.1. Пусть связность Γ на $\mathbb{M}$ вещественно аналитична. Тогда в окрестности начала нормальной системы координат все производные

$$\partial_{\{\gamma_1 \dots \gamma_{k-1}\}}^{k-1} \overset{\circ}{\Gamma}_{\gamma_k}{}^\alpha{}^\beta, \quad (16.78)$$

где, в отличие от (16.71), симметризация проводится только по части нижних индексов, выражаются через тензоры кривизны, кручения и их ковариантные производные.

Доказательство. Лемма доказывается прямыми, но громоздкими вычислениями. Поэтому мы ограничимся только первыми двумя слагаемыми.

Используя свойства (16.71), нетрудно доказать равенства

$$\begin{aligned} \partial_{\{\gamma_1} \overset{\circ}{\Gamma}_{\gamma_2}{}^\alpha{}^\beta &= \frac{1}{3} \left(\partial_{\{\gamma_1} \overset{\circ}{\Gamma}_{\alpha\gamma_2}{}^\beta - \partial_\alpha \overset{\circ}{\Gamma}_{\{\gamma_1\gamma_2\}}{}^\beta + 2\partial_{\{\gamma_1} \overset{\circ}{T}_{\gamma_2\}}{}^\alpha{}^\beta \right), \\ \partial_{\{\gamma_1\gamma_2}^2 \overset{\circ}{\Gamma}_{\gamma_3}{}^\alpha{}^\beta &= \frac{1}{3} \left(\partial_{\{\gamma_1\gamma_2}^2 \overset{\circ}{\Gamma}_{\alpha\gamma_3}{}^\beta - \partial_\alpha^2 \overset{\circ}{\Gamma}_{\gamma_2\gamma_3}{}^\beta + 2\partial_{\{\gamma_1\gamma_2}^2 \overset{\circ}{T}_{\gamma_3\}}{}^\alpha{}^\beta \right). \end{aligned} \quad (16.79)$$

Здесь и в дальнейшем фигурные скобки обозначают симметризацию *только* по индексам $\gamma_1, \gamma_2, \dots$.

В точке x_0 аффинная связность полностью определяется тензором кручения

$$\overset{\circ}{\Gamma}_{\gamma\alpha}{}^\beta = \frac{1}{2} \overset{\circ}{T}_{\gamma\alpha}{}^\beta. \quad (16.80)$$

Далее, прямые вычисления приводят к следующим равенствам

$$\begin{aligned}
\mathring{R}_{\alpha\{\gamma_1\gamma_2\}}{}^\beta &= \partial_\alpha \mathring{\Gamma}_{\{\gamma_1\gamma_2\}}{}^\beta - \partial_{\{\gamma_1\gamma_2\}} \mathring{\Gamma}_{\alpha\gamma_2}{}^\delta - \frac{1}{4} \mathring{T}_{\alpha\{\gamma_1}{}^\delta \mathring{T}_{\gamma_2\}}{}^\beta, \\
\nabla_{\{\gamma_1\gamma_2\}} \mathring{R}_{\alpha\gamma_2\gamma_3}{}^\beta &= \partial_\alpha^2 \mathring{\Gamma}_{\gamma_2\gamma_3}{}^\beta - \partial_{\{\gamma_1\gamma_2\}}^2 \mathring{\Gamma}_{\alpha\gamma_3}{}^\beta + \frac{2}{3} \mathring{R}_{\alpha\{\gamma_1\gamma_2\}}{}^\delta \mathring{T}_{\gamma_3\}}{}^\beta + \frac{2}{3} \mathring{T}_{\alpha\{\gamma_1}{}^\delta \mathring{R}_{\delta\gamma_2\gamma_3\}}{}^\beta + \\
&\quad + \frac{1}{6} \nabla_{\{\gamma_1\gamma_2\}} \mathring{T}_{\gamma_2\alpha}{}^\delta \mathring{T}_{\gamma_3\}}{}^\beta - \frac{1}{3} \mathring{T}_{\alpha\{\gamma_1}{}^\delta \nabla_{\gamma_2\}} \mathring{T}_{\gamma_3\}}{}^\beta, \\
\nabla_{\{\gamma_1\gamma_2\}} \mathring{T}_{\gamma_2\alpha}{}^\beta &= \partial_{\{\gamma_1\gamma_2\}} \mathring{T}_{\gamma_2\alpha}{}^\beta, \\
\nabla_{\{\gamma_1\gamma_2\}} \nabla_{\gamma_2\}} \mathring{T}_{\gamma_3\}}{}^\beta &= \partial_{\{\gamma_1\gamma_2\}}^2 \mathring{T}_{\gamma_3\}}{}^\beta + \frac{1}{6} \mathring{R}_{\alpha\{\gamma_1\gamma_2\}}{}^\delta \mathring{T}_{\gamma_3\}}{}^\beta + \frac{1}{6} \mathring{T}_{\alpha\{\gamma_1}{}^\delta \mathring{R}_{\delta\gamma_2\gamma_3\}}{}^\beta + \\
&\quad + \frac{1}{6} \nabla_{\{\gamma_1\gamma_2\}} \mathring{T}_{\gamma_2\alpha}{}^\delta \mathring{T}_{\gamma_3\}}{}^\beta + \frac{1}{6} \mathring{T}_{\alpha\{\gamma_1}{}^\delta \nabla_{\gamma_2\}} \mathring{T}_{\gamma_3\}}{}^\beta.
\end{aligned}$$

Используя полученные формулы, правые части (16.79) можно записать в ковариантном виде

$$\partial_{\{\gamma_1\gamma_2\}} \mathring{\Gamma}_{\gamma_2\alpha}{}^\beta = \frac{1}{3} \left(-\mathring{R}_{\alpha\{\gamma_1\gamma_2\}}{}^\beta + 2\nabla_{\{\gamma_1\gamma_2\}} \mathring{T}_{\gamma_2\alpha}{}^\beta - \frac{1}{4} \mathring{T}_{\alpha\{\gamma_1}{}^\delta \mathring{T}_{\gamma_2\}}{}^\beta \right), \quad (16.81)$$

$$\begin{aligned}
\partial_{\{\gamma_1\gamma_2\}}^2 \mathring{\Gamma}_{\gamma_3\alpha}{}^\beta &= \frac{1}{3} \left(-\nabla_{\{\gamma_1\gamma_2\}} \mathring{R}_{\alpha\gamma_2\gamma_3}{}^\beta + \frac{1}{3} \mathring{R}_{\alpha\{\gamma_1\gamma_2\}}{}^\delta \mathring{T}_{\gamma_3\}}{}^\beta + \frac{1}{3} \mathring{T}_{\alpha\{\gamma_1}{}^\delta \mathring{R}_{\delta\gamma_2\gamma_3\}}{}^\beta + \right. \\
&\quad \left. + 2\nabla_{\{\gamma_1\gamma_2\}} \nabla_{\gamma_2\}} \mathring{T}_{\gamma_3\}}{}^\beta - \frac{1}{6} \nabla_{\{\gamma_1\gamma_2\}} \mathring{T}_{\gamma_2\alpha}{}^\delta \mathring{T}_{\gamma_3\}}{}^\beta - \frac{2}{3} \mathring{T}_{\alpha\{\gamma_1}{}^\delta \nabla_{\gamma_2\}} \mathring{T}_{\gamma_3\}}{}^\beta \right), \quad (16.82)
\end{aligned}$$

где, напомним, симметризация проводится только по индексам $\gamma_1, \gamma_2, \dots$. Аналогичным образом все частные производные от связности вида (16.78) можно выразить через ковариантные объекты. $\square$

Теперь можно доказать утверждение о разложении компонент произвольного тензорного поля в окрестности начала нормальной системы координат в ряд Тейлора. Для определенности рассмотрим произвольный ковариантный тензор второго ранга $A_{\alpha\beta}$ в нормальных координатах и разложим его в ряд Тейлора вблизи начала координат:

$$A_{\alpha\beta}(y) = \mathring{A}_{\alpha\beta} + \partial_\gamma \mathring{A}_{\alpha\beta} y^\gamma + \frac{1}{2} \partial_{\gamma_1\gamma_2}^2 \mathring{A}_{\alpha\beta} y^{\gamma_1} y^{\gamma_2} + \dots \quad (16.83)$$

Эта процедура явно не ковариантна, потому что координаты y^γ не являются компонентами вектора, и коэффициенты этого ряда также не ковариантны. Основным достоинством нормальных координат является то, что все коэффициенты полученного ряда тем не менее можно выразить через ковариантные величины. Для этого необходимо выразить все частные производные через ковариантные и выразить компоненты аффинной связности в начале координат через тензор кривизны, кручения и их ковариантные производные. Начнем с первой производной

$$\partial_\gamma \mathring{A}_{\alpha\beta} = \nabla_\gamma \mathring{A}_{\alpha\beta} + \mathring{\Gamma}_{\gamma\alpha}{}^\delta \mathring{A}_{\delta\beta} + \mathring{\Gamma}_{\gamma\beta}{}^\delta \mathring{A}_{\alpha\delta} = \nabla_\gamma \mathring{A}_{\alpha\beta} + \frac{1}{2} \mathring{T}_{\gamma\alpha}{}^\delta \mathring{A}_{\delta\beta} + \frac{1}{2} \mathring{T}_{\gamma\beta}{}^\delta \mathring{A}_{\alpha\delta},$$

где мы воспользовались формулами (16.80). Прямые, но более громоздкие выкладки

позволяют представить вторые частные производные также в ковариантном виде:

$$\begin{aligned} \partial_{\gamma_1 \gamma_2}^2 \overset{\circ}{A}_{\alpha\beta} = & \nabla_{\{\gamma_1} \nabla_{\gamma_2\}} \overset{\circ}{A}_{\alpha\beta} + \overset{\circ}{T}_{\{\gamma_1 \alpha}{}^{\delta} \nabla_{\gamma_2\}} \overset{\circ}{A}_{\delta\beta} + \overset{\circ}{T}_{\{\gamma_1 \beta}{}^{\delta} \nabla_{\gamma_2\}} \overset{\circ}{A}_{\alpha\delta} + \frac{1}{2} \overset{\circ}{T}_{\{\gamma_1 \alpha}{}^{\delta} \overset{\circ}{T}_{\gamma_2\}}{}^{\epsilon} \overset{\circ}{A}_{\delta\epsilon} + \\ & + \frac{1}{3} \left(-\overset{\circ}{R}_{\alpha\{\gamma_1 \gamma_2\}}{}^{\delta} + 2\nabla_{\{\gamma_1} \overset{\circ}{T}_{\gamma_2\}}{}^{\alpha}{}^{\delta} - \overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\epsilon} \overset{\circ}{T}_{\gamma_2\}}{}^{\delta}{}^{\epsilon} \right) \overset{\circ}{A}_{\delta\beta} + \\ & + \frac{1}{3} \left(-\overset{\circ}{R}_{\beta\{\gamma_1 \gamma_2\}}{}^{\delta} + 2\nabla_{\{\gamma_1} \overset{\circ}{T}_{\gamma_2\}}{}^{\beta}{}^{\delta} - \overset{\circ}{T}_{\beta\{\gamma_1}{}^{\epsilon} \overset{\circ}{T}_{\gamma_2\}}{}^{\delta}{}^{\epsilon} \right) \overset{\circ}{A}_{\alpha\delta}. \end{aligned}$$

Эту процедуру можно продолжить вплоть до произвольного порядка. Однако уже в третьем порядке формулы настолько громоздки, что нет смысла их приводить в явном виде.

Ясно, что аналогичное представление справедливо для тензорного поля произвольного ранга и типа. Поэтому справедлива следующая

Теорема 16.9.3. *Компоненты произвольного вещественно аналитического тензорного поля в некоторой окрестности начала нормальной системы координат представимы в виде ряда Тейлора, коэффициенты которого определяются ковариантными производными компонент данного тензорного поля, а также тензором кривизны, тензором кручения и их ковариантными производными, взятыми в начале координат.*

В математической физике такое представление часто бывает очень удобным при проведении вычислений.

16.9.2 Нормальные координаты в (псевдо)римановом пространстве

Основное достоинство нормальных координат заключается в том, что вещественно аналитические тензорные поля в окрестности $\mathbb{U}_0 \subset \mathbb{M}$ произвольной точки $x_0 \in \mathbb{M}$ представляются в виде ряда Тейлора, коэффициенты которого задаются только ковариантными объектами: ковариантными производными данного тензорного поля, а также тензорами кривизны и кручения и их ковариантными производными. В предыдущем разделе были явно вычислены первые два члена этого ряда для ковариантного тензора второго ранга. Эти члены содержат много слагаемых с тензором кручения. Поэтому в (псевдо)римановой геометрии, где кручение тождественно равно нулю, формулы упрощаются, и можно продвинуться значительно дальше в вычислениях.

Кроме того, если на многообразии $\mathbb{M}$ задана аффинная геометрия $(\mathbb{M}, g, \Gamma)$, то в качестве определяющего пучка кривых, проходящих через точку $x_0 \in \mathbb{M}$ можно выбрать не геодезические линии, а экстремали, которые определяются исключительно метрикой. Это также дает основание рассмотреть нормальные координаты в (псевдо)римановой геометрии более подробно.

Два слова об обозначениях. Мы часто используем знак тильды для геометрических объектов в (псевдо)римановой геометрии. Поскольку значок окружности над символом уже используется для обозначения геометрических объектов, рассматриваемых в точке x_0 , то, чтобы не загромождать обозначений, знак тильды в настоящем разделе мы опустим.

В (псевдо)римановой геометрии в нормальных координатах $\overset{\circ}{\Gamma}_{\alpha\beta}{}^{\gamma} = 0$, и ковариантная производная произвольного тензора в этой точке совпадает с частной производной. При этом в выражении для тензора кривизны пропадают квадратичные

слагаемые по связности:

$$\overset{\circ}{R}_{\alpha\beta\gamma}{}^{\delta} = \partial_{\alpha}\overset{\circ}{\Gamma}_{\beta\gamma}{}^{\delta} - \partial_{\beta}\overset{\circ}{\Gamma}_{\alpha\gamma}{}^{\delta}$$

или

$$\overset{\circ}{R}_{\alpha\beta\gamma\delta} = \frac{1}{2}(\partial_{\alpha\gamma}^2\overset{\circ}{g}_{\beta\delta} - \partial_{\alpha\delta}^2\overset{\circ}{g}_{\beta\gamma} - \partial_{\beta\gamma}^2\overset{\circ}{g}_{\alpha\delta} + \partial_{\beta\delta}^2\overset{\circ}{g}_{\alpha\gamma}).$$

Вычисления, аналогичные тем, что привели к формулам (16.81), (16.82), в (псевдо)римановой геометрии дают равенства

$$\begin{aligned}\partial_{\{\gamma_1}\overset{\circ}{\Gamma}_{\gamma_2\}\alpha}{}^{\beta} &= -\frac{1}{3}\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2\}}{}^{\beta}, \\ \partial_{\{\gamma_1\gamma_2}\overset{\circ}{\Gamma}_{\gamma_3\}\alpha}{}^{\beta} &= -\frac{1}{2}\nabla_{\{\gamma_1}\overset{\circ}{R}_{\alpha\gamma_2\gamma_3\}}{}^{\beta}, \\ \partial_{\{\gamma_1\gamma_2\gamma_3}\overset{\circ}{\Gamma}_{\gamma_4\}\alpha}{}^{\beta} &= -\frac{3}{5}\nabla_{\{\gamma_1}\nabla_{\gamma_2}\overset{\circ}{R}_{\alpha\gamma_3\gamma_4\}}{}^{\beta} - \frac{2}{15}\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2}{}^{\delta}\overset{\circ}{R}_{\delta\gamma_3\gamma_4\}}{}^{\beta},\end{aligned}\tag{16.84}$$

где симметризация проводится только по индексам $\gamma_1, \gamma_2, \dots$. Здесь мы дополнительно вычислили третью симметризованную производную от связности.

Разложим ковариантный тензор $A_{\alpha_1\dots\alpha_s}$ произвольного ранга s в окрестности точки x_0 в ряд по нормальным координатам:

$$A_{\alpha_1\dots\alpha_s} = \overset{\circ}{A}_{\alpha_1\dots\alpha_s} + \partial_{\gamma}\overset{\circ}{A}_{\alpha_1\dots\alpha_s}y^{\gamma} + \frac{1}{2}\partial_{\gamma_1\gamma_2}^2\overset{\circ}{A}_{\alpha_1\dots\alpha_s}y^{\gamma_1}y^{\gamma_2} + \dots$$

Преобразовав частные производные в ковариантные и исключив слагаемые со связностью с помощью формул (16.84), получим следующий ряд

$$\begin{aligned}A_{\alpha_1\dots\alpha_s} &= \overset{\circ}{A}_{\alpha_1\dots\alpha_s} + \nabla_{\gamma}\overset{\circ}{A}_{\alpha_1\dots\alpha_s}y^{\gamma} + \\ &+ \frac{1}{2!}\left(\nabla_{\gamma_1}\nabla_{\gamma_2}\overset{\circ}{A}_{\alpha_1\dots\alpha_s} - \frac{1}{3}\sum_{k=1}^s\overset{\circ}{R}_{\alpha_k\gamma_1\gamma_2}{}^{\beta_k}\overset{\circ}{A}_{\alpha_1\dots\beta_k\dots\alpha_s}\right)y^{\gamma_1}y^{\gamma_2} + \\ &+ \frac{1}{3!}\left(\nabla_{\gamma_1}\nabla_{\gamma_2}\nabla_{\gamma_3}\overset{\circ}{A}_{\alpha_1\dots\alpha_s} - \sum_{k=1}^s\overset{\circ}{R}_{\alpha_k\gamma_1\gamma_2}{}^{\beta_k}\nabla_{\gamma_3}\overset{\circ}{A}_{\alpha_1\dots\beta_k\dots\alpha_s} - \right. \\ &\quad \left. - \frac{1}{2}\sum_{k=1}^s\nabla_{\gamma_1}\overset{\circ}{R}_{\alpha_k\gamma_2\gamma_3}{}^{\beta_k}\overset{\circ}{A}_{\alpha_1\dots\beta_k\dots\alpha_s}\right)y^{\gamma_1}y^{\gamma_2}y^{\gamma_3} + \dots\end{aligned}\tag{16.85}$$

Под знаком суммы подразумевается суммирование по некому индексу β_k , который стоит на k -том месте в $\overset{\circ}{A}_{\alpha_1\dots\beta_k\dots\alpha_s}$. В полученном выражении симметризацию по индексам $\gamma_1, \gamma_2, \dots$ можно не указывать, т.к. происходит свертка с симметричным произведением $y^{\gamma_1}\dots y^{\gamma_k}$.

Аналогичные ряды можно построить для тензоров, содержащих произвольные наборы ковариантных и контравариантных индексов. Во всех случаях коэффициенты ряда вплоть до любого порядка могут быть выражены только через ковариантные объекты в точке x_0 .

Нормальные координаты особенно удобны при анализе (псевдо)римановой метрики. В этом случае возникает дополнительное упрощение, поскольку все ковариантные производные от метрики тождественно равны нулю. Несложные вычисления

показывают, что ряд (16.85) для метрики в нормальных координатах принимает вид

$$g_{\alpha\beta} = \eta_{\alpha\beta} - \frac{1}{3} \overset{\circ}{R}_{\alpha\gamma_1\gamma_2\beta} y^{\gamma_1} y^{\gamma_2} - \frac{1}{3!} \nabla_{\gamma_1} \overset{\circ}{R}_{\alpha\gamma_2\gamma_3\beta} y^{\gamma_1} y^{\gamma_2} y^{\gamma_3} + \\ + \frac{1}{5!} \left(-6 \nabla_{\gamma_1} \nabla_{\gamma_2} \overset{\circ}{R}_{\alpha\gamma_3\gamma_4\beta} + \frac{16}{3} \overset{\circ}{R}_{\alpha\gamma_1\gamma_2}{}^{\delta} \overset{\circ}{R}_{\delta\gamma_3\gamma_4\beta} \right) y^{\gamma_1} y^{\gamma_2} y^{\gamma_3} y^{\gamma_4} + \dots, \quad (16.86)$$

где мы также вычислили слагаемое четвертого порядка.

16.9.3 (Псевдо)римановы пространства постоянной кривизны

Выражение для метрики в нормальных координатах (16.86) принимает особенно простой вид для пространств постоянной кривизны, которые определяются равенством $\nabla_{\epsilon} R_{\alpha\beta\gamma}{}^{\delta} = 0$.

Предложение 16.9.2. *В пространстве постоянной кривизны (псевдо)риманова метрика в окрестности произвольной точки $x_0 \in \mathbb{M}$ в нормальных координатах y^{α} имеет вид*

$$g_{\alpha\beta} = \eta_{\alpha\beta} + \frac{1}{2} \sum_{k=1}^{\infty} \frac{(-1)^k 2^{2k+2}}{(2k+2)!} V_{\alpha}{}^{\epsilon_1} V_{\epsilon_1}{}^{\epsilon_2} \dots V_{\epsilon_{k-1}}{}^{\epsilon_k} \eta_{\epsilon_k\beta}, \quad (16.87)$$

где

$$V_{\epsilon_{k-1}}{}^{\epsilon_k} := \overset{\circ}{R}_{\epsilon_{k-1}\gamma_1\gamma_2}{}^{\epsilon_k} y^{\gamma_1} y^{\gamma_2}, \quad \epsilon_0 = \alpha.$$

Доказательство. Сначала проверяем, что первые два члена суммы действительно совпадают с (16.86) для пространств постоянной кривизны. Далее доказательство проводится по индукции. $\square$

Ряд для метрики (16.87) можно просуммировать для пространств постоянной кривизны специального вида (8.50). В этом случае в начале координат справедливо равенство

$$\overset{\circ}{R}_{\alpha\beta\gamma\delta} = -K(\eta_{\alpha\gamma}\eta_{\beta\delta} - \eta_{\alpha\delta}\eta_{\beta\gamma}), \quad n := \dim \mathbb{M}, \quad K = \text{const}. \quad (16.88)$$

Тогда матрица $V_{\alpha}{}^{\beta}$ пропорциональна проекционному оператору:

$$V_{\alpha}{}^{\beta} = as \left(\delta_{\alpha}^{\beta} - \frac{y_{\alpha} y^{\beta}}{s} \right),$$

где

$$a := \frac{2K}{n(n-1)}, \quad s := y^{\alpha} y_{\alpha}, \quad y_{\alpha} := y^{\beta} \eta_{\alpha\beta},$$

и ряд для метрики принимает вид

$$g_{\alpha\beta} = \eta_{\alpha\beta} + \frac{1}{2} \left(\eta_{\alpha\beta} - \frac{y_{\alpha} y_{\beta}}{s} \right) \sum_{k=1}^{\infty} \frac{2^{2k+2}}{(2k+2)!} (-as)^k. \quad (16.89)$$

Это выражение для метрики определено и для $s = 0$, т.к. ряд начинается с линейного по s члена. Теперь ряд можно просуммировать, что приводит к следующему выражению для метрики

$$g_{\alpha\beta} = f(s) \Pi_{\alpha\beta}^T + \Pi_{\alpha\beta}^L = f \eta_{\alpha\beta} + (1-f) \frac{y_{\alpha} y_{\beta}}{s}, \quad (16.90)$$

где метрика представлена в виде суммы проекционных операторов

$$\Pi_{\alpha\beta}^T := \eta_{\alpha\beta} - \frac{y_\alpha y_\beta}{s}, \quad \Pi_{\alpha\beta}^L := \frac{y_\alpha y_\beta}{s},$$

а функция $f(s)$ определена рядом

$$f(s) := \sum_{k=0}^{\infty} \frac{2^{2k+2}}{(2k+2)!} (-as)^k, \quad (16.91)$$

который сходится на всей комплексной плоскости s . Тем самым доказано следующее утверждение.

Теорема 16.9.4. Пусть на многообразии задана (псевдо)риманова метрика g класса C^2 такая, что многообразие является пространством постоянной кривизны специального вида:

$$R_{\alpha\beta\gamma\delta} = -K(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}). \quad (16.92)$$

Тогда метрика g на $\mathbb{M}$ вещественно аналитична.

Доказательство. Прямая проверка показывает, что вещественно аналитическая метрика (16.90), (16.91) описывает многообразие постоянной кривизны (16.92). Единственность метрики следует из единственности решения задачи Коши для уравнений геодезических. $\square$

Ряд (16.91) можно просуммировать. Для псевдоримановых пространств положительной кривизны $K > 0, a > 0$, и функция $f(s)$ имеет вид

$$f(s) = \begin{cases} \frac{\sin^2 \sqrt{as}}{as}, & s > 0, \\ 1, & s = 0, \\ -\frac{\operatorname{sh}^2 \sqrt{-as}}{as}, & s < 0. \end{cases} \quad (16.93)$$

Функция $f(s)$, как нетрудно проверить, аналитична. Напомним, что $\sin ix = i \operatorname{sh} x$.

Для псевдоримановых пространств отрицательной кривизны $K < 0, a < 0$ имеем равенство

$$f(s) = \begin{cases} -\frac{\operatorname{sh}^2 \sqrt{-as}}{as}, & s > 0, \\ 1, & s = 0, \\ \frac{\sin^2 \sqrt{as}}{as}, & s < 0. \end{cases} \quad (16.94)$$

Отметим, что на световом конусе, $s = 0$, метрика в обоих случаях совпадает с метрикой Минковского.

Если кривизна псевдориманова пространства постоянной кривизны равна нулю, $K = 0$, то $f = 1$.

Для римановых пространств постоянной кривизны всегда $s > 0$, и

$$f(s) = \begin{cases} \frac{\sin^2 \sqrt{as}}{as}, & K > 0, \\ 1, & K = 0, \\ -\frac{\operatorname{sh}^2 \sqrt{-as}}{as}, & K < 0. \end{cases} \quad (16.95)$$

Тот факт, что функция $f(s)$ действительно приводит к ряду (16.89) проверяется прямой проверкой. Не зная функции $f(s)$, ряд (16.89) можно просуммировать следующим образом. Представление (16.89) задает тензорную структуру метрики. В разделе 15.6 для метрики более общего вида, которая параметризуется двумя произвольными функциями $f(s)$ и $g(s)$, были вычислены символы Кристоффеля и тензор кривизны. Рассматриваемому случаю соответствует условие

$$g = \frac{(f + f's)^2}{f - afs} = 1,$$

где штрих обозначает дифференцирование по s . Последнее равенство задает обыкновенное дифференциальное уравнение первого порядка на $f(s)$. Это уравнение легко решается после подстановки

$$U = fs \quad (16.96)$$

Постоянная интегрирования находится из условия $U(0) = 0$ (ограниченность метрики).

Рассмотрим римановы пространства постоянной кривизны, для которых $g_{\alpha\beta}(0) = \delta_{\alpha\beta}$. Метрика (16.90) принимает особенно простой вид в сферической системе координат евклидова пространства $\mathbb{R}^n$. В сферических координатах $s = y^\alpha y^\beta \delta_{\alpha\beta} = r^2$, где r – радиальная координата, и справедливы тождества

$$\Pi_{\alpha\beta}^L dy^\alpha dy^\beta = dr^2, \quad \Pi_{\alpha\beta}^T dy^\alpha dy^\beta = r^2 d\Omega,$$

где $d\Omega$ – элемент телесного угла в евклидовом пространстве $\mathbb{R}^n$. Таким образом метрику пространств постоянной кривизны в нормальных координатах можно записать в виде

$$ds^2 = dr^2 + r^2 f d\Omega. \quad (16.97)$$

Проанализируем пространство $\mathbb{R}^n$ с метрикой (16.97) подробнее. Объем сферы радиуса $r\sqrt{f}$ в $\mathbb{R}^n$ с метрикой (16.97) равен

$$S_r^{n-1} = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})} (r^2 f)^{(n-1)/2}.$$

Мы видим, что нули функции $f(r^2) = 0$ определяют те сферы в $\mathbb{R}^n$, площадь которых равна нулю. Поскольку площадь поверхности является инвариантным объектом, то это означает, что на самом деле эти сферы соответствуют отдельным точкам пространства постоянной кривизны.

Нормальные координаты для многообразий постоянной кривизны определены для всех $\{y^\alpha\} \in \mathbb{R}^n$. При этом все геодезические, проходящие через точку x_0 , полны, т.к. канонический параметр пробегает всю вещественную прямую $t \in (-\infty, \infty)$.

Проведенное рассмотрение доказывает

Теорема 16.9.5. *Нормальные координаты на (псевдо)римановом многообразии $\mathbb{M}$ постоянной кривизны вида (16.92) в каждой точке $x_0 \in \mathbb{M}$ задают гладкое сюръективное отображение*

$$\mathbb{R}^n \rightarrow \mathbb{M}. \quad (16.98)$$

Для римановых пространств нулевой кривизны нормальные координаты совпадают с декартовыми.

Замечание. В дальнейшем мы увидим, что отображение (16.98) не является накрытием. $\square$

В общем случае отображение (16.98) не является взаимно однозначным. Поэтому в области определения нормальных координат $\mathbb{R}^n$ можно задать отношение эквивалентности, отождествив те точки, которые отображаются на одну и ту же точку из $\mathbb{M}$. Таким образом, пространство постоянной кривизны вида (16.92) можно рассматривать как евклидово пространство $\mathbb{R}^n$, в котором задано некоторое отношение эквивалентности между точками.

Пример 16.9.2. Рассмотрим двумерную сферу $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ единичного радиуса в качестве пространства постоянной положительной кривизны. В этом случае $K = 1, a = 1, n = 2$. Функция $f(r^2)$ в полярных координатах на плоскости $\mathbb{R}^2$ имеет вид

$$f = \frac{\sin^2 r}{r^2}.$$

Это соответствует метрике

$$ds^2 = dr^2 + \sin^2 r d\varphi^2.$$

Длина окружности на плоскости $\mathbb{R}^2$ радиуса r с центром в начале координат равна

$$L = \int_0^{2\pi} d\varphi \sin r = 2\pi \sin r.$$

Отсюда следует, что окружности радиусов $r = \pi k, k = 1, 2, \dots$ отображаются в одну точку сферы. При этом плоскость $\mathbb{R}^2$ бесконечное число раз “накрывает” сферу $\mathbb{S}^2$. Если условиться, что начало координат соответствует южному полюсу сферы, то при отображении $\mathbb{R}^2 \rightarrow \mathbb{S}^2$ все окружности радиуса $r = 2\pi k, k = 0, 1, \dots$, соответствуют южному полюсу, а все окружности радиуса $r = \pi + 2\pi k$ – северному. В рассматриваемом случае между точками евклидовой плоскости возникает отношение эквивалентности

$$y^\alpha \sim y^\alpha + \frac{y^\alpha}{r} 2\pi k, \quad k = 0, 1, \dots \quad \square$$

Нормальные координаты были определены таким образом, что геодезические линии в них совпадают с прямыми. В (псевдо)римановом пространстве экстремали совпадают с геодезическими и поэтому также являются прямыми. Проверим это для пространств постоянной кривизны, которые были рассмотрены выше. Уравнения для экстремалей, определяемых метрикой (16.90), можно проинтегрировать. Из выражения для символов Кристоффеля (15.43), которые в рассматриваемом случае имеют вид

$$\Gamma_{\alpha\beta}^\gamma = \frac{f'}{f} (y_\alpha \Pi_\beta^{\tau\gamma} + y_\beta \Pi_\alpha^{\tau\gamma}) + \frac{1 - f - f's}{s} y^\gamma \Pi_{\alpha\beta}^\tau, \quad (16.99)$$

следуют уравнения для экстремалей (16.21)

$$\ddot{y}^\alpha = -2\frac{f'}{f} \dot{y}^\alpha y_\beta \dot{y}^\beta + 2\frac{f'}{f} y^\alpha \frac{(y_\beta \dot{y}^\beta)^2}{s} + \frac{1 - f - f's}{s^2} y^\alpha (y_\beta \dot{y}^\beta)^2 - \frac{1 - f - f's}{s} y^\alpha (\dot{y}_\beta \dot{y}^\beta). \quad (16.100)$$

Эти уравнения имеют интеграл (16.29)

$$C_0 = \dot{y}^\alpha \dot{y}^\beta g_{\alpha\beta} = \dot{y}_\alpha \dot{y}^\alpha f - \frac{(y_\alpha \dot{y}^\alpha)^2}{s} (f - 1) = \text{const.}$$

Прямая проверка показывает, что все прямые линии, проходящие через начало координат,

$$y^\alpha = v^\alpha t, \quad v^\alpha = \text{const}, \quad t \in \mathbb{R},$$

являются экстремалими. Эти экстремали, очевидно, полны. Отметим, что символы Кристоффеля (16.99) равны нулю только в начале координат. В близких точках они отличны от нуля, и среди экстремалей прямыми являются только те, которые проходят через начало координат.

Уравнения для экстремалей (16.100) можно проинтегрировать и в общем случае. То есть найти те экстремали, которые не проходят через начало координат. Для этой цели рассмотрим зависимость $s := y^\alpha y^\beta \eta_{\alpha\beta}$ от t . Учитывая равенства

$$\begin{aligned}\dot{s} &= 2y_\alpha \dot{y}^\alpha, \\ \ddot{s} &= 2y_\alpha \ddot{y}^\alpha + 2\dot{y}_\alpha \dot{y}^\alpha,\end{aligned}$$

из уравнений для экстремалей получаем обыкновенное дифференциальное уравнение на $s(t)$:

$$\ddot{s} = \left(\frac{1}{2s} - \frac{U'}{2U} \right) \dot{s}^2 + \frac{2C_0 U' s}{U},$$

где функция $U(s)$ определена равенством (16.96)

$$U(s) = \frac{\sin^2 \sqrt{as}}{a}.$$

Рассмотрим случай $C_0 a > 0$. Введя новую переменную $z^2 := as$, при $as > 0$, и растянув канонический параметр $t \rightarrow t/\sqrt{C_0 a}$, приходим к уравнению

$$\ddot{z} = (1 - \dot{z}^2) \operatorname{ctg} z,$$

которое можно явно проинтегрировать. Общее решение этого уравнения имеет вид

$$\cos \sqrt{as} = \sqrt{1 - \frac{1}{C_1}} \sin(t + t_0), \quad C_1 = \text{const} > 1, \quad t_0 = \text{const},$$

Аналогично можно рассмотреть все остальные случаи.

Таким образом, для пространств постоянной кривизны специального вида (16.88) в нормальных координатах можно найти и проанализировать поведение всех экстремалей. В дальнейшем мы рассмотрим эту задачу в ряде конкретных случаев.

16.9.4 Нормальные координаты и экспоненциальное отображение

Пусть задана аффинная связность Γ на многообразии $\mathbb{M}$ и $\gamma = x(t)$ – геодезическая. Если t – канонический параметр вдоль геодезической, то координатные функции $\{x^\alpha(t)\}$ удовлетворяют системе уравнений (16.3). Зафиксируем канонический параметр t каким либо образом и начальную точку $x_0 = x(0)$. Тогда касательный вектор скорости $u_0 := \dot{x}_0 \in \mathbb{T}_0(\mathbb{M})$ к геодезической в точке x_0 определен однозначно. Верно также обратное утверждение. Если задана точка $x_0 \in \mathbb{M}$ и касательный вектор $u_0 \in \mathbb{T}_0(\mathbb{M})$, то существует единственная геодезическая, проходящая через x_0 с начальным вектором скорости u_0 , при этом канонический параметр определен с точностью до сдвига. Таким образом, каждая геодезическая вместе с каноническим параметром однозначно определена парой (x_0, u_0) .

В разделе 2.6.5 было определено экспоненциальное отображение для гладких полных векторных полей на многообразии. Геодезическая линия $x(t)$ является интегральной кривой для векторного поля скорости $u(t) := \dot{x}(t)$. Тогда для полных геодезических определено экспоненциальное отображение

$$\exp(tu_0) : x_0 \mapsto x(t), \quad (16.101)$$

которое отображает начальную точку x_0 в точку $x(t)$. Если вектор скорости u_0 принимает все возможные направления в касательном пространстве $T_0(M)$, то экспоненциальное отображение (16.101) можно рассматривать, как отображение касательного пространства в многообразие

$$\exp(tu_0) : T_0(M) \ni tu_0 \mapsto x(t) \in M, \quad (16.102)$$

для которого мы сохраним прежнее обозначение.

Замечание. В разделе 2.6.5 экспоненциальное отображение было определено для каждого полного векторного поля. При этом вопрос о том, каким образом данное векторное поле задано, не рассматривался. Если векторное поле одно, то экспоненциальное отображение нельзя рассматривать, как отображение касательного пространства (16.102), т.к. в точке x_0 имеется только один вектор. В рассматриваемом случае ситуация другая. На многообразии M задана связность Γ , которая определяет все множество геодезических. Для фиксированной точки x_0 мы рассматриваем множество геодезических, проходящих через данную точку во всех возможных направлениях. То есть вектор скорости u_0 в (16.102) принимает все возможные направления. Следовательно, экспоненциальное отображение можно рассматривать, как отображение касательного пространства. $\square$

По построению, каждая прямая в касательном пространстве $T_0(M)$, проходящая через начало координат, отображается в соответствующую геодезическую. Ясно, что такое отображение можно построить для каждой точки $x_0 \in M$.

Если геодезическая неполна, то экспоненциальное отображение (16.102) определено только для некоторого интервала значений канонического параметра $-\epsilon_1 < t < \epsilon_2$, где $\epsilon_{1,2} > 0$. В результате экспоненциальное отображение будет определено в некоторой окрестности начала координат касательного пространства. Если связность на M класса C^∞ , то экспоненциальное отображение будет того же класса гладкости C^∞ . Поскольку дифференциал экспоненциального отображения в точке x_0 невырожден, то существует окрестность $U_0 \ni x_0$, такая, что экспоненциальное отображение (16.102) является диффеоморфизмом $U_0 \rightarrow V_0 \subset T_0(M)$, где V_0 – некоторая окрестность касательного пространства, содержащая начало координат. Выберем координатный репер $\{\partial_\alpha\}$ в точке x_0 и предположим, что векторы скорости принимают значение на единичной сфере

$$\sum_{\alpha=1}^n (u_0^\alpha)^2 = 1.$$

Теперь отождествим касательное пространство $T_0(M)$ с евклидовым пространством $\mathbb{R}^n$ естественным образом, отождествив координаты касательного вектора $\{tu_0^\alpha\}$ с декартовыми координатами точки $\{y^\alpha\}$ в $\mathbb{R}^n$. В результате получим координатную систему, определенную на U_0 .

Определение. Система координат в окрестности $\mathbb{U}_0 \subset \mathbb{M}$ точки x_0 , определенная экспоненциальным отображением (16.102),

$$\varphi: \mathbb{M} \supset \mathbb{U}_0 \ni \{x^\alpha(t)\} \mapsto \{y^\alpha := tu_0^\alpha\} \in \mathbb{V}_0 \subset \mathbb{R}^n,$$

называется *нормальной*. □

Замечание. Подчеркнем, что нормальная система координат определена исключительно связностью, а не метрикой, которой вообще может не быть на многообразии. □

Если на многообразии $\mathbb{M}$ помимо связности Γ задана также метрика g , то исходную систему координат x^α в окрестности точки x_0 можно всегда выбрать таким образом, что координатный базис ∂_α будем ортонормальным в данной точке x_0 , т.е. $g_{\alpha\beta}(x_0) = \eta_{\alpha\beta}$. Мы всегда предполагаем, что для нормальной системы координат при наличии метрики данное условие выполнено.

Теорема 16.9.6 (Уайтхед). Пусть y^α – нормальная система координат в окрестности точки $x_0 \in \mathbb{M}$. Определим окрестность $\mathbb{U}_0(\rho)$ точки x_0 равенством $\sum_\alpha (y^\alpha)^2 < \rho^2$. Тогда существует положительное число r такое, что если $0 < \rho < r$, то:

- 1) окрестность $\mathbb{U}_0(\rho)$ является геодезически выпуклой, т.е. любые две точки из $\mathbb{U}_0(\rho)$ можно соединить геодезической, целиком лежащей в $\mathbb{U}_0(\rho)$;
- 2) каждая точка из $\mathbb{U}_0(\rho)$ имеет нормальную координатную окрестность, содержащую $\mathbb{U}_0(\rho)$.

Доказательство. См. [114]. □

Замечание. Данная теорема справедлива для произвольных достаточно гладких связностей, не зависимо от того задана ли на многообразии метрика или нет. □

16.10 Полнота римановых многообразий

Пусть задано риманово многообразие $(\mathbb{M}, g)$. В настоящем разделе положительная определенность метрики существенна, и мы будем ее предполагать.

Определение. Рассмотрим две произвольные точки многообразия $p, q \in \mathbb{M}$. Расстоянием между этими точками называется нижняя грань интегралов

$$l(p, q) := \inf \int_p^q ds, \quad ds := \sqrt{\dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}} \quad (16.103)$$

по всем кусочно дифференцируемым кривым класса $\mathcal{C}^1$, соединяющим эти точки. □

Функция расстояния определяет отображение

$$l: \mathbb{M} \times \mathbb{M} \ni (p, q) \mapsto l(p, q) \in \mathbb{R}_+.$$

Нетрудно проверить, что все свойства топологической метрики, рассмотренной в разделе (1.3.2) выполнены и, следовательно, функция $l(p, q)$ задает топологическую метрику на $\mathbb{M}$.

Определение. Множество точек

$$\mathbb{B}_r(p) := \{x \in \mathbb{M} : l(x, p) < r\} \quad (16.104)$$

называется *шаром* радиуса r с центром в точке $p \in \mathbb{M}$. $\square$

Предложение 16.10.1. *Функция расстояния (16.103) непрерывна и семейство метрических шаров для всех $p \in \mathbb{M}$ и $r \in \mathbb{R}_+$ образует базу исходной топологии многообразия.*

Доказательство. Следует из непрерывной зависимости интеграла от пределов интегрирования. $\square$

Для связностей Леви–Чивиты на римановых пространствах справедлива

Теорема 16.10.1 (Майерс, Стинрод). *Пусть $(\mathbb{M}_1, g_1)$ и $(\mathbb{M}_2, g_2)$ – римановы многообразия. Пусть l_1 и l_2 – функции расстояния на $\mathbb{M}_1$ и $\mathbb{M}_2$, соответственно. Если задано отображение $f : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ (которое не предполагается непрерывным или дифференцируемым), такое, что $l_1(p, q) = l_2(f(p), f(q))$ для всех $p, q \in \mathbb{M}_1$, то f есть диффеоморфизм из $\mathbb{M}_1$ на $\mathbb{M}_2$, который отображает метрику g_1 в g_2 .*

В частности, каждое отображение f из $\mathbb{M}$ на себя, сохраняющее функцию расстояния l , есть изометрия, т.е. отображение f сохраняет метрику g .

Доказательство. См. [115]. $\square$

Напомним, что топологическое пространство $\mathbb{M}$ и, в частности, многообразие, называется метрически полным, если любая фундаментальная последовательность в $\mathbb{M}$ сходится к некоторой точке из $\mathbb{M}$.

С другой стороны, в разделе 16.2 мы определили экстремали как линии, для которых первая вариация интеграла (16.103) равна нулю. Экстремали являются одновременно геодезическими линиями для связности Леви–Чивиты. Эта связность является полной, если любую экстремаль можно продолжить в обе стороны до бесконечных значений канонического параметра. В этом случае мы говорим, что риманово многообразие полно.

Таким образом, было введено два понятия полноты многообразия: метрическая полнота и полнота связности Леви–Чивиты. Оба эти понятия определяются одним геометрическим объектом – римановой метрикой и поэтому между ними существует тесная связь, которую мы здесь рассмотрим.

Теорема 16.10.2 (Хопф, Ринов). *Для связного риманова многообразия следующие условия эквивалентны:*

- 1) $(\mathbb{M}, g)$ – полное риманово многообразие;
- 2) $(\mathbb{M}, l)$ – полное метрическое пространство;
- 3) каждый замкнутый метрический шар $\overline{\mathbb{B}}_r(p)$ в $\mathbb{M}$ компактен;
- 4) для каждой точки $x_0 \in \mathbb{M}$ экспоненциальное отображение (16.102) определено на всем касательном пространстве $T_0(\mathbb{M})$.

Доказательство. См. [116]. $\square$

Теорема 16.10.3. *Если $\mathbb{M}$ – связное полное риманово многообразие, то $\mathbb{M}$ геодезически выпукло, т.е. любые две точки $p, q \in \mathbb{M}$ можно соединить минимизирующей экстремалью.*

Доказательство. См., например, [48], теорема 4.2. $\square$

Следствие. Если все геодезические, исходящие из любой выбранной точки p связного риманова многообразия полны, то $(\mathbb{M}, l)$ – геодезически и метрически полно.

Доказательство. См., например, [48], следствие 4.3. $\square$

Следствие. Каждое компактное риманово многообразие $(\mathbb{M}, g)$ метрически полно.

Доказательство. Следствие импликации $3) \rightarrow 1)$ в теореме 16.10.2. $\square$

Напомним, что риманово пространство $(\mathbb{M}, g)$ называется *однородным*, если группа изометрий действует на $\mathbb{M}$ транзитивно.

Теорема 16.10.4. Каждое однородное риманово многообразие $(\mathbb{M}, g)$ полно.

Доказательство. Пусть x – точка однородного риманова пространства $(\mathbb{M}, g)$. Тогда существует положительное число $r > 0$ такое, что для каждого единичного вектора $u \in T_x(\mathbb{M})$ геодезическая $\exp tu$ определена для каждого $|t| \leq r$. Пусть $\gamma = x(t)$, $0 \leq t \leq s$, – произвольная геодезическая в $\mathbb{M}$ с каноническим параметром t . Покажем, что эта геодезическая γ может быть продолжена до геодезической, определенной при $0 \leq t \leq s + r$. Пусть ι – изометрия многообразия $\mathbb{M}$, которая отображает точку x в $x(s)$. Тогда дифференциал обратной изометрии ι_*^{-1} отображает касательный вектор $\dot{x}(s)$ в u

$$u = \iota_*^{-1} \dot{x}(s).$$

Поскольку $\exp tu$ есть геодезическая, проходящая через точку x , то $\iota(\exp tu)$ – геодезическая, проходящая через $x(s)$. Положим

$$x(s+t) := \iota(\exp tu) \quad \text{для} \quad 0 \leq t \leq r.$$

Тогда кривая $\gamma = x(t)$ при $0 \leq t \leq s + r$ является продолжением геодезической. $\square$

Замечание. Доказанная теорема следует также из того общего факта, что каждое локально компактное метрическое однородное пространство полно. $\square$

Теорема 16.10.5. Пусть $\tilde{\mathbb{M}}$ и $\mathbb{M}$ – связные римановы многообразия одинаковой размерности. Допустим, что существует изометрическое погружение $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$. Тогда:

- 1) Если $\tilde{\mathbb{M}}$ полно, то отображение $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ является накрытием и $\mathbb{M}$ полно.
- 2) Обратно, если $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ – накрытие и $\mathbb{M}$ полно, то $\tilde{\mathbb{M}}$ также полно.

Доказательство. См., например, [48], теорема 4.6. $\square$

Следствие. Пусть $\tilde{\mathbb{M}}$ и $\mathbb{M}$ – связные многообразия одинаковой размерности и $p : \tilde{\mathbb{M}} \rightarrow \mathbb{M}$ – погружение. Тогда, если многообразие $\tilde{\mathbb{M}}$ компактно, то $\mathbb{M}$ также компактно, а p – накрывающее отображение.

Доказательство. Возьмем любую риманову метрику g на $\mathbb{M}$. Тогда p^*g – единственная метрика на $\tilde{\mathbb{M}}$ такая, что p – изометрическое погружение. Поскольку $\tilde{\mathbb{M}}$ компактно, то, по следствию 16.10, оно полно. Тогда из теоремы 16.10.5 следует, что p – накрывающее отображение и отсюда $\mathbb{M}$ компактно. $\square$

Замечание. В предыдущих теореме и следствии требование одинаковой размерности многообразий $\tilde{\mathbb{M}}$ и $\mathbb{M}$ является излишним, т.к. погружение возможно только для многообразий одинаковой размерности. $\square$

Определение. Говорят, что связное риманово пространство *непродолжаемо*, если его нельзя изометрически вложить в другое связное риманово пространство как собственное открытое подмногообразие. $\square$

Теорема 16.10.5 показывает, что каждое полное связное риманово многообразие непродолжаемо. Обратное утверждение неверно.

Пример 16.10.1. Пусть $\mathbb{M}$ есть евклидова плоскость с выколотым началом координат, а $\tilde{\mathbb{M}}$ – его универсальное накрывающее пространство. Как открытое подмногообразие евклидовой плоскости $\mathbb{M}$ имеет естественную евклидову метрику, которая, очевидно, неполна. На универсальной накрывающей $\tilde{\mathbb{M}}$ также задана естественная евклидова метрика. Риманово многообразие $\mathbb{M}$ неполно по теореме 16.10.5. Может быть доказано, что $\tilde{\mathbb{M}}$ непродолжаемо. Таким образом, неполное связное риманово многообразие в общем случае может быть непродолжаемо. $\square$

Следствие. Пусть $\mathbb{I}(\mathbb{M})$ – группа изометрий связного риманова многообразия $(\mathbb{M}, g)$. Если орбита $x\mathbb{I}$ точки $x \in \mathbb{M}$ содержит открытое подмножество из $\mathbb{M}$, то орбита $x\mathbb{I}$ совпадает со всем $\mathbb{M}$. Тем самым риманово многообразие $\mathbb{M}$ однородно.

16.11 Формулы Френе

Рассмотрим произвольную гладкую кривую $\gamma = \{x^\alpha(t)\}$ в трехмерном римановом многообразии $\mathbb{M}$, $\dim \mathbb{M} = 3$, с положительно определенной метрикой g . Тогда длина кривой (6.9) отлична от нуля. Выберем длину кривой в качестве канонического параметра t вдоль кривой. Предположим также, что на многообразии задана метрическая связность Γ (геометрия Римана–Картана). Единичный касательный вектор к кривой определяется вектором скорости

$$u^\alpha := \dot{x}^\alpha,$$

где точка обозначает дифференцирование по каноническому параметру t . Дифференцируя тождество $u^2 = 1$, получим равенство

$$\nabla_\beta u^2 = 2u_\alpha \nabla_\beta u^\alpha = 0. \quad (16.105)$$

Отсюда вытекает, что ковариантная производная (6.10) от u^α вдоль кривой γ (ускорение кривой),

$$\frac{Du^\alpha}{dt} = u^\beta \nabla_\beta u^\alpha = \frac{1}{\rho} n^\alpha, \quad (16.106)$$

где $\rho(t)$ – некоторая функция вдоль кривой, ортогональна вектору скорости. Здесь мы предполагаем, что $u^\beta \nabla_\beta u^\alpha \neq 0$, т.е. кривая γ не является геодезической. Вектор n^α всегда можно выбрать единичным:

$$(u, n) = 0, \quad n^2 = 1. \quad (16.107)$$

Определение. Единичное векторное поле $n^\alpha(s) \in \mathcal{X}(\mathbb{M}, \gamma)$, определенное вдоль кривой, называется *главной нормалью кривой*. Функция $1/\rho(t)$ называется *кривизной кривой*.

Поскольку пространство трехмерно, то дополним векторы u^α и n^α до ортонормированного базиса в касательном пространстве с помощью вектора *бинормали* к кривой, определяемого следующими соотношениями:

$$(u, b) = 0, \quad (n, b) = 0, \quad b^2 = 1. \quad \square \quad (16.108)$$

Ковариантные производные от n^α и b^α можно разложить по этому базису с некоторыми коэффициентами

$$\begin{aligned}\frac{Dn^\alpha}{dt} &= au^\alpha + bn^\alpha + cb^\alpha, \\ \frac{Db^\alpha}{dt} &= a'u^\alpha + b'n^\alpha + c'b^\alpha.\end{aligned}\tag{16.109}$$

Дифференцируя тождества (16.107), (16.108) вдоль кривой, получим условия на коэффициенты разложения:

$$\frac{1}{\rho} + a = 0, \quad b = 0, \quad a' = 0, \quad c + b' = 0, \quad c' = 0.$$

Тогда из соотношений (16.106) и (16.109) следуют *формулы Френе*:

$$\begin{aligned}\frac{Du^\alpha}{dt} &= \frac{1}{\rho}n^\alpha, \\ \frac{Dn^\alpha}{dt} &= -\frac{1}{\rho}u^\alpha + \frac{1}{\tau}b^\alpha, \\ \frac{Db^\alpha}{dt} &= -\frac{1}{\tau}n^\alpha,\end{aligned}\tag{16.110}$$

где введено обозначение

$$c = -b' = \frac{1}{\tau}.$$

Функция $1/\tau(t)$ называется *кручением* кривой.

Если кривая задана, то при желании можно найти явные выражения для векторных полей u^α , n^α и b^α , а также вычислить кривизну и кручение кривой. В трехмерном евклидовом пространстве $\mathbb{R}^3$ верно также обратное утверждение (см. [7], глава 1, §5, пункт 2). А именно, если известны кривизна и кручение как функции канонического параметра вдоль кривой, то можно восстановить кривую в $\mathbb{R}^3$ с точностью до движений (сдвигов, вращений и отражений) всего пространства. Таким образом, кривизна и кручение кривой в трехмерном евклидовом пространстве представляют собой полный набор геометрических инвариантов кривой.

Пример 16.11.1. Кривизна и кручение прямой линии в трехмерном евклидовом пространстве $\mathbb{R}^3$ равны нулю.

Пример 16.11.2. Рассмотрим спираль в трехмерном евклидовом пространстве $\mathbb{R}^3$, ось которой совпадает с осью z :

$$\begin{aligned}x &= R \cos \frac{t}{\sqrt{R^2 + v^2}}, \\ y &= R \sin \frac{t}{\sqrt{R^2 + v^2}}, \\ z &= \frac{vt}{\sqrt{R^2 + v^2}},\end{aligned}\tag{16.111}$$

где $R > 0$ и v – постоянные и параметр $t \in \mathbb{R}$ совпадает с длиной спирали. Касательный вектор к спирали имеет следующие компоненты:

$$\begin{aligned} u^x &= -\frac{R}{\sqrt{R^2 + v^2}} \sin \frac{t}{\sqrt{R^2 + v^2}}, \\ u^y &= \frac{R}{\sqrt{R^2 + v^2}} \cos \frac{t}{\sqrt{R^2 + v^2}}, \\ u^z &= \frac{v}{\sqrt{R^2 + v^2}}. \end{aligned}$$

Дифференцирование этих равенств по t приводит к следующему вектору главной нормали:

$$\begin{aligned} n^x &= -\cos \frac{t}{\sqrt{R^2 + v^2}}, \\ n^y &= -\sin \frac{t}{\sqrt{R^2 + v^2}}, \\ n^z &= 0. \end{aligned}$$

При этом кривизна спирали равна

$$\frac{1}{\rho} = \frac{R}{R^2 + v^2}.$$

Дальнейшее дифференцирование полученных равенств по t определяет вектор бинормали:

$$\begin{aligned} b^x &= \frac{v}{\sqrt{R^2 + v^2}} \sin \frac{t}{\sqrt{R^2 + v^2}}, \\ b^y &= -\frac{v}{\sqrt{R^2 + v^2}} \cos \frac{t}{\sqrt{R^2 + v^2}}, \\ b^z &= \frac{R}{\sqrt{R^2 + v^2}} \end{aligned}$$

и кручение спирали

$$\frac{1}{\tau} = \frac{v}{R^2 + v^2}.$$

Таким образом, вычислены все характеристики спирали.

Если $v = 0$, то спираль вырождается в окружность. Для окружности кривизна и кручение равны

$$\frac{1}{\rho} = \frac{1}{R}, \quad \frac{1}{\tau} = 0. \quad \square$$

Замечание. Кривизна $1/\rho$ и кручение $1/\tau$ кривой зависят от метрики g и тензора кручения T трехмерного многообразия $\mathbb{M}$, что следует из определения метрической связности. При выводе формул Френе условие метричности связности важно, т.к. при отличной от нуля неметричности формула (16.105) неверна и ковариантная производная (16.106) не будет ортогональна вектору скорости. Заметим также, что кривизна и кручение кривой являются понятиями, отличными от кривизны и кручения аффинной связности, введенных ранее. $\square$

Посмотрим на геодезические линии с точки зрения введенных выше понятий. Следующее утверждение очевидно.

Предложение 16.11.1. *Кривая γ в трехмерном пространстве Римана–Картана является геодезической тогда и только тогда, когда ее кривизна $1/\rho$ равна нулю.*

Для геодезической линии на $\mathbb{M}$ правая часть уравнения (16.106) равна нулю, и, следовательно, вектор нормали n к геодезической нельзя определить соотношением (16.106). Кручение геодезической линии также неопределено.

При смещении вдоль кривой, отличной от геодезической, на расстояние dt компоненты касательного вектора и вектора главной нормали получают приращение

$$Du^\alpha = \frac{dt}{\rho} n^\alpha, \quad Dn^\alpha = -\frac{dt}{\rho} u^\alpha + \frac{dt}{\tau} b^\alpha.$$

Отсюда следует, что при параллельном переносе вдоль кривой на бесконечно малое расстояние касательный вектор и вектор главной нормали остаются в плоскостях, натянутых на векторы u^α и n^α , тогда и только тогда, когда кручение кривой равно нулю, $1/\tau = 0$. В этом случае векторы u^α и n^α поворачиваются на угол $d\varphi = -dt/\rho$.

Если рассматривать кривую γ на двумерном многообразии с заданной римановой метрикой и метрической связностью, то вектор бинормали тождественно равен нулю, а формулы Френе принимают вид

$$\begin{aligned} \frac{Du^\alpha}{dt} &= \frac{1}{\rho} n^\alpha, \\ \frac{Dn^\alpha}{dt} &= -\frac{1}{\rho} u^\alpha. \end{aligned}$$

В этом случае кривые нулевой кривизны и только они являются геодезическими.

Глава 17

Разложения кривизны и кручения

Исследование моделей гравитации во многом упрощается и становится более прозрачным, если кривизну и кручение разложить на неприводимые компоненты относительно действия группы общих преобразований координат. В настоящей главе мы рассмотрим разложение компонент тензоров кривизны и кручения в геометрии Римана–Картана, которая определяется метрикой и кручением. Общий случай аффинной геометрии принципиальных трудностей не представляет, но является более громоздким, и мы его опускаем.

В моделях математической физики важную роль играют группы преобразований, которые оставляют инвариантным действие. В настоящей главе мы рассмотрим также три группы преобразований в моделях гравитации. Первая группа преобразований – конформная – является подгруппой общих преобразований координат, а две другие действуют непосредственно на метрику и связность, не затрагивая координаты.

17.1 Разложение тензора кривизны

Разложение произвольного тензора на неприводимые компоненты сводится к построению неприводимых представлений симметрической группы, т.е. группы перестановок некоторого набора символов, в роли которых выступают индексы рассматриваемого тензора. В общем случае эта процедура довольно сложна (см., например, [117], глава 7), и мы не будем ее рассматривать во всех деталях. Тензор кривизны содержит только четыре индекса, и для него разложение на неприводимые компоненты является интуитивно понятным.

Начнем с геометрии Римана–Картана (M, g, T) , когда на многообразии M произвольной размерности $\dim M = n$ задана метрика g и кручение T . Для определенности предположим, что метрика имеет лоренцеву сигнатуру, и компоненты тензоров будем рассматривать относительно ортонормального базиса касательного пространства e_a , $a = 0, 1, \dots, n - 1$. Ортонормальный базис допускает произвольные локальные вращения относительно полной группы Лоренца $O(1, n - 1)$, включающей обращение времени и пространственные отражения.

В разделе 6.8 было показано, что в геометрии Римана–Картана тензор кривизны R_{abcd} со всеми опущенными индексами антисимметричен относительно перестановки первой и второй пары индексов:

$$R_{abcd} = -R_{bacd}, \quad R_{abcd} = -R_{abdc}. \quad (17.1)$$

Другой симметрии относительно перестановки индексов он не имеет. В (псевдо)римановой геометрии добавляется еще одна симметрия относительно циклической перестановки первых трех индексов:

$$\tilde{R}_{abcd} + \tilde{R}_{bcad} + \tilde{R}_{cabd} = 0, \quad (17.2)$$

где знак тильды обозначает, что тензор кривизны построен для аффинной связности с нулевым кручением. Вместе с симметрией (17.1) равенство (17.2) влечет за собой симметрию относительно перестановки первой пары индексов со второй:

$$\tilde{R}_{abcd} = \tilde{R}_{cdab}. \quad (17.3)$$

Разложим тензор кривизны на неприводимые компоненты относительно действия полной группы Лоренца в общем случае, когда размерность пространства-времени достаточно большая, $n \geq 4$. Случаи низших размерностей имеют специфику, и будут рассмотрены отдельно.

Схема построения неприводимых компонент следующая. Из тензора кривизны выделяем след (тензор Риччи)

$$R_{ab} := R_{acb}{}^c.$$

Этот тензор имеет три неприводимые компоненты: симметричную бесследовую часть R_{ab}^s , антисимметричную часть R_{ab}^A и след (скалярную кривизну) $R := R_a{}^a$:

$$R_{ab} = R_{ab}^s + R_{ab}^A + \frac{1}{n} R \eta_{ab}, \quad (17.4)$$

где

$$R_{ab}^s := \frac{1}{2}(R_{ab} + R_{ba}) - \frac{1}{n} R \eta_{ab}, \quad (17.5)$$

$$R_{ab}^A := \frac{1}{2}(R_{ab} - R_{ba}). \quad (17.6)$$

Теперь бесследовый тензор четвертого ранга, который остается после выделения тензора Риччи, делим на компоненты, симметричные и антисимметричные относительно перестановки пар индексов, и выделяем полностью антисимметричную часть. В результате получаем шесть независимых неприводимых компонент при $n \geq 4$:

C_{abcd}^s – симметричный тензор Вейля,

C_{abcd}^A – антисимметричный тензор Вейля,

R_{ab}^s – симметричный бесследовый тензор Риччи,

R_{ab}^A – антисимметричный тензор Риччи,

R – скалярная кривизна,

$R_{[abcd]}$ – полностью антисимметричный тензор кривизны.

Разложение имеет вид

$$\begin{aligned} R_{abcd} = & C_{abcd}^s + C_{abcd}^A + \\ & + \frac{1}{n-2}(R_{ac}^s \eta_{bd} - R_{ad}^s \eta_{bc} - R_{bc}^s \eta_{ad} + R_{bd}^s \eta_{ac}) + \\ & + \frac{1}{n-2}(R_{ac}^A \eta_{bd} - R_{ad}^A \eta_{bc} - R_{bc}^A \eta_{ad} + R_{bd}^A \eta_{ac}) + \\ & - \frac{1}{(n-1)(n-2)} R (\eta_{ac} \eta_{bd} - \eta_{ad} \eta_{bc}) + R_{[abcd]}, \end{aligned} \quad (17.7)$$

где тензор конформной кривизны или тензор Вейля $C_{abcd} := C_{abcd}^S + C_{abcd}^A$ вычисляется по тензору кривизны, тензору Риччи, скалярной кривизне и полностью антисимметричной части:

$$R_{[abcd]} := \frac{1}{6}(R_{abcd} - R_{bcda} + R_{cdab} - R_{dabc} + R_{acdb} - R_{bdab}). \quad (17.8)$$

Явное выражение для тензора Вейля имеет вид

$$\begin{aligned} C_{abcd} := & R_{abcd} - \frac{1}{n-2}(R_{ac}\eta_{bd} - R_{ad}\eta_{bc} - R_{bc}\eta_{ad} + R_{bd}\eta_{ac}) + \\ & + \frac{1}{(n-1)(n-2)}R(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc}) - R_{[abcd]}. \end{aligned} \quad (17.9)$$

При этом он удовлетворяет следующим соотношениям:

$$\begin{aligned} C_{abcd} &= -C_{bacd} = -C_{abdc}, \\ C_{abc}{}^b &= 0, \quad C_{[abcd]} = 0, \end{aligned} \quad (17.10)$$

где квадратные скобки обозначают антисимметризацию по всем индексам, как в формуле (17.8). Тензор Вейля (17.9) взаимно однозначно определяется алгебраическими соотношениями (17.10). В разложении (17.7) тензор Вейля разложен на симметричную, C_{abcd}^S , и антисимметричную, C_{abcd}^A , части относительно перестановки первой и второй пары индексов,

$$C_{abcd}^S := \frac{1}{2}(C_{abcd} + C_{cdab}), \quad (17.11)$$

$$C_{abcd}^A := \frac{1}{2}(C_{abcd} - C_{cdab}). \quad (17.12)$$

В (псевдо)римановой геометрии (M, g) , когда тензор кручения равен нулю, $T_{abc} = 0$, тензор кривизны имеет только три неприводимые компоненты, т.к. три другие компоненты тождественно обращаются в нуль:

$$C_{abcd}^A \equiv 0, \quad R_{ac}^A \equiv 0, \quad R_{[abcd]} \equiv 0. \quad (17.13)$$

В двух и трехмерном пространстве тензор кривизны имеет меньшее число неприводимых компонент, что будет рассмотрено ниже.

Для положительно определенной метрики разложение на неприводимые компоненты (17.7) остается прежним, необходимо только метрику Лоренца η_{ab} заменить на евклидову метрику δ_{ab} .

В (псевдо)римановой геометрии вид тензора Вейля упрощается:

$$\begin{aligned} C_{abcd} := & \tilde{R}_{abcd} - \frac{1}{n-2}(\tilde{R}_{ac}\eta_{bd} - \tilde{R}_{ad}\eta_{bc} - \tilde{R}_{bc}\eta_{ad} + \tilde{R}_{bd}\eta_{ac}) + \\ & + \frac{1}{(n-1)(n-2)}\tilde{R}(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc}) \end{aligned} \quad (17.14)$$

и, так же, как и тензор кривизны, обладает дополнительной симметрией относительно перестановки первой и второй пары индексов, $C_{abcd} = C_{cdab}$.

Если тензор Риччи удовлетворяет вакуумным уравнениям Эйнштейна

$$R_{ab} - g_{ab}\Lambda = 0, \quad (17.15)$$

где Λ – космологическая постоянная, то тензор Вейля принимает вид

$$C_{abcd} = R_{abcd} - \frac{\Lambda}{(n-1)}(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc}) - R_{[abcd]}. \quad (17.16)$$

При нулевом кручении, т.е. в геометрии Римана последнее слагаемое отсутствует.

В (псевдо)римановом пространстве n измерений тензор кривизны удовлетворяет условиям (17.1), (17.2) и поэтому имеет

$$\frac{n^2(n^2-1)}{12} \quad (17.17)$$

независимых компонент, каждая из которых зависит от точки многообразия. При $n \geq 4$ число независимых компонент больше числа независимых компонент у тензора Риччи, и его нельзя однозначно восстановить по тензору Риччи. В двумерном и трехмерном пространствах, как будет следовать из дальнейших построений, это возможно.

Четырехмерное пространство-время, $n = 4$

В четырехмерном пространстве-времени существует полностью антисимметричный псевдотензор четвертого ранга,

$$\varepsilon_{abcd} = \varepsilon_{[abcd]}, \quad \varepsilon_{0123} = 1,$$

где квадратные скобки обозначают антисимметризацию по всем индексам, а приставка псевдо указывает, что при инверсии пространственных координат или обращении времени он, по определению, меняет знак. С его помощью ряд формул упрощается.

При $n = 4$ антисимметричная компонента тензора Вейля C_{abcd}^A (17.12) взаимно однозначно параметризуется симметричным бесследовым псевдотензором второго ранга:

$$D_{ab} = D_{ba}, \quad D_a^a = 0.$$

Соответствующие формулы имеют вид

$$C_{abcd}^A = -\frac{1}{8}(\varepsilon_{abce}D^e_d - \varepsilon_{abde}D^e_c - \varepsilon_{cdae}D^e_b + \varepsilon_{cdbe}D^e_a),$$

$$D_{de} := C_{abcd}^A \varepsilon^{abc}_e.$$

В последней формуле симметричность тензора D_{de} проверяется сверткой с ε^{defg} .

Полностью антисимметричная компонента тензора кривизны параметризуется псевдоскалярной кривизной R^* :

$$R_{[abcd]} = \varepsilon_{abcd}R^* \quad \Leftrightarrow \quad R^* := -\frac{1}{24}R_{abcd}\varepsilon^{abcd}. \quad (17.18)$$

Знаменатель 24 возник из-за того, что полностью антисимметричный тензор четвертого ранга имеет 24 нетривиальные компоненты.

Симметричная часть тензора Вейля, тензор Риччи и скалярная кривизна в четырехмерном пространстве-времени остаются прежними.

В (псевдо)римановой геометрии тензор кривизны имеет три неприводимые компоненты: симметричный тензор Вейля $\tilde{C}_{abcd}^s$, симметричный бесследовый тензор Риччи $\tilde{R}_{ab}^s$ и скалярную кривизну $\tilde{R}$. Наличие полностью антисимметричного тензора четвертого ранга не приводит к каким либо упрощениям.

Если метрика положительно определена, то в аффинной геометрии общего вида существует дополнительная возможность разложить тензор кривизны на самодуальную и антисамодуальную части. По определению, тензор кривизны является 2-формой

$$R^{ab} = \frac{1}{2} dx^\alpha \wedge dx^\beta R_{\alpha\beta}{}^{ab}, \quad R_{\alpha\beta ab} := e_\alpha{}^c e_\beta{}^d R_{cdab},$$

и его компоненты антисимметричны по первым двум индексам. Определим тензор, дуальный к 2-форме кривизны, с помощью оператора Ходжа (см. раздел 3.5)

$$* R_{\alpha\beta}{}^{ab} := \frac{1}{2} \varepsilon_{\alpha\beta\gamma\delta} R^{\gamma\delta ab}, \quad (17.19)$$

где индексы поднимаются и опускаются с помощью метрики. Поскольку метрика риманова, то справедлива формула (3.53):

$$**R = R, \quad (17.20)$$

где мы, для краткости, опустили тензорные индексы. Поэтому можно ввести самодуальный R^D и антисамодуальный R^A тензоры кривизны:

$$R^D := \frac{1}{2}(R + *R), \quad R^A := \frac{1}{2}(R - *R). \quad (17.21)$$

Из определения следует, что эти тензоры удовлетворяют уравнениям самодуальности и антисамодуальности:

$$* R^D = R^D, \quad * R^A = -R^A. \quad (17.22)$$

Как легко видеть, эти уравнения инвариантны относительно действия оператора Ходжа (умножение на звездочку слева) в силу равенства (17.20). Из построения также следует, что разложение тензора кривизны на самодуальную и антисамодуальную компоненты инвариантно относительно общих преобразований координат.

В геометрии Римана–Картана тензор кривизны антисимметричен также относительно перестановки второй пары индексов. Поэтому аналогичное разложение можно провести для второй пары индексов.

Замечание. Если метрика четырехмерного пространства-времени имеет лоренцеву сигнатуру, то квадрат оператора Ходжа равен минус единице: $**R = -R$, и это вносит существенные изменения в конструкцию. По прежнему тензор кривизны можно представить в виде суммы двух слагаемых (17.21). При действии оператора Ходжа эти компоненты перемешиваются:

$$* R^D = -R^A, \quad * R^A = R^D. \quad (17.23)$$

Если подействовать оператором Ходжа на полученные соотношения еще раз, то возникают алгебраические уравнения, такие же как и для полного тензора кривизны:

$$** R^D = -R^D, \quad ** R^A = -R^A,$$

которые не несут новой информации. Мы видим, что для псевдоримановых многообразий разложение тензора кривизны на самодуальную и антисамодуальную части невозможно. $\square$

Трехмерная геометрия Римана–Картана, $n = 3$

В трехмерном случае возникают существенные упрощения ввиду наличия полностью антисимметричного тензора третьего ранга:

$$\varepsilon_{abc} = \varepsilon_{[abc]}, \quad \varepsilon_{012} = \varepsilon^{012} = 1 \quad \text{или} \quad \varepsilon_{123} = \varepsilon^{123} = 1.$$

Предложение 17.1.1. *В геометрии Римана–Картана при $n = 3$ полный тензор кривизны взаимно однозначно определяется тензором Риччи и скалярной кривизной:*

$$\begin{aligned} R_{abcd} &= -\varepsilon_{abe}\varepsilon_{cdf}R^{fe} + \frac{1}{2}(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc})R = \\ &= \eta_{ac}R_{bd} - \eta_{ad}R_{bc} - \eta_{bc}R_{ad} + \eta_{bd}R_{ac} - \frac{1}{2}(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc})R. \end{aligned} \quad (17.24)$$

Доказательство. Тензор кривизны при $n = 3$ можно представить в виде

$$R_{abcd} = A\varepsilon_{abe}\varepsilon_{cdf}X^{ef}, \quad (17.25)$$

где A – постоянная и X^{ef} – некоторый тензор второго ранга, не обязательно симметричный. Тогда тензор Риччи примет вид

$$R_{ac} = A(\eta_{ac}X - X_{ac}), \quad X := X^{bd}\eta_{bd}. \quad (17.26)$$

Последующая свертка дает скалярную кривизну:

$$R = 2AX.$$

Отсюда следует, что, без ограничения общности, можно положить

$$A = \frac{1}{2}, \quad X = R.$$

Подстановка полученных значений в равенство (17.26) позволяет определить неизвестный тензор

$$X_{ac} = \eta_{ac}R - 2R_{ca}.$$

С учетом этого выражения равенство (17.25) принимает вид (17.24), где мы учли формулу для свертки антисимметричного тензора третьего ранга (37.67). $\square$

Теперь нетрудно проверить, что в рассматриваемом случае три неприводимые компоненты тензора кривизны тождественно равны нулю:

$$C_{abcd}^s \equiv 0, \quad C_{abcd}^A \equiv 0, \quad R_{[abcd]} \equiv 0.$$

Таким образом, в трехмерном случае у тензора кривизны в геометрии Римана–Картана имеется только три неприводимые компоненты: симметричная бесследовая часть тензора Риччи R_{ab}^s , антисимметричная часть тензора Риччи R_{ab}^A и скалярная кривизна R .

В (псевдо)римановой геометрии при нулевом кручении тензор кривизны имеет только две неприводимые компоненты: симметричный бесследовый тензор Риччи $\tilde{R}_{ab}^s$ и скалярную кривизну $\tilde{R}$. Он полностью определяется своим симметричным тензором Риччи по формуле (17.24).

Двумерная геометрия Римана–Картана

При $n = 2$ существует полностью антисимметричный тензор второго ранга

$$\varepsilon_{ab} = \varepsilon_{[ab]}, \quad \varepsilon_{01} = -\varepsilon^{01} = 1 \quad \text{или} \quad \varepsilon_{12} = \varepsilon^{12} = 1. \quad (17.27)$$

Предложение 17.1.2. В геометрии Римана–Картана при $n = 2$ полный тензор кривизны взаимно однозначно определяется скалярной кривизной

$$R_{abcd} = \pm \frac{1}{2} \varepsilon_{ab} \varepsilon_{cd} R = \frac{1}{2} (\eta_{ac} \eta_{bd} - \eta_{ad} \eta_{bc}) R, \quad (17.28)$$

где знак $\pm$ соответствует евклидовой и лоренцевой сигнатуре метрики.

Доказательство. Тензор кривизны при $n = 2$ можно представить виде

$$R_{abcd} = \varepsilon_{ab} \varepsilon_{cd} X,$$

где X – некоторая функция. Свертка по индексам b и d приводит к равенству

$$R_{ac} = \pm \eta_{ac} X,$$

где знак $\pm$ соответствует евклидовой и лоренцевой сигнатуре (см. формулы (37.70) и (37.76). Последующая свертка индексов позволяет найти неизвестную функцию

$$X = \frac{1}{2} R.$$

Что и дает выражение (17.28). □

Таким образом, в двумерном случае у тензора кривизны есть только одна неприводимая компонента – скалярная кривизна R .

17.2 Разложение тензора кручения

Рассмотрим теперь разложение тензора кручения. Из него можно выделить след T_a и полностью антисимметричную часть $T_{[abc]}$. Таким образом, при $n \geq 3$ тензор кривизны имеет три неприводимые компоненты:

$$T_{abc} = S_{abc} + \frac{1}{n-1} (\eta_{ac} T_b - \eta_{bc} T_a) + T_{[abc]}, \quad (17.29)$$

где

$$\begin{aligned} T_b &:= T_{ab}^a, \\ T_{[abc]} &:= \frac{1}{3} (T_{abc} + T_{cab} + T_{bca}). \end{aligned}$$

Неприводимая компонента S_{abc} , по определению, удовлетворяет следующим лоренц ковариантным условиям:

$$S_{abc} = -S_{bac}, \quad S_{ab}^a = 0, \quad S_{abc} + S_{bca} + S_{cab} = 0.$$

Выписанные выше разложения тензоров кривизны и кручения на неприводимые компоненты существенно упрощаются в пространствах низших размерностей. Помимо этого при $n = 2$ и 4 справедлива формула Гаусса–Бонне и ее обобщение, которые позволяют исключить из лагранжиана одно слагаемое. Специфические свойства пространств низших размерностей подробно рассмотрим в соответствующих разделах.

Четырехмерное пространство-время, $n = 4$

В четырехмерной геометрии Римана–Картана упрощается вид одной из неприводимых компонент тензора кручения. А именно, полностью антисимметричная часть тензора кручения параметризуется псевдовекторным полем T^{*d} :

$$T_{[abc]} = \varepsilon_{abcd} T^{*d} \quad \Leftrightarrow \quad T^{*d} := -\frac{1}{6} T_{abc} \varepsilon^{abcd}.$$

На четырехмерном многообразии с римановой метрикой 2-форму кручения можно разложить на самодуальную и антисамодуальную компоненты, так же, как это было сделано для 2-формы кривизны.

Трехмерное пространство-время, $n = 3$

У тензора кручения в трехмерном случае, так же, как и в большем числе измерений, все три неприводимые компоненты отличны от нуля. Упрощается только вид полностью антисимметричной части, которая параметризуется псевдоскаляром T^* :

$$T_{[abc]} = \varepsilon_{abc} T^*, \quad \Leftrightarrow \quad T^* := \frac{1}{6} T_{abc} \varepsilon^{abc}.$$

Еще одно формальное упрощение касается лоренцевой связности, поскольку справедлива следующая параметризация

$$\omega_\alpha^{ab} = \omega_{\alpha c} \varepsilon^{cab} \quad \Leftrightarrow \quad \omega_{\alpha c} := \frac{1}{2} \omega_\alpha^{ab} \varepsilon_{abc}.$$

Соответствующий тензор кривизны примет вид

$$R_{\alpha\beta c} := R_{\alpha\beta}^{ab} \varepsilon_{abc} = 2(\partial_\alpha \omega_{\beta c} - \partial_\beta \omega_{\alpha c} + \omega_\alpha^a \omega_\beta^b \varepsilon_{abc}).$$

Использование компонент $\omega_{\alpha c}$ вместо компонент ω_α^{ab} упрощает многие формулы и широко используется при изучении моделей трехмерной гравитации.

Двумерное пространство-время, $n = 3$

В двумерном случае тензор кручения имеет только одну неприводимую компоненту – след:

$$T_{abc} = \eta_{ac} T_b - \eta_{bc} T_a, \quad T_b := T_{ab}^a.$$

Его также можно представить в виде

$$T_{abc} = \bigoplus \frac{1}{2} \varepsilon_{ab} T_c^* \quad \Leftrightarrow \quad T_c^* := T_{abc} \varepsilon^{ab},$$

где знак $\bigoplus$ соответствует римановой и лоренцевой сигнатуре метрики.

Лоренцева или $\mathbb{SO}(2)$ связность параметризуются ковекторным полем:

$$\omega_\alpha^{ab} = \omega_\alpha \varepsilon^{ab} \quad \Leftrightarrow \quad \omega_\alpha = \bigoplus \frac{1}{2} \omega_\alpha^{ab} \varepsilon_{ab},$$

что также приводит к упрощению многих формул.

17.3 Конформные преобразования

Рассмотрим многообразие M , $\dim M = n$, на котором задана метрика g . Группа общих преобразований координат (см. раздел 1.5) содержит подгруппу, при действии которой метрика умножается на некоторый ненулевой множитель.

Определение. Преобразование координат $x^\alpha \mapsto x'^\alpha = f^\alpha(x)$ назовем *конформным*, если метрика после преобразования

$$g_{\alpha\beta}(x) \mapsto g'_{\alpha\beta}(x') := \frac{\partial x^\gamma}{\partial x'^\alpha} \frac{\partial x^\delta}{\partial x'^\beta} g_{\gamma\delta}(x)$$

умножается на некоторый отличный от нуля множитель:

$$\frac{\partial x^\gamma}{\partial x'^\alpha} \frac{\partial x^\delta}{\partial x'^\beta} g_{\gamma\delta}(x) = \Omega^2(x') g_{\alpha\beta}(x'). \quad (17.30)$$

Здесь в правой части уравнения стоит старая метрика, взятая в точке x' , и $\Omega(x') > 0$ – произвольный конформный множитель (достаточно гладкая строго положительная функция). $\square$

Замечание. Конформный множитель не является скалярным полем (функцией) в строгом смысле слова, т.к. определен не для всех преобразований координат. $\square$

Ясно, что множество конформных преобразований образует группу, которая по построению является подгруппой общих преобразований координат.

Вычислив определитель от обеих частей равенства (17.30), получим равенство

$$g(x) J^{-2}(x) = g(x) \Omega^{2n}(x), \quad g := \det g_{\alpha\beta}, \quad (17.31)$$

или

$$\Omega = |J|^{-1/n},$$

т.е. конформный множитель Ω , если он существует, для невырожденных преобразований координат всегда отличен от нуля и однозначно определяется модулем якобиана преобразования.

При конформных преобразованиях координат в (псевдо)римановом пространстве произвольной размерности углы между векторами сохраняются, а меняются только длины векторов. Заметим, что поскольку конформные преобразования являются частным случаем общих преобразований координат, то ориентация векторов в каждой точке лоренцева многообразия сохраняется.

Пример 17.3.1. В двумерном римановом пространстве метрику всегда локально можно привести к конформно плоскому виду (теорема 28.3.1)

$$ds^2 = e^{2\phi}(dx^2 + dy^2) = e^{2\phi} dz d\bar{z},$$

где $z := x + iy$ – комплексные координаты и $\phi(x, y)$ – некоторая “функция”. При конформных преобразованиях комплексной плоскости $z \mapsto w(z)$ метрика умножается на множитель:

$$ds^2 = e^{2\phi} \frac{dz}{dw} \frac{d\bar{z}}{d\bar{w}} dw d\bar{w}.$$

Следовательно, конформные преобразования комплексной плоскости являются конформными в смысле определения, данного в начале раздела. Это служит оправданием названия преобразования координат (17.30) конформными. $\square$

В аффинной геометрии при конформных преобразованиях метрика преобразуется просто, а компоненты связности, кручения, неметричности и тензора кривизны преобразуются так, как требуется общими преобразованиями координат. В общем случае эти преобразования не имеют простого вида.

Бесконечно малое конформное преобразование $x'^\alpha = x^\alpha + \epsilon K^\alpha(x)$, $\epsilon \ll 1$, генерируется *конформным* векторным полем Киллинга $K(x)$, которое является обобщением векторного поля Киллинга (15.4). Формально оно определяется следующим уравнением

$$\mathcal{L}_K g_{\alpha\beta} = \lambda g_{\alpha\beta}, \quad (17.32)$$

где производная Ли берется вдоль конформного векторного поля Киллинга, и $\lambda(x)$ – некоторая функция. Это уравнение является следствием равенства (17.30) при $\Omega^2(x) = 1 + \epsilon\lambda(x)$ в первом порядке по ϵ . В развернутом виде при заданной метрике возникает уравнение на конформное поле Киллинга

$$g_{\alpha\gamma}\partial_\beta K^\gamma + g_{\beta\gamma}\partial_\alpha K^\gamma + K^\gamma\partial_\gamma g_{\alpha\beta} = \lambda g_{\alpha\beta}.$$

Это уравнение можно переписать в явно ковариантном виде

$$\tilde{\nabla}_\alpha K_\beta + \tilde{\nabla}_\beta K_\alpha = \lambda g_{\alpha\beta}. \quad (17.33)$$

Оно представляет собой простое обобщение равенства (15.4) для обычных векторных полей Киллинга.

В математической физике любая модель, инвариантная относительно общих преобразований координат, инвариантна также относительно конформных преобразований, что следует из определения. Однако требование инвариантности относительно общих преобразований координат можно ослабить, потребовав инвариантность только относительно конформных преобразований. Такие модели нашли широкое применение в современной математической физике.

17.4 Преобразование Вейля в (псевдо)римановой геометрии

Рассмотрим (псевдо)риманово многообразие $(\mathbb{M}, g)$, $\dim \mathbb{M} = n$, произвольной размерности. Для краткости, мы не будем отмечать знаком тильды геометрические объекты в этом и следующем разделе.

Определение. Преобразованием Вейля метрики $g \mapsto \bar{g}$ называется ее умножение на функцию, которая нигде не обращается в нуль [56]. Для определенности, будем считать эту функцию строго положительной и параметризуем ее с помощью экспоненты:

$$\bar{g}_{\alpha\beta} = e^{2\phi} g_{\alpha\beta} \quad \Leftrightarrow \quad \bar{g}^{\alpha\beta} = e^{-2\phi} g^{\alpha\beta}, \quad (17.34)$$

где $\phi = \phi(x)$ – достаточно гладкая функция. Функция $e^{2\phi}$ называется *конформным множителем*. $\square$

Преобразование Вейля можно записать в инвариантном виде для интервалов:

$$d\bar{s}^2 = e^{2\phi} ds^2. \quad (17.35)$$

Это преобразование меняет длины векторов (4.9), но не изменяет углы между ними (4.10). В этом отношении преобразование Вейля похоже на конформное преобразование, рассмотренное в разделе 17.3. Однако при преобразовании Вейля координаты

не преобразуются, и это важно. В теории струн, например, присутствуют оба вида симметрии, и их следует различать.

Метрика $\bar{g}_{\alpha\beta}$ после преобразования Вейля конформно эквивалентна метрике $g_{\alpha\beta}$ в смысле определения (4.11). Заметим также, что ориентация векторных полей и тип кривых при преобразовании Вейля для псевдоримановых многообразий не меняется. В частности, если метрика $g_{\alpha\beta}$ имеет лоренцеву сигнатуру, то вейлевские преобразования сохраняют световые конусы во всех точках многообразия. Это значит, что преобразования Вейля не меняют причинную структуру пространства-времени.

Вейлевские преобразования (17.34) сохраняют световой конус во всех точках лоренцева многообразия, поскольку уравнения

$$ds^2 = 0 \quad \Leftrightarrow \quad d\bar{s}^2 = 0$$

эквивалентны. Интервал для пространственно- и времениподобных кривых не инвариантен относительно преобразования (17.34). Следовательно, требование инвариантности физической модели относительно преобразований Вейля приводит к тому, что физический смысл имеют только относительные измерения времен и расстояний.

При вейлевских преобразованиях определитель метрики преобразуется по правилу:

$$\det(\bar{g}_{\alpha\beta}) = e^{2n\phi} \det(g_{\alpha\beta}). \quad (17.36)$$

Преобразование Вейля (17.34) определяет преобразование символов Кристоффеля (6.24) и соответствующего тензора кривизны (6.74). Прямые вычисления приводят к следующему правилу преобразования символов Кристоффеля

$$\bar{\Gamma}_{\alpha\beta}^{\gamma} = \Gamma_{\alpha\beta}^{\gamma} + \partial_{\alpha}\phi\delta_{\beta}^{\gamma} + \partial_{\beta}\phi\delta_{\alpha}^{\gamma} - g_{\alpha\beta}g^{\gamma\delta}\partial_{\delta}\phi. \quad (17.37)$$

Это, в свою очередь, приводит к преобразованию тензора кривизны:

$$e^{-2\phi}\bar{R}_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta} + g_{\alpha\gamma}\sigma_{\beta\delta} - g_{\alpha\delta}\sigma_{\beta\gamma} - g_{\beta\gamma}\sigma_{\alpha\delta} + g_{\beta\delta}\sigma_{\alpha\gamma} + (g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma})\partial\phi^2, \quad (17.38)$$

где введены сокращенные обозначения:

$$\sigma_{\alpha\beta} = \sigma_{\beta\alpha} := \nabla_{\alpha}\nabla_{\beta}\phi - \partial_{\alpha}\phi\partial_{\beta}\phi, \quad \partial\phi^2 := g^{\alpha\beta}\partial_{\alpha}\phi\partial_{\beta}\phi.$$

Выше, для краткости, опущены знаки тильды, которые используются для обозначения геометрических объектов, построенных только по метрике.

Тензор Риччи и скалярная кривизна преобразуются по правилам:

$$\bar{R}_{\alpha\beta} = R_{\alpha\beta} + (n-2)(\sigma_{\alpha\beta} + g_{\alpha\beta}\partial\phi^2) + g_{\alpha\beta}\square\phi, \quad (17.39)$$

$$e^{2\phi}\bar{R} = R + 2(n-1)\square\phi + (n-1)(n-2)\partial\phi^2, \quad (17.40)$$

где $\square := g^{\alpha\beta}\nabla_{\alpha}\nabla_{\beta}$ – оператор Лапласа–Бельтрами (волновой оператор) (3.62). Используя соотношение между метриками (17.34), закон преобразования скалярной кривизны можно записать в следующем виде:

$$\bar{g}_{\alpha\beta}R = g_{\alpha\beta} [R + 2(n-1)\square\phi + (n-1)(n-2)\partial\phi^2]. \quad (17.41)$$

Приведем также закон преобразования бесследовой части тензора Риччи (17.5)

$$\bar{R}_{\alpha\beta}^s = R_{\alpha\beta}^s + (n-2)\sigma_{\alpha\beta} - \frac{n-2}{n}(\square\phi - \partial\phi^2)g_{\alpha\beta}. \quad (17.42)$$

Заметим, что символы Кристоффеля до и после вейлевского преобразования (17.37) нельзя связать между собой преобразованием (16.9), не меняющим форму экстремалей. Это означает, что форма экстремалей (которые совпадают с геодезическими в (псевдо)римановом пространстве) меняется. Исключение составляют светоподобные экстремали, форма которых остается без изменения, а меняется лишь канонический параметр.

17.5 Вейлевская эквивалентность (псевдо)римановых пространств

Определение. Два (псевдо)римановых пространства (M, g) и $(\bar{M}, \bar{g})$ называются *конформно эквивалентными* или *эквивалентными относительно преобразований Вейля*, если они диффеоморфны, $M \approx \bar{M}$, и их метрики связаны преобразованием Вейля (17.35). $\square$

Сначала докажем необходимое условие вейлевской эквивалентности двух (псевдо)римановых пространств. При $n \geq 3$ из правила преобразования тензора Риччи (17.39) получаем формулу

$$\sigma_{\alpha\beta} = \frac{1}{n-2}(\bar{R}_{\alpha\beta} - R_{\alpha\beta}) - \frac{1}{2(n-1)(n-2)}(\bar{g}_{\alpha\beta}\bar{R} - g_{\alpha\beta}R) - \frac{1}{2}g_{\alpha\beta}\partial\phi^2. \quad (17.43)$$

Теперь подставим найденное выражение в закон преобразования тензора кривизны (17.38) с одним поднятым индексом. После несложных вычислений получим равенство

$$\bar{C}_{\alpha\beta\gamma}{}^\delta = C_{\alpha\beta\gamma}{}^\delta, \quad (17.44)$$

где

$$\begin{aligned} C_{\alpha\beta\gamma}{}^\delta := & R_{\alpha\beta\gamma}{}^\delta - \frac{1}{n-2}(g_{\alpha\gamma}R_{\beta}{}^\delta - \delta_{\alpha}^{\delta}R_{\beta\gamma} - g_{\beta\gamma}R_{\alpha}{}^\delta + \delta_{\beta}^{\delta}R_{\alpha\gamma}) + \\ & + \frac{1}{2(n-1)(n-2)}(g_{\alpha\gamma}\delta_{\beta}^{\delta} - \delta_{\alpha}^{\delta}g_{\beta\gamma} - g_{\beta\gamma}\delta_{\alpha}^{\delta} + \delta_{\beta}^{\delta}g_{\alpha\gamma})R \end{aligned} \quad (17.45)$$

– тензор Вейля (17.9). Если у тензора Вейля опущены все индексы, то он преобразуется нетривиально:

$$e^{-2\phi}\bar{C}_{\alpha\beta\gamma\delta} = C_{\alpha\beta\gamma\delta}.$$

Таким образом, необходимым условием вейлевской эквивалентности (псевдо)римановых пространств является равенство тензоров Вейля с одним контравариантным индексом для заданных метрик (17.44). Этот результат не зависит от сигнатуры метрики.

Теперь поставим следующую задачу. Найти критерий того, что данное (псевдо)риманово многообразие является локально конформно (псевдо)евклидовым. Это значит, что необходимо найти такую функцию ϕ , чтобы после преобразования Вейля (17.34) полный тензор кривизны обратился в нуль в некоторой окрестности произвольной точки: $\bar{R}_{\alpha\beta\gamma}{}^\delta = 0$. Для этого необходимо, чтобы функция ϕ при $n \geq 3$ удовлетворяла уравнению

$$\nabla_{\alpha}\nabla_{\beta}\phi = \partial_{\alpha}\phi\partial_{\beta}\phi - \frac{1}{2}g_{\alpha\beta}\partial\phi^2 - \frac{1}{n-2}\left[R_{\alpha\beta} - \frac{1}{2(n-1)}g_{\alpha\beta}R\right]. \quad (17.46)$$

Оно является следствием формулы (17.43) при $\bar{R}_{\alpha\beta} = 0$ и $\bar{R} = 0$.

Из предыдущего рассмотрения вытекает, что необходимым условием конформной (псевдо)евклидовости является равенство нулю тензора Вейля, $C_{\alpha\beta\gamma}{}^\delta = 0$.

Для того чтобы найти необходимые и достаточные условия локальной разрешимости системы уравнений в частных производных (17.46), рассмотрим систему уравнений

$$\partial_{\alpha}\partial_{\beta}\phi = F_{\alpha\beta},$$

где $F_{\alpha\beta}(\phi, x)$ – некоторая достаточно гладкая функция своих аргументов. Запишем ее в виде системы уравнения первого порядка:

$$\partial_\alpha \phi := G_\alpha, \quad \partial_\alpha G_\beta = F_{\alpha\beta}.$$

Необходимым и достаточным условием локальной разрешимости данной системы уравнений являются два условия:

$$\partial_\alpha G_\beta = \partial_\beta G_\alpha \Leftrightarrow F_{\alpha\beta} = F_{\beta\alpha}, \quad \partial_\alpha F_{\beta\gamma} = \partial_\beta F_{\alpha\gamma}.$$

Поскольку правая часть уравнения (17.46) симметрична по индексам α, β , то первое условие интегрируемости выполнено. Второе условие локальной интегрируемости запишем в ковариантном виде:

$$\nabla_\alpha \nabla_\beta \nabla_\gamma \phi - \nabla_\beta \nabla_\alpha \nabla_\gamma \phi = [\nabla_\alpha, \nabla_\beta] \nabla_\gamma \phi = -R_{\alpha\beta\gamma}{}^\delta \nabla_\delta \phi, \quad (17.47)$$

где мы воспользовались видом коммутатора ковариантных производных от ковекторного поля (6.94). Подстановка в левую часть полученного соотношения равенства (17.46) дает

$$\nabla_\beta \phi \nabla_\alpha \nabla_\gamma \phi - \frac{1}{2} g_{\beta\gamma} \nabla_\alpha \partial \phi^2 - \nabla_\alpha \phi \nabla_\beta \nabla_\gamma \phi + \frac{1}{2} g_{\alpha\gamma} \nabla_\beta \partial \phi^2 - \frac{1}{n-2} C_{\alpha\beta\gamma}, \quad (17.48)$$

где введен *тензор Коттона* [118]

$$C_{\alpha\beta\gamma} := \nabla_\alpha R_{\beta\gamma} - \nabla_\beta R_{\alpha\gamma} + \frac{1}{2(n-1)} (g_{\alpha\gamma} \nabla_\beta R - g_{\beta\gamma} \nabla_\alpha R). \quad (17.49)$$

Снова подставим в формулу (17.48) выражение для вторых производных из равенства (17.46). В результате первые четыре слагаемых примут вид

$$\begin{aligned} & -\frac{1}{n-2} (R_{\alpha\gamma} \nabla_\beta \phi - R_{\beta\gamma} \nabla_\alpha \phi - g_{\beta\gamma} R_\alpha{}^\delta \nabla_\delta \phi + g_{\alpha\gamma} R_\beta{}^\delta \nabla_\delta \phi) + \\ & + \frac{1}{(n-1)(n-2)} (g_{\alpha\gamma} \nabla_\beta \phi - g_{\beta\gamma} \nabla_\alpha \phi) R. \end{aligned}$$

Нетрудно проверить, что это – в точности правая часть равенства (17.47), куда подставлено выражение для тензора кривизны

$$\begin{aligned} R_{\alpha\beta\gamma}{}^\delta &= \frac{1}{n-2} (R_{\alpha\gamma} \delta_\beta^\delta - R_{\beta\gamma} \delta_\alpha^\delta - g_{\beta\gamma} R_\alpha{}^\delta + g_{\alpha\gamma} R_\beta{}^\delta) - \\ & - \frac{1}{(n-1)(n-2)} (g_{\alpha\gamma} \delta_\beta^\delta - g_{\beta\gamma} \delta_\alpha^\delta) R, \end{aligned} \quad (17.50)$$

которое получено из равенства нулю тензора Вейля (17.45). Таким образом, необходимым и достаточным условием локальной разрешимости системы уравнений на конформный множитель (17.46) при $n \geq 3$ является равенство нулю тензора Коттона, $C_{\alpha\beta\gamma} = 0$.

Покажем, что если размерность пространства превосходит три, $n \geq 4$, то равенство нулю тензора Коттона вытекает из равенства нулю тензора Вейля. Из тождеств Бианки для тензора кривизны (6.116),

$$\nabla_\alpha R_{\beta\gamma\delta}{}^\epsilon + \nabla_\beta R_{\gamma\alpha\delta}{}^\epsilon + \nabla_\gamma R_{\alpha\beta\delta}{}^\epsilon = 0,$$

вытекает тождество для тензора Вейля:

$$\begin{aligned} \nabla_\alpha C_{\beta\gamma\delta}^\epsilon + \nabla_\beta C_{\gamma\alpha\delta}^\epsilon + \nabla_\gamma C_{\alpha\beta\delta}^\epsilon = \\ = \frac{1}{n-2} (-C_{\alpha\beta\delta} \delta_\gamma^\epsilon - C_{\beta\gamma\delta} \delta_\alpha^\epsilon - C_{\gamma\alpha\delta} \delta_\beta^\epsilon + C_{\alpha\beta}^\epsilon g_{\gamma\delta} + C_{\beta\gamma}^\epsilon g_{\alpha\delta} + C_{\gamma\alpha}^\epsilon g_{\beta\delta}). \end{aligned} \quad (17.51)$$

Свернем полученное равенство по индексам γ и ϵ . Поскольку свертка тензора Вейля равна нулю, то получаем тождество

$$\nabla_\gamma C_{\alpha\beta\delta}^\gamma = -\frac{n-3}{n-2} C_{\alpha\beta\delta}. \quad (17.52)$$

Поэтому при $n > 3$ равенство $C_{\alpha\beta\gamma}^\delta = 0$ влечет за собой равенство $C_{\alpha\beta\gamma} = 0$.

С точки зрения дифференцируемости метрики, в рассмотренных выше случаях, $n \geq 3$, существование тензора Коттона является достаточным условием, т.е. $g \in \mathcal{C}^3$.

Рассмотрим случай $n = 2$. В двумерном (псевдо)римановом пространстве из равенства (17.38) при $\bar{R}_{\alpha\beta\gamma\delta} = 0$ следует уравнение на конформный множитель

$$g_{\alpha\beta} \square \phi = R_{\alpha\beta}. \quad (17.53)$$

Его свертка с метрикой приводит к уравнению Пуассона

$$\square \phi = -\frac{1}{2} R. \quad (17.54)$$

Это уравнение, как известно, локально разрешимо при дифференцируемой правой части как для евклидовой, так и для лоренцевой сигнатуры метрики (просто непрерывности не достаточно). При этом правая часть уравнения (17.38) тождественно равна нулю. Следовательно, уравнение (17.53) также выполнено. В этом нетрудно убедиться и прямой проверкой с учетом тождества $R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = 0$.

Следовательно, для локального существования решения уравнения (17.54) достаточно существования дифференцируемой скалярной кривизны, т.е. метрика должна быть трижды непрерывно дифференцируемой функцией $g \in \mathcal{C}^3$.

Таким образом, мы доказали критерий того, что (псевдо)риманово многообразие является локально (псевдо)евклидовым.

Теорема 17.5.1. Пусть задано (псевдо)риманово многообразие $(\mathbb{M}, g)$ с трижды непрерывно дифференцируемой метрикой $g \in \mathcal{C}^3(\mathbb{M})$. Для того, чтобы это многообразие было локально (псевдо)евклидовым необходимо и достаточно выполнение следующих условий:

$$n \geq 4 : \quad C_{\alpha\beta\gamma}^\delta = 0, \quad - \text{тензор Вейля}, \quad (17.55)$$

$$n = 3 : \quad C_{\alpha\beta\gamma} = 0, \quad - \text{тензор Коттона}, \quad (17.56)$$

$$n = 2 : \quad \text{всегда.}$$

Вейль показал, что исчезновение тензора конформной кривизны $C_{\alpha\beta\gamma}^\delta$ является необходимым условием локальной евклидовости [119]. Схоутен доказал, что это условие является также достаточным при $n > 3$ [120]. Он также нашел необходимое и достаточное условие при $n = 3$.

Из определения (17.49) следует антисимметрия тензора Коттона

$$C_{\alpha\beta\gamma} = -C_{\beta\alpha\gamma}.$$

Кроме этого, легко проверяется равенство

$$C_{[\alpha\beta\gamma]} = \frac{1}{3}(C_{\alpha\beta\gamma} + C_{\beta\gamma\alpha} + C_{\gamma\alpha\beta}) = 0. \quad (17.57)$$

Поскольку в трехмерном случае, $n = 3$, полный тензор кривизны взаимно однозначно определяется тензором Риччи (17.24):

$$\begin{aligned} R_{\alpha\beta\gamma\delta} &= \frac{1}{2}\varepsilon_{\alpha\beta\epsilon}\varepsilon_{\gamma\delta\zeta}R^{\epsilon\zeta} = \\ &= \frac{1}{2}\left[-g_{\alpha\gamma}R_{\beta\delta} + g_{\alpha\delta}R_{\beta\gamma} + g_{\beta\gamma}R_{\alpha\delta} - g_{\beta\delta}R_{\alpha\gamma} + (g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma})R\right], \end{aligned} \quad (17.58)$$

то, как нетрудно проверить, тензор Вейля тождественно равен нулю, $C_{\alpha\beta\gamma}{}^{\delta} = 0$.

В двумерном случае полный тензор кривизны взаимно однозначно определяется скалярной кривизной (17.28), где знак $\pm$ соответствует евклидовой и лоренцевой сигнатуре. Поэтому и тензор Вейля, и Коттона тождественно равны нулю, $C_{\alpha\beta\gamma}{}^{\delta} = 0$ и $C_{\alpha\beta\gamma} = 0$. Любая (псевдо)риманова поверхность, $n = 2$, конформно эквивалентна (псевдо)евклидовой плоскости. Следовательно любые две поверхности конформно эквивалентны, если соответствующие метрики имеют одинаковую сигнатуру.

Пример 17.5.1. Если M – пространство постоянной кривизны, тензор кривизны которого имеет вид (15.47), то тензор Вейля, как легко проверить, тождественно равен нулю при $n > 3$. Если $n = 3$, то тензор Коттона также равен нулю. Поэтому все пространства постоянной кривизны такого типа являются конформно (псевдо)евклидовыми. Именно такой вид имеет метрика пространств постоянной кривизны в стереографических координатах (18.28). $\square$

17.6 Преобразование Вейля в аффинной геометрии

В (псевдо)римановой геометрии (M, g) все геометрические объекты однозначно определяются метрикой многообразия g . Поэтому вейлевское преобразование метрики (17.34) определяет вид преобразования компонент связности $\tilde{\Gamma}$ (символов Кристоффеля) и соответствующей кривизны $\tilde{R}$. В аффинной геометрии общего вида (M, g, T, Q) помимо преобразования метрики необходимо дополнительно задать закон преобразования тензора кручения T и неметричности Q . В качестве определения примем следующее условие. При вейлевском преобразовании метрики (17.34), кручение и тензор неметричности преобразуются таким образом, что соответствующая аффинная связность не изменяется:

$$\bar{\Gamma}_{\alpha\beta}{}^{\gamma} = \Gamma_{\alpha\beta}{}^{\gamma}. \quad (17.59)$$

Если индекс γ опустить, то вейлевский закон преобразования примет вид

$$\bar{\Gamma}_{\alpha\beta\gamma} = e^{2\phi}\Gamma_{\alpha\beta\gamma}. \quad (17.60)$$

Поскольку порядок индексов и знаки слагаемых с метрикой и тензором неметричности в выражении для аффинной связности (6.17) одинаковы, а слагаемые с кручением имеют другие знаки, то с помощью алгебраических выкладок можно доказать, что при вейлевских преобразованиях кручение не меняется, а тензор неметричности преобразуется по правилу:

$$\bar{T}_{\alpha\beta}{}^{\gamma} = T_{\alpha\beta}{}^{\gamma} \quad (17.61)$$

$$\bar{Q}_{\alpha\beta\gamma} = e^{2\phi}(Q_{\alpha\beta\gamma} - 2\partial_{\alpha}\phi g_{\beta\gamma}). \quad (17.62)$$

Определение. Преобразования метрики (17.34) вместе с преобразованиями кручения (17.61) и неметричности (17.62) называются *преобразованиями Вейля* в аффинной геометрии общего вида. $\square$

Поскольку связность при этом не меняется (17.59), то тензор кривизны и тензор Риччи инвариантны относительно преобразований Вейля,

$$\bar{R}_{\alpha\beta\gamma}{}^{\delta} = R_{\alpha\beta\gamma}{}^{\delta}, \quad \bar{R}_{\alpha\beta} = R_{\alpha\beta}. \quad (17.63)$$

Соответствующая скалярная кривизна преобразуется как обратная метрика

$$\bar{R} = e^{-2\phi} R. \quad (17.64)$$

В геометрии Римана–Картана–Вейля, когда тензор неметричности имеет специальный вид (6.20), форма Вейля преобразуется по правилу

$$\bar{W}_{\alpha} = W_{\alpha} - 2\partial_{\alpha}\phi, \quad (17.65)$$

которое имеет тот же вид, что и калибровочное преобразование в электродинамике. Согласованность этого преобразования с правилом преобразования связности (17.60) особенно просто проверяется в геометрии Римана–Вейля, когда связность записана в виде (6.23).

Отметим, что вейлевские преобразования в пространстве аффинной связности не меняют геодезических, поскольку не меняют связности. В то же время форма экстремалей меняется, как показано в разделе 17.4.

Преобразования Вейля меняют только метрику и неметричность, не затрагивая кручения. Это означает, что на кручение можно наложить любые ограничения, в частности, положить его равным нулю.

Отметим также, что тензор неметричности является приводимым относительно действия группы $\mathrm{GL}(n, \mathbb{R})$, и преобразование Вейля меняет только одну его неприводимую компоненту – форму Вейля (6.20). Таким образом “минимальная” геометрия, в которой можно рассматривать преобразования Вейля, не затрагивающие аффинной связности, – это геометрия Римана–Вейля.

Преобразования метрики (17.34) и формы (17.65) были рассмотрены Г. Вейлем [56] при попытке объединить общую теорию относительности с электродинамикой.

Если постулировать, что физическая теория, т.е. действие, должна быть инвариантна относительно преобразований Вейля, то действие должно быть безразмерным числом. Поскольку элемент объема в четырехмерном пространстве-времени преобразуется по правилу

$$\sqrt{\bar{g}} = e^{4\phi} \sqrt{g}, \quad (17.66)$$

то лагранжиан теории может содержать слагаемые, квадратичные по кривизне и четвертой степени по кручению и неметричности.

Преобразование Вейля для картановых переменных $e_{\alpha}{}^a$ и $\omega_{\alpha a}{}^b$, введенных в разделе 6.9, определим следующим образом:

$$\begin{aligned} \bar{e}_{\alpha}{}^a &:= e^{\phi} e_{\alpha}{}^a & \Leftrightarrow & \quad \bar{e}^{\alpha}{}_a := e^{-\phi} e^{\alpha}{}_a, \\ \bar{\omega}_{\alpha a}{}^b &:= \omega_{\alpha a}{}^b - \partial_{\alpha}\phi \delta_a^b. \end{aligned} \quad (17.67)$$

Легко проверить, что преобразование картановых переменных (17.67) определяется преобразованием метрики (17.34) и неметричности (17.62) однозначно с точностью до

локальных $\mathrm{GL}(n, \mathbb{R})$ вращений репера. Заметим, что преобразование Вейля (17.67) меняет только симметричную часть линейной связности.

В геометрии Римана–Картана, где неметричность равна нулю, преобразование Вейля метрики невозможно дополнить преобразованием кручения таким образом, чтобы связность осталась без изменения. Если кручение при преобразовании Вейля не меняется, то нетрудно найти соответствующие преобразования компонент аффинной связности, тензора кривизны и других геометрических объектов. Эти формулы мы выписывать не будем.

17.7 Алгебраически независимые компоненты тензора кривизны

Для изучения геометрических свойств (псевдо)риманова пространства важно знать, сколько алгебраически независимых компонент имеет тензор кривизны в фиксированной точке многообразия.

Рассмотрим n -мерное (псевдо)риманово пространство. Как было показано в разделе 6.8, тензор кривизны $R_{\alpha\beta\gamma\delta}$ в этом случае имеет $n^2(n^2 - 1)/12$ независимых компонент. При этом метрика имеет $n(n + 1)/2$ компоненту. Это значит, что в фиксированной точке пространства из тензора кривизны и метрики можно построить не более

$$\frac{n^2(n^2 - 1)}{12} + \frac{n(n + 1)}{2}$$

алгебраически независимых инвариантов. На самом деле независимых инвариантов меньше, т.к. существует свобода в выборе системы координат. В фиксированной точке тензоры умножаются на матрицы Якоби, которые имеют n^2 компонент, и могут быть выбраны произвольным образом. Следовательно, в данной точке пространства число алгебраически независимых инвариантов, построенных из кривизны и метрики, равно

$$\frac{n^2(n^2 - 1)}{12} + \frac{n(n + 1)}{2} - n^2 = \frac{1}{12}n(n - 1)(n - 2)(n + 3). \quad (17.68)$$

Эта утверждение верно при $n \geq 3$.

Случай $n = 2$ является вырожденным, т.к. группа общих преобразований координат содержит подгруппу, которая не меняет $g_{\alpha\beta}$ и $R_{\alpha\beta\gamma\delta}$. Поэтому при $n = 2$ из тензора кривизны можно построить один независимый инвариант (а не нуль). В качестве такого инварианта удобно выбрать скалярную кривизну R .

Если $n = 3$, то из формулы (17.68) следует, что существует три инварианта. В качестве инвариантов можно выбрать, например, корни секулярного уравнения

$$\det(R_{\alpha\beta} - \lambda g_{\alpha\beta}) = 0,$$

или, что эквивалентно, три инварианта:

$$R, \quad R^{\alpha\beta} R_{\alpha\beta}, \quad \frac{\det R_{\alpha\beta}}{\det g_{\alpha\beta}}, \quad (17.69)$$

где $R_{\alpha\beta}$ – тензор Риччи.

Если $n = 4$, то алгебраически независимых инвариантов четырнадцать. Здесь существует много способов выбрать независимые инварианты, которые мы перечислять не будем, поскольку существует много вырожденных случаев.

Формула (17.68) дает число алгебраически независимых инвариантов. Среди этих инвариантов, вообще говоря, имеются связи, содержащие ковариантные производные тензора кривизны. Поэтому число функционально независимых инвариантов тензора кривизны может быть меньше, чем вычисленное по формуле (17.68).

17.8 Классификация Петрова

Важная задача, которая стоит в любой модели гравитации, заключается в том, чтобы установить являются ли две метрики, заданные на одном многообразии своими компонентами, одинаковыми или нет. То есть существует ли такое преобразование координат, которое переводит одну метрику в другую? Явное нахождение преобразования координат сводится к сложной системе дифференциальных уравнений в частных производных, которую можно решить лишь в простейших случаях. Поэтому большое значение имеют конструктивные методы сравнения метрик, которые сводятся к простым вычислениям. Например, можно вычислить скалярные кривизны, которые являются геометрическими инвариантами, для двух метрик. Ясно, что если скалярные кривизны различны, то метрики не могут быть связаны никаким преобразованием координат. Обратное утверждение, конечно, неверно. Классификация Петрова [121] (см. также [122]) является одним из таких конструктивных методов, который сводится к алгебраической классификации тензора Вейля. Классификация Петрова является алгебраической и относится только к четырехмерному пространству-времени.

Изоморфизм $\mathbb{SO}_0(1, 3) \simeq \mathbb{SO}(3, \mathbb{C})$

Сначала рассмотрим пример, который поясняет преобразования тензора Вейля, удобные для классификации Петрова. Пусть задана произвольная 2-форма с компонентами $X_{\alpha\beta} = -X_{\beta\alpha}$ на четырехмерном псевдоримановом многообразии $(\mathbb{M}, g)$ с метрикой сигнатуры $\text{sign } g = (+ - - -)$. Ее также называют ковариантным *бивектором*. Определим дуальную 2-форму

$$Y_{\alpha\beta} := *X_{\alpha\beta} := \frac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}X^{\gamma\delta}.$$

Если применить оператор Ходжа еще раз, то получим равенства (3.53)

$$**X = -X \quad \Leftrightarrow \quad *Y = -X, \quad (17.70)$$

где мы, для простоты, опустили тензорные индексы. Введем комплексную 2-форму

$$Z := X - iY \quad \Leftrightarrow \quad Z_{\alpha\beta} := X_{\alpha\beta} - iY_{\alpha\beta}. \quad (17.71)$$

Применение оператора Ходжа к данному определению приводит к равенству

$$*Z = iZ. \quad (17.72)$$

Обратное преобразование имеет простой вид: $X = \text{re } Z$.

Нетрудно проверить, что комплексно сопряженная 2-форма $\bar{Z} := X + iY$ удовлетворяет равенству

$$*\bar{Z} = -i\bar{Z}. \quad (17.73)$$

Определение. Комплексные 2-формы, удовлетворяющие равенствам (17.72) и (17.73), называются, соответственно, *самодуальными* и *антисамодуальными*. $\square$

Замечание. Соотношение (17.72) не является уравнением на что либо. Если комплексная 2-форма определена равенством (17.71), то формула (17.72) представляет собой тождество, которое выполнено для любой 2-формы X . Для евклидовой сигнатуры метрики также можно ввести комплексную 2-форму с помощью (17.71). Тогда она будет комплексно антисамодуальной, т.к. для римановой метрики $**X = X$. При этом комплексная антисамодуальность также не является уравнением. Для римановой метрики можно ввести понятие самодуальности в вещественной случае. Тогда условие самодуальности будет действительно уравнением. $\square$

Таким образом построен изоморфизм между вещественными 2-формами и самодуальными комплексными 2-формами на четырехмерном псевдоримановом многообразии.

Посчитаем компоненты. Вещественная 2-форма в четырехмерном пространстве-времени имеет 6 независимых компонент. Столько же компонент имеет самодуальная комплексная 2-форма, которые, как мы увидим ниже, можно объединить в комплексный 3-вектор. Поэтому между ними и возможен построенный изоморфизм.

Чтобы построить изоморфизм в явном виде, предположим, что на $\mathbb{M}$ задано времениподобное векторное поле $u = u^\alpha \partial_\alpha$ единичной длины: $u^2 = 1$. Введем операторы

$$\Omega_\alpha^{\beta\gamma} = -\Omega_\alpha^{\gamma\beta} := \frac{1}{2}(\delta_\alpha^\beta u^\gamma - \delta_\alpha^\gamma u^\beta) + \frac{i}{2}\varepsilon_{\delta\alpha}^{\beta\gamma} u^\delta. \quad (17.74)$$

Эти операторы ставят в соответствие каждой комплексной 2-форме комплексную 1-форму по правилу

$$Z_\alpha := \Omega_\alpha^{\beta\gamma} Z_{\beta\gamma} = 2Z_{\alpha\beta} u^\beta. \quad (17.75)$$

Полученная 1-форма обладает следующим очевидным свойством

$$Z_\alpha u^\alpha = 0, \quad (17.76)$$

т.е. лежит в ортогональном дополнении к векторному полю u .

Введем оператор, действующий в обратную сторону:

$$\Omega^{\alpha\beta\gamma}_\delta := g^{\alpha\delta} \Omega_\delta^{\epsilon\zeta} g_{\epsilon\beta} g_{\zeta\gamma}, \quad (17.77)$$

который получается простым поднятием и опусканием индексов. Положим

$$Z_{\alpha\beta} := Z_\gamma \Omega^\gamma_{\alpha\beta} = \frac{1}{2}(Z_\alpha u_\beta - Z_\beta u_\alpha) + \frac{i}{2}\varepsilon_{\alpha\beta\gamma\delta} u^\gamma Z^\delta. \quad (17.78)$$

Предложение 17.8.1. Формула (17.78) является тождеством, если выполнено равенство (17.75).

Доказательство. Подставим выражение (17.75) в правую часть равенства (17.78). Тогда вещественная часть становится тождеством

$$X_{\alpha\beta} = -u_\alpha X_{\beta\gamma} u^\gamma + u_\beta X_{\alpha\gamma} u^\gamma + \varepsilon_{\alpha\beta\gamma\delta} u^\gamma \frac{1}{2} \varepsilon^{\delta\epsilon\zeta\eta} X_{\zeta\eta} u_\epsilon = X_{\alpha\beta},$$

где использована формула свертки антисимметричных тензоров (37.72). Это же верно и для мнимой части. $\square$

Таким образом, по заданной комплексной 2-форме и вещественному времениподобному векторному полю построена комплексная 1-форма. Верно также обратное утверждение: по заданному векторному полю и комплексной 1-форме можно построить соответствующую ей комплексную 2-форму по формуле (17.78). При этом

$$X_{\alpha\beta} = \operatorname{re} Z_{\alpha\beta}, \quad Y_{\alpha\beta} = \operatorname{im} Z_{\alpha\beta}. \quad (17.79)$$

Тем самым доказана

Теорема 17.8.1. *Если на четырехмерном псевдоримановом многообразии (M, g) задано единичное времениподобное векторное поле u , то между вещественной 2-формой $\frac{1}{2}dx^\alpha \wedge dx^\beta X_{\alpha\beta}$ и комплекснозначной 1-формой $dx^\alpha Z_\alpha$, удовлетворяющей условию $Z_\alpha u^\alpha = 0$, существует взаимно однозначное соответствие. Оно задается равенствами (17.75) и (17.79).*

Посчитаем компоненты. 2-форма $X_{\alpha\beta}$ имеет шесть независимых компонент ввиду антисимметрии. Комплексная 1-форма Z_α имеет восемь вещественных компонент, на которые наложено два условия (17.76). Тем самым число независимых компонент 2-формы $X_{\alpha\beta}$ и комплексной 1-формы Z_α совпадает. Поэтому между ними существует изоморфизм, который осуществляется операторами (17.74) и (17.77). Изоморфное отображение зависит от выбора векторного поля u . Например, если $u = (u^0, 0, 0, 0)$, то $Z_0 = 0$. В то же время при $u = (u^0, u^1, 0, 0)$ на компоненты комплексной 1-формы наложено другое условие $Z_0 u^0 + Z_1 u^1 = 0$.

Предложение 17.8.2. *Операторы Ω имеют следующие свойства:*

$$u^\alpha \Omega_\alpha^{\beta\gamma} = 0, \quad (17.80)$$

$$\Omega_\alpha^{\gamma\delta} \Omega_\gamma^\beta = \delta_\alpha^\beta - u_\alpha u^\beta, \quad (17.81)$$

$$\Omega_\epsilon^{\alpha\beta} \Omega_\gamma^\epsilon = \frac{1}{4}(\delta_\gamma^\alpha \delta_\delta^\beta - \delta_\gamma^\beta \delta_\delta^\alpha) + \frac{i}{4}(\varepsilon^{\alpha\beta}{}_\epsilon \gamma u^\epsilon u_\delta - \varepsilon^{\alpha\beta}{}_\delta \epsilon u^\epsilon u_\gamma + \varepsilon^{\epsilon\alpha}{}_\gamma \delta u_\epsilon u^\beta - \varepsilon^{\epsilon\beta}{}_\gamma \delta u_\epsilon u^\alpha). \quad (17.82)$$

Доказательство. Прямая проверка. $\square$

Метрика на M порождает метрику в пространстве 2-форм:

$$X_{\alpha\beta} X^{\alpha\beta} = g^{(\alpha\beta)(\gamma\delta)} X_{\alpha\beta} X_{\gamma\delta},$$

где

$$g^{(\alpha\beta)(\gamma\delta)} := \frac{1}{2}(g^{\alpha\gamma} g^{\beta\delta} - g^{\alpha\delta} g^{\beta\gamma}). \quad (17.83)$$

В свою очередь эта метрика порождает метрику в пространстве комплексных 1-форм

$$h_{\alpha\beta} := \Omega_\alpha^{\gamma\delta} \Omega_\beta^{\epsilon\zeta} g_{(\gamma\delta)(\epsilon\zeta)} = g_{\alpha\beta} - u_\alpha u_\beta, \quad (17.84)$$

что является прямым следствием равенства (17.80). Эта метрика вырождена, т.к. u – нетривиальный собственный вектор с нулевым собственным значением.

Построенное отображение, очевидно, не зависит от выбора системы координат. Для дальнейшего анализа удобно перейти к ортонормированному базису кокасательного пространства $e_a = e^\alpha_a \partial_\alpha$, $a = 0, 1, 2, 3$, (тетраде). Это означает замену греческих индексов в предыдущих формулах на латинские.

Выберем базис e_a таким образом, чтобы вектор u имел только одну отличную от нуля компоненту: $u = (1, 0, 0, 0)$. Тогда метрика (17.84) в пространстве комплексных 1-форм примет вид

$$h_{ab} = \begin{pmatrix} 0 & 0 \\ 0 & \eta_{ij} \end{pmatrix}, \quad \eta_{ij} = -\delta_{ij} = -\text{diag}(1, 1, 1). \quad (17.85)$$

То есть в пространстве комплексных 1-форм с точностью до знака мы имеем обычную евклидову метрику в трехмерном подпространстве.

Пример 17.8.1. Рассмотрим электромагнитное поле в пространстве Минковского. Тензор напряженности электромагнитного поля $F_{ab} = -F_{ba}$ является 2-формой и выражается через напряженность электрического и магнитного полей (25.76). Его дуальный тензор также вычислен (25.157). Выберем единичное времениподобное векторное поле $u = (1, 0, 0, 0)$. Тогда соответствующая комплекснозначная 2-форма имеет компоненты

$$Z_{ab} := F_{ab} - i *F_{ab} = \frac{1}{2} \begin{pmatrix} 0 & -Z_1 & -Z_2 & -Z_3 \\ Z_1 & 0 & i\bar{Z}_3 & -i\bar{Z}_2 \\ Z_2 & -i\bar{Z}_3 & 0 & i\bar{Z}_1 \\ Z_3 & i\bar{Z}_2 & -i\bar{Z}_1 & 0 \end{pmatrix}, \quad (17.86)$$

где

$$Z_a := 2Z_{ab}u^a = 2Z_{a0} \quad \Leftrightarrow \quad Z_0 = 0, \quad Z_i = 2(E_i + iH_i). \quad (17.87)$$

То есть напряженность электромагнитного поля F_{ab} взаимно однозначно отображается на комплексный 3-вектор $\mathbf{E} + i\mathbf{H}$. $\square$

Проведенное выше построение существенно опирается на единичное времениподобное векторное поле. Два таких векторных поля u^a и v^a связаны между собой некоторым преобразованием компонент из группы Лоренца:

$$v^a = u^b S_a^{-1b}, \quad S \in \mathbb{SO}_0(1, 3).$$

При этом 2-форма преобразуется по правилу

$$Z_{ab} = S_a^c S_b^d Z_{cd}.$$

Эта матрица преобразований Лоренца 2-форм $S_a^{[c} S_b^{d]}$ отображается в комплексную матрицу преобразований комплексных 1-форм

$$T_a^b := \Omega_a^{cd} S_c^e S_d^f \Omega_{ef}^b.$$

Если выбран базис так, что $u = (1, 0, 0, 0)$, то

$$\begin{aligned} \Omega_0^{0i} &= \Omega_0^{i0} = \Omega_0^{ij} = 0, \\ \Omega_i^{j0} &= -\Omega_i^{0j} = \frac{1}{2}\delta_i^j, \quad \Omega_i^{jk} = \frac{i}{2}\varepsilon_{oi}^{jk}. \end{aligned} \quad (17.88)$$

Следовательно,

$$\begin{aligned} T_0^0 &= T_0^i = T_i^0 = 0, \\ T_i^j &= \Omega_i^{ab} S_a^c S_b^d \Omega_{cd}^j. \end{aligned}$$

Прямые вычисления приводят к равенству

$$T_i^j = \frac{1}{2}(S_0^0 S_i^j - S_i^0 S_0^j) + \frac{i}{2}(\varepsilon_{0i}^{kl} S_k^j S_l^0 + \varepsilon_0^{jkl} S_i^k S_0^l) + \\ + \frac{1}{4}[\delta_i^j(S_k^k S_l^l - S_k^l S_l^k) + 2S_i^k S_j^k - 2S_i^j S_k^k], \quad (17.89)$$

где подъем и опускание пространственных индексов $i, j, \dots$ осуществляется с помощью отрицательно определенной метрики $\eta_{ij} = -\delta_{ij}$. По построению, матрица T сохраняет метрику (17.85):

$$\eta_{ij} = T_i^k T_j^l \eta_{kl},$$

и, следовательно, принадлежит комплексной группе вращений $\mathbb{SO}(3, \mathbb{C})$. Прямая проверка того факта, что матрица (17.89) ортогональна весьма затруднительна. Нетрудно также вычислить матрицу обратного отображения

$$S_a^{[c} S_b^{d]} = \Omega_{ab}^i T_i^j \Omega_j^{cd}.$$

Таким образом, комплексная матрица (17.89) является образом изоморфного отображения $\mathbb{SO}_0(1, 3) \Leftrightarrow \mathbb{SO}(3, \mathbb{C})$.

Тензор Вейля

Теперь рассмотрим тензор Вейля $C_{\alpha\beta\gamma\delta}$, который задан на четырехмерном псевдоримановом многообразии (M, g) . Он антисимметричен и по первой, и по второй паре индексов. Кроме этого он симметричен относительно перестановки пар индексов. Из предыдущего рассмотрения следует, что тензор Вейля изоморфен комплексной симметричной 4×4 матрице, на которую наложены некоторые условия. Построим этот изоморфизм.

Тензор, дуальный к тензору Вейля, определяется следующим образом

$$*C_{\alpha\beta\gamma\delta} := \frac{1}{2}\varepsilon_{\alpha\beta\epsilon\zeta} C^{\epsilon\zeta}_{\gamma\delta} = \frac{1}{2}\varepsilon_{\alpha\beta\epsilon\zeta} C_{\gamma\delta}^{\epsilon\zeta}. \quad (17.90)$$

Предложение 17.8.3. *Тензор, дуальный к тензору Вейля, обладает следующими симметриями относительно перестановки индексов:*

$$*C_{\alpha\beta\gamma\delta} + *C_{\beta\gamma\alpha\delta} + *C_{\gamma\alpha\beta\delta} = 0, \quad (17.91)$$

$$*C_{\alpha\beta\gamma\delta} = *C_{\gamma\delta\alpha\beta}. \quad (17.92)$$

Доказательство. Вычисляем

$$*C_{\alpha\beta\gamma\delta} \varepsilon^{\beta\delta\epsilon\zeta} = C^{\epsilon\zeta}_{\gamma\alpha},$$

где использованы равенства (37.72). Умножив полученное тождество на $\varepsilon_{\eta\theta\epsilon\zeta}$, получим равенство (17.91). Теперь это равенство применим трижды:

$$*C_{\alpha\beta\gamma\delta} = -*C_{\beta\gamma\alpha\delta} - *C_{\gamma\alpha\beta\delta} = *C_{\gamma\delta\alpha\beta} + *C_{\delta\beta\alpha\gamma} + *C_{\alpha\delta\beta\gamma} + *C_{\delta\gamma\beta\alpha} = \\ = 2 *C_{\gamma\delta\alpha\beta} - *C_{\beta\alpha\delta\gamma}.$$

Перенос последнего слагаемого в левую часть приводит к равенству (17.92). $\square$

Таким образом, тензор (17.90), дуальный к тензору Вейля, обладает всеми симметриями самого тензора Вейля относительно перестановки индексов. Поэтому определение дуального тензора можно переписать в эквивалентном виде

$$*C_{\alpha\beta\gamma\delta} = \frac{1}{2}\varepsilon_{\gamma\delta\epsilon\zeta}C_{\alpha\beta}{}^{\epsilon\zeta}. \quad (17.93)$$

Теперь введем комплексный тензор Вейля

$$Z_{\alpha\beta\gamma\delta} := C_{\alpha\beta\gamma\delta} - i *C_{\alpha\beta\gamma\delta}.$$

Он, очевидно, обладает всеми симметриями тензора Вейля.

Определение. Положим

$$Q_{\alpha\gamma} := \Omega_{\alpha}{}^{\beta\delta}Z_{\beta\delta\epsilon\zeta}\Omega_{\gamma}{}^{\epsilon\zeta} = 4Z_{\alpha\beta\gamma\delta}u^{\beta}u^{\delta} =: E_{\alpha\gamma} + iH_{\alpha\gamma}, \quad (17.94)$$

где мы воспользовались равенством (17.75). Тензоры второго ранга

$$E_{\alpha\gamma} = 4C_{\alpha\beta\gamma\delta}u^{\beta}u^{\delta}, \quad (17.95)$$

$$H_{\alpha\gamma} = 4 *C_{\alpha\beta\gamma\delta}u^{\beta}u^{\delta} \quad (17.96)$$

называются *электрической* и *магнитной компонентами* тензора Вейля $C_{\alpha\beta\gamma\delta}$.

Предложение 17.8.4. Введенные тензоры обладают следующими свойствами:

$$\begin{aligned} E_{\alpha\gamma} &= E_{\gamma\alpha}, & H_{\alpha\gamma} &= H_{\gamma\alpha}, & Q_{\alpha\gamma} &= Q_{\gamma\alpha}, \\ E_{\alpha\gamma}u^{\gamma} &= 0, & H_{\alpha\gamma}u^{\gamma} &= 0, & Q_{\alpha\gamma}u^{\gamma} &= 0, \\ E_{\alpha\gamma}g^{\alpha\gamma} &= 0, & H_{\alpha\gamma}g^{\alpha\gamma} &= 0, & Q_{\alpha\gamma}g^{\alpha\gamma} &= 0. \end{aligned} \quad (17.97)$$

Доказательство. Все свойства являются очевидными следствиями симметрий тензора Вейля (17.14) за исключением симметрии магнитной части, $H_{\alpha\gamma} = H_{\gamma\alpha}$. Последнее свойство следует из равенства (17.92). $\square$

Таким образом тензор Вейля однозначно определяет компоненты тензора $Q_{\alpha\beta}$, которые образуют собой комплексную симметричную 4×4 матрицу с нулевым следом, на которую наложено условие $Q_{\alpha\beta}u^{\beta} = 0$.

Прямые вычисления доказывают обратную формулу

$$\begin{aligned} Z_{\alpha\beta\gamma\delta} &= \Omega_{\alpha\beta}{}^{\epsilon\zeta}Q_{\epsilon\zeta}\Omega_{\gamma\delta}{}^{\eta\theta} = \\ &= \frac{1}{4}(-g_{\alpha\gamma}Q_{\beta\delta} + g_{\alpha\delta}Q_{\beta\gamma} + g_{\beta\gamma}Q_{\alpha\delta} - g_{\beta\delta}Q_{\alpha\gamma}) + \\ &+ \frac{1}{2}(u_{\alpha}u_{\gamma}Q_{\beta\delta} - u_{\alpha}u_{\delta}Q_{\beta\gamma} - u_{\beta}u_{\gamma}Q_{\alpha\delta} + u_{\beta}u_{\delta}Q_{\alpha\gamma}) + \\ &+ \frac{i}{4}(\varepsilon_{\alpha\beta\epsilon\zeta}u^{\epsilon}Q^{\zeta}{}_{\gamma}u_{\delta} - \varepsilon_{\alpha\beta\epsilon\zeta}u^{\epsilon}Q^{\zeta}{}_{\delta}u_{\gamma} + \varepsilon_{\gamma\delta\epsilon\zeta}u^{\epsilon}Q^{\zeta}{}_{\alpha}u_{\beta} - \varepsilon_{\gamma\delta\epsilon\zeta}u^{\epsilon}Q^{\zeta}{}_{\beta}u_{\alpha}). \end{aligned} \quad (17.98)$$

Тем самым тензор $Q_{\alpha\beta}$ однозначно определяет комплексный тензор Вейля и, следовательно, сам тензор Вейля $C_{\alpha\beta\gamma\delta} = \text{re } Z_{\alpha\beta\gamma\delta}$. Тем самым доказана

Теорема 17.8.2. Если на четырехмерном псевдоримановом многообразии $(\mathbb{M}, g)$ задано единичное временноподобное векторное поле u , то между тензором Вейля $C_{\alpha\beta\gamma\delta}$ и комплексным симметричным бесследовым тензором $Q_{\alpha\beta}$, удовлетворяющим условию $Q_{\alpha\beta}u^{\beta} = 0$, существует взаимно однозначное соответствие. Это соответствие задается равенствами (17.94) и (17.98).

Из доказанной теоремы вытекает, что алгебраическая классификация тензора Вейля сводится к классификации комплексной 4×4 матрицы $Q_{\alpha\beta}$ со свойствами (17.97).

Предыдущие построения были выполнены в произвольной системе координат. Для дальнейшего анализа удобно перейти к ортонормальному базису кокасательного пространства $e_a := e^\alpha_a \partial_\alpha$. Тогда комплексная матрица $Q_{ab} := Q_{\alpha\beta} e^\alpha_a e^\beta_b$ удовлетворяет следующим условиям:

$$Q_{ab} = Q_{ba}, \quad Q_{ab} \eta^{ab} = 0, \quad Q_{ab} u^b = 0, \quad (17.99)$$

где $\eta^{ab} = \text{diag}(+ - - -)$.

Выберем ортонормальный базис так, что $u = (1, 0, 0, 0)$. Тогда равенство $Q_{ab} u^b = 0$ примет вид $Q_{a0} = 0$. С учетом симметрии это означает, что в данной системе координат у матрицы Q_{ab} отличен от нуля только пространственный 3×3 блок Q_{ij} , $i, j = 1, 2, 3$.

Любые два единичных времениподобных вектора, направленных в будущее, связаны некоторым преобразованием из собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, 3)$. При этом тензор Вейля преобразуется с помощью той же матрицы лоренцевых вращений, а комплексный тензор Q – с помощью матрицы соответствующего комплексного 3×3 представления (17.89). Тем самым алгебраическая классификация тензора Вейля в данной точке $x \in \mathbb{M}$ сводится к классификации возможного вида комплексной 3×3 матрицы с нулевым следом с точностью до преобразований Лоренца.

Поставим задачу на собственные значения:

$$Q_i^j X_j = \lambda X_i, \quad Q_i^j := Q_{ik} \eta^{kj}.$$

В общем случае, собственные векторы X и собственные значения λ комплексны. Нам необходимо классифицировать матрицу Q с точностью до преобразования подобия $Q \mapsto TQT^{-1}$, где матрица T (17.89) – трехмерное комплексное представление группы Лоренца $\mathbb{SO}_0(1, 3)$.

Для классификации напомним

Определение. Пусть задана $n \times n$ матрица $A \in \text{mat}(n, \mathbb{C})$. Многочлен степени n

$$p_A(t) := \det(1t - A) \quad (17.100)$$

называется *характеристическим*. Нормированный многочлен

$$q_A(t) = t^k + a_{k-1}t^{k-1} + \dots + a_1t + a_0 \quad (17.101)$$

минимальной степени, аннулирующий матрицу A , т.е. такой, что $q_A(A) = 0$, называется *минимальным*. $\square$

Теорема 17.8.3. Для каждой матрицы A существует единственный минимальный многочлен $q_A(t)$. Его степень не превосходит n . Любой многочлен $p(t)$, такой, что $p(A) = 0$, делится на q_A . Подобные матрицы имеют один и тот же минимальный многочлен. Кроме того, $q_A(\lambda) = 0$ тогда и только тогда, когда λ есть собственное значение матрицы A , т.е. $p_A(\lambda) = 0$.

Доказательство. См., например, [23], теорема 3.3.1, следствия 3.3.3 и 3.3.4. $\square$

Следовательно, если характеристический многочлен разложен на множители,

$$p_A(t) = \sum_{i=1}^m (t - \lambda_i)^{s_i}, \quad 1 \leq s_i \leq n, \quad s_1 + s_2 + \dots + s_m = n,$$

где s_i – кратность собственного значения λ_i , и все собственные значения различны, то минимальный многочлен имеет вид

$$q_A(t) = \sum_{i=1}^m (t - \lambda_i)^{r_i}, \quad 1 \leq r_i \leq s_i,$$

т.е. кратность собственных значений может быть ниже. Ясно, что если все собственные значения матрицы A различны, то минимальный и характеристический многочлены совпадают.

Теорема 17.8.4. Пусть $A \in \mathbf{mat}(n, \mathbb{C})$ имеет $m \leq n$ различных собственных значений $\lambda_1, \lambda_2, \dots, \lambda_m$. Тогда матрица A диагонализируема в том, и только в том случае, когда минимальный многочлен имеет вид

$$q_A(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_m)$$

Доказательство. См., например, [23], следствие 3.3.8. □

В общих теоремах матрицы классифицируются с точностью до преобразования подобия $Q \mapsto S Q S^{-1}$, где S – произвольная невырожденная матрица. В рассматриваемом случае матрицей преобразования T является матрица представления элемента группы Лоренца (17.89), а их гораздо меньше. Поэтому канонический вид матрицы Q отличается от того, который приведен в учебниках по линейной алгебре.

Вернемся к классификации Петрова. Поскольку след матрицы $Q = (Q_i^j)$ равен нулю, то сумма собственных значений равна нулю. Поэтому возможны три случая.

Тип I. Все три собственных значения матрицы Q различны: λ_1, λ_2 и $\lambda_3 = -\lambda_1 - \lambda_2$. Тогда характеристический и минимальный многочлены совпадают,

$$p_Q(t) = q_Q(t) = (t - \lambda_1)(t - \lambda_2)(t - \lambda_3),$$

и матрица Q диагонализируема с помощью лоренцева преобразования подобия. Она имеет следующий канонический вид

$$Q = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & -\lambda_1 - \lambda_2 \end{pmatrix} \quad (17.102)$$

В этом случае существует три линейно независимых собственных вектора:

$$X_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Совпадение собственных чисел $\lambda_1 = \lambda_2$ соответствует **типу D**. В этом случае также есть три собственных вектора.

Тензор Вейля типа I параметризуется четырьмя инвариантами: вещественные и мнимые части собственных чисел λ_1 и λ_2 . Это – максимальное число алгебраически

независимых инвариантов. Действительно, тензор Вейля имеет 10 независимых компонент. Существует свобода лоренцевых вращений, которые параметризуются шестью параметрами. Поэтому число независимых инвариантов не может превышать $10 - 6 = 4$.

Для конструктивного вычисления собственных значений при заданной метрике можно воспользоваться следующим способом. Нетрудно проверить равенство

$$Z_{ab}{}^{cd} Z_{cd}{}^{ab} = 2(C_{ab}{}^{cd} C_{cd}{}^{ab} - i C_{ab}{}^{cd} * C_{cd}{}^{ab}). \quad (17.103)$$

С другой стороны, учитывая равенства (17.98) и (17.81), получаем

$$Z_{ab}{}^{cd} Z_{cd}{}^{ab} = \Omega^e{}_{ab} Q_{ef} \Omega^f{}^{cd} \Omega^g{}_{cd} Q_{gh} \Omega^h{}^{ab} = Q_a{}^b Q_b{}^a.$$

Если $u = (1, 0, 0, 0)$, то $Z_{ab}{}^{cd} Z_{cd}{}^{ab} = Q_i{}^j Q_j{}^i$, и можно выразить комплексный инвариант через собственные значения:

$$I_1 := \frac{1}{6} Z_{abcd} Z^{abcd} = \frac{1}{6} (\lambda_1^2 + \lambda_2^2 + \lambda_3^2) = \frac{1}{3} (\lambda_1^2 + \lambda_2^2 + \lambda_1 \lambda_2), \quad (17.104)$$

где использован канонический вид матрицы $\mathbf{Q}$ (17.102). Аналогично, посчитаем инвариант третьего порядка

$$Z_{ab}{}^{cd} Z_{cd}{}^{ef} Z_{ef}{}^{ab} = 4(C_{ab}{}^{cd} C_{cd}{}^{ef} C_{ef}{}^{ab} - i C_{ab}{}^{cd} C_{cd}{}^{ef} * C_{ef}{}^{ab}).$$

С другой стороны,

$$Z_{ab}{}^{cd} Z_{cd}{}^{ef} Z_{ef}{}^{ab} = Q_a{}^b Q_b{}^c Q_c{}^a.$$

Поэтому для второго комплексного инварианта получаем выражение

$$I_2 := -\frac{1}{6} Z_{ab}{}^{cd} Z_{cd}{}^{ef} Z_{ef}{}^{ab} = -\frac{1}{6} (\lambda_1^3 + \lambda_2^3 + \lambda_3^3) = \frac{1}{2} \lambda_1 \lambda_2 (\lambda_1 + \lambda_2). \quad (17.105)$$

Тем самым по заданной метрике можно вычислить два комплексных инварианта I_1 и I_2 , а затем найти собственные значения λ_1 и λ_2 путем решения алгебраических уравнений (17.104), (17.105).

Тип II. Два собственных значения совпадают: λ , λ и -2λ . Тогда характеристический многочлен имеет вид.

$$p_{\mathbf{Q}}(t) = (t - \lambda)^2 (t + 2\lambda).$$

Для минимального многочлена возникают две возможности:

$$q_{\mathbf{Q}}(t) = (t - \lambda)^2 (t + 2\lambda) \quad \text{или} \quad q_{\mathbf{Q}}(t) = (t - \lambda)(t + 2\lambda) \quad (\text{тип D}).$$

В последнем случае, который называется типом **D**, матрица $\mathbf{Q}$ диагонализируема. Канонический вид матрицы $\mathbf{Q}$ следующий

$$\mathbf{Q} = \begin{pmatrix} \lambda - \mu & i\mu & 0 \\ i\mu & \lambda + \mu & 0 \\ 0 & 0 & -2\lambda \end{pmatrix}, \quad (17.106)$$

где $\mu \in \mathbb{R}$ – произвольный вещественный параметр. Если $\mu \neq 0$, то имеется только два линейно независимых собственных вектора:

$$X_1 = \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (17.107)$$

Типу **D** соответствует значение $\mu = 0$.

Для тензора Вейля типа **II** существует два независимых алгебраических инварианта: $\operatorname{re} \lambda$ и $\operatorname{im} \lambda$. При этом $I_1 = \lambda^2$ и $I_2 = \lambda^3$. Поэтому между комплексными инвариантами (17.104) и (17.105) существует зависимость: $I_1^3 - I_2^2$.

При $\lambda = 0$ возникает вырожденный случай, который называется **типом N**. Ему соответствует два независимых собственных вектора (17.107).

Тип III. Все три собственных значения равны нулю. Тогда $p_Q(t) = t^3$, и для минимального многочлена существует три возможности:

$$q_Q(t) = t^3, \quad q_Q(t) = t^2 \quad (\text{тип N}), \quad q_Q(t) = t.$$

Предпоследняя возможность соответствует типу **N**. Последний случай соответствует нулевой матрице **Q** и нулевому тензору Вейля. Матрица **Q** имеет следующий канонический вид

$$Q = \begin{pmatrix} -\mu & i\mu & -\sigma \\ i\mu & \mu & i\sigma \\ -\sigma & i\sigma & 0 \end{pmatrix}, \quad \mu, \sigma \in \mathbb{R}, \quad \mu\sigma \neq 0. \quad (17.108)$$

При $\mu\sigma \neq 0$ существует только один собственный вектор

$$X_1 = \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}.$$

с нулевым собственным значением. При этом $I_1 = 0$ и $I_2 = 0$. Здесь возникает интересная ситуация: пространство-время искривлено, но все инварианты, построенные из тензора Вейля равны нулю.

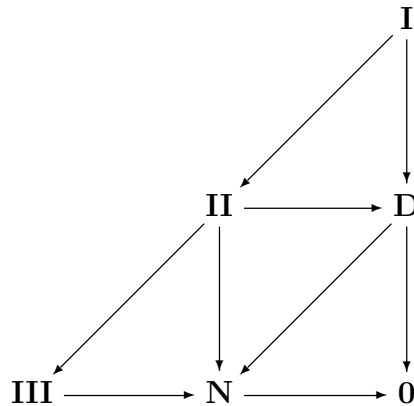
Нетрудно вычислить

$$Q^2 = \begin{pmatrix} \sigma^2 & -i\sigma^2 & 0 \\ -i\sigma^2 & -\sigma^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Q^3 = 0.$$

Мы видим, что при $\sigma = 0$ возникает вырожденный случай, $Q^2 = 0$, который называется **типом N**. Если $\sigma = 0$ и $\mu = 0$ одновременно, то тензор $Q = 0$. Соответствующий тензор Вейля также равен нулю (тип **0**). Если тензор Вейля равен нулю в некоторой области, то согласно теореме 17.5.1 пространство-время является конформно плоским. В этом случае тензор кривизны полностью определяется тензором Риччи.

Таким образом, существует шесть возможностей, которые перечислены в таблице 17.1.

Тип **I** по Петрову является случаем общего положения и все остальные типы возникают из типа **I** в предельных случаях совпадения собственных значений. Иерархию типов по Петрову часто представляют в виде следующей диаграммы



Тип	Минимальный многочлен
I	$(t - \lambda_1)(t - \lambda_2)(t - \lambda_3)$
II	$(t - \lambda)^2(t + 2\lambda)$
D	$(t - \lambda)(t + 2\lambda)$
III	t^3
N	t^2
0	t

Таблица 17.1: Классификация Петрова тензора Вейля.

Типам тензора Вейля, стоящим на одной диагонали, соответствуют минимальные многочлены одной степени. На каждой вертикали расположены типы с одинаковым числом линейно независимых собственных векторов, число которых равно двум для средней вертикальной линии и трем – для правой.

Если так случилось, что матрица $\mathbf{Q}$ для данной метрики оказалась вещественной, то она имеет только простые элементарные делители. Следовательно, тензор Вейля может быть типа **I**, **D** или **0** по Петрову.

Алгебраическая классификация Петрова тензора Вейля относится к фиксированной точке пространства-времени. В принципе, тип тензора Вейля может меняться от точки к точке. Однако все известные автору решения уравнений Эйнштейна имеют фиксированный тип по Петрову всюду в той области, где они определены.

Если решаются вакуумные уравнения Эйнштейна $R_{\alpha\beta} = 0$, то полный тензор кривизны для соответствующей метрики совпадает с тензором Вейля. Поэтому классификация тензора Вейля дает классификацию вакуумных решений уравнений Эйнштейна. Приведем два общих утверждения.

Теорема 17.8.5. *Вакуумное решение уравнений Эйнштейна типа **I**, для которого одно из собственных значений λ равно нулю в некоторой области, является плоским в этой области.*

Доказательство. См. [123]. □

Теорема 17.8.6. *Вакуумные решения уравнений Эйнштейна, которые удовлетворяют уравнению*

$$g^{\epsilon\zeta} \nabla_\epsilon \nabla_\zeta R_{\alpha\beta\gamma\delta} = \omega R_{\alpha\beta\gamma\delta},$$

*относятся либо к типу **N** ($\omega = 0$), либо к типу **D** ($\omega \neq 0$).*

Доказательство. См. [124, 125]. □

Пример 17.8.2. Решение Керра – осесимметричное решение вакуумных уравнений Эйнштейна – имеет тип **D** по классификации Петрова во всей области определения. В частном случае, решение Шварцшильда также типа **D**. □

Пример 17.8.3. Метрика Фридмана (35.70) для однородной и изотропной вселенной имеет вид

$$g = \begin{pmatrix} 1 & 0 \\ 0 & a^2 g_{\mu\nu} \end{pmatrix}, \quad \mu, \nu, = 1, 2, 3, \quad (17.109)$$

где $a(t) > 0$ – масштабный множитель и $\overset{\circ}{g}_{\mu\nu}$ – метрика трехмерного пространства постоянной кривизны. Компоненты тензора кривизны для этой метрики уже были вычислены (35.72):

$$R_{0\mu 0\nu} = \frac{\ddot{a}}{a} \overset{\circ}{g}_{\mu\nu}, \quad R_{\mu\nu\rho\sigma} = \left(\frac{\dot{a}^2}{a^2} - \frac{K}{a^2} \right) (\overset{\circ}{g}_{\mu\rho} \overset{\circ}{g}_{\nu\sigma} - \overset{\circ}{g}_{\mu\sigma} \overset{\circ}{g}_{\nu\rho}), \quad (17.110)$$

где $K = \text{const}$ – гауссова кривизна пространственных сечений. Соответствующий тензор Риччи с скалярная кривизна имеют следующие нетривиальные компоненты:

$$R_{00} = 3\frac{\ddot{a}}{a}, \quad R_{\mu\nu} = \left(\frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2} - \frac{2K}{a^2} \right) \overset{\circ}{g}_{\mu\nu}, \quad R = 6 \left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} - \frac{K}{a^2} \right).$$

Прямые вычисления компонент тензора Вейля (17.14) показывают, что все они равны нулю. Следовательно, метрика Фридмана относится к типу **0** по классификации Петрова. $\square$

По построению, классификация Петрова тензора Вейля является алгебраической и не зависит от выбора системы координат. Это значит, что если две метрики относятся к разным типам по Петрову, то они неэквивалентны. Если же они относятся к одному типу по Петрову, то вопрос остается открытым, и для их различия требуется привлечение более тонких методов.

17.9 Преобразование Эйнштейна–Кауфман

Аффинная геометрия общего вида допускает локальные преобразования связности, которые не меняют тензора кривизны и формы геодезических. Эти преобразования рассматривались Эйнштейном и Кауфман при построении единой теории поля [126, 127].

Пусть на $\mathbb{M}$ задана аффинная геометрия.

Определение. *Преобразованием Эйнштейна–Кауфман* называется следующее локальное преобразование компонент связности, при этом метрика остается без изменения:

$$g'_{\alpha\beta} = g_{\alpha\beta}, \quad (17.111)$$

$$\Gamma'_{\alpha\beta}{}^{\gamma} = \Gamma_{\alpha\beta}{}^{\gamma} + \partial_{\alpha}\phi\delta_{\beta}^{\gamma}, \quad (17.112)$$

где $\phi = \phi(x)$ – произвольная достаточно гладкая функция. $\square$

Используя определение тензора кривизны (6.74), можно проверить, что преобразование связности (17.112) оставляет полный тензор кривизны без изменения

$$R'_{\alpha\beta\gamma}{}^{\delta} = R_{\alpha\beta\gamma}{}^{\delta}.$$

Поскольку преобразование симметричной части связности имеет вид (16.9), то в силу теоремы 16.1.1 преобразование Эйнштейна–Кауфман меняет только параметризацию геодезических, не изменяя их форму. Так как метрика инвариантна при преобразовании (17.111), то экстремали при этом также инвариантны.

Получим закон преобразования кручения и неметричности при преобразовании Эйнштейна–Кауфман. Из определения (6.15) следует, что компоненты тензора кручения преобразуются по правилу

$$T'_{\alpha\beta}{}^{\gamma} = T_{\alpha\beta}{}^{\gamma} + \partial_{\alpha}\phi\delta_{\beta}^{\gamma} - \partial_{\beta}\phi\delta_{\alpha}^{\gamma}. \quad (17.113)$$

Подставляя закон преобразования кручения (17.113) и связности (17.111) в выражение для связности (6.17), получим закон преобразования компонент тензора неметричности

$$Q'_{\alpha\beta\gamma} = Q_{\alpha\beta\gamma} + 2\partial_\alpha \phi g_{\beta\gamma}. \quad (17.114)$$

При преобразовании Эйнштейна–Кауфман метрика и, следовательно, элемент объема не меняются. Это значит, что любой лагранжиан, построенный только по тензору кривизны будет инвариантен относительно локального преобразования (17.111), (17.112).

Нетрудно проверить, что преобразование Эйнштейна–Кауфман коммутирует с преобразованием Вейля, и результат их совместного действия на неметричность имеет вид

$$\bar{Q}'_{\alpha\beta\gamma} = e^{2\phi_1} (Q_{\alpha\beta\gamma} + 2g_{\beta\gamma} \partial_\alpha (\phi_2 - \phi_1)),$$

где ϕ_1 и ϕ_2 – параметры преобразования Вейля и Эйнштейна–Кауфман, соответственно. При $\phi_1 = \phi_2$ тензор неметричности просто умножается на конформный множитель. Отсюда следует, что совместное действие преобразования Вейля и Эйнштейна–Кауфман с одним и тем же параметром можно рассматривать в геометрии Римана–Картана, положив неметричность равной нулю. Это преобразование имеет вид (17.34), (17.112) и (17.113) и не приводит к появлению неметричности. Впервые оно рассматривалось в [128, 129]. Совместное преобразование не меняет тензора кривизны и формы геодезических. Экстремали при этом меняются.

Глава 18

Симметрические пространства

Теория симметрических пространств или пространств постоянной кривизны является классическим разделом дифференциальной геометрии и играет важнейшую роль в приложениях. Достаточно отметить, что к симметрическим пространствам относятся сферы, различные гиперboloиды, группы Ли и другие важные многообразия. В настоящее время симметрические пространства классифицированы, но эта классификация сложна, и мы ее описывать не будем. В начальных разделах главы будут даны определения и описаны общие свойства симметрических пространств. В заключительном разделе будут подробно описаны псевдосферы, которые представляют собой обширный класс (псевдо)римановых пространств постоянной кривизны, которые особенно важны в математической физике.

18.1 Определения и общие свойства

Рассмотрим аффинное многообразие (M, Γ) , где Γ – аффинная связность, произвольной размерности n . Наличие метрики на многообразии M пока не предполагаем. Каждая связность определяет геодезические линии, которые позволяют построить нормальную систему координат в некоторой окрестности U_p произвольной точки $p \in M$ (см. раздел 16.9.4). А именно, геодезические линии определяют экспоненциальное отображение

$$\exp X : T_p(M) \ni X \mapsto (\exp X)p \in M, \quad (18.1)$$

произвольного вектора из некоторой окрестности $N_p \subset T_p(M)$ нулевого вектора касательного пространства в некоторую окрестность U_p точки p . При этом экспоненциальное отображение является диффеоморфизмом.

Определение. Симметрией s_p в точке $p \in M$ называется диффеоморфизм окрестности U_p на себя такой, что отображение $\exp X$ переходит в $\exp(-X)$. $\square$

Если есть две симметрии, определенные в окрестностях U_p и V_p , то они совпадают в пересечении $U_p \cap V_p$, и можно говорить о симметрии в точке. Пусть x^α – нормальная система координат в окрестности точки p , тогда симметрия s_p отражает точки:

$$s_p : U_p \ni (x^1, \dots, x^n) \mapsto (-x^1, \dots, -x^n) \in U_p. \quad (18.2)$$

Это отображение является инволюцией, т.к. $s_p^2 = \text{id}$. Из формулы (18.2) вытекает, что дифференциал отображения s_p равен $-\text{id}_p$, где id_p – тождественное преобразование касательного пространства $T_p(M)$.

Определение. Если симметрия s_p является аффинным преобразованием, т.е. не меняет аффинную связность Γ , в окрестности каждой точки $p \in \mathbb{M}$, то $\mathbb{M}$ называется *аффинным локально симметрическим пространством* (многообразием). Аффинное пространство $(\mathbb{M}, \Gamma)$ называется *аффинным симметрическим*, если симметрия s_p может быть продолжена до глобального аффинного преобразования $\mathbb{M}$. Аффинные симметрические пространства называются также *пространствами постоянной кривизны*. $\square$

Предложение 18.1.1. На аффинном локально симметрическом пространстве любое тензорное поле нечетного ранга r , инвариантное относительно преобразования симметрии s_p для всех $p \in \mathbb{M}$, тождественно равно нулю.

Доказательство. Так как дифференциал отображения симметрии s_p равен $-\text{id}_p$, то произвольный тензор K ранга r преобразуется в тензор $(-1)^r K$. С другой стороны, он инвариантен относительно преобразования симметрии s_p . Для нечетных r такое возможно только при $K = 0$. $\square$

Теорема 18.1.1. Многообразие $\mathbb{M}$ с аффинной связностью Γ является аффинным локально симметрическим пространством тогда и только тогда, когда выполнены равенства

$$T_{\alpha\beta}{}^\gamma = 0, \quad \nabla_\epsilon R_{\alpha\beta\gamma}{}^\delta = 0, \quad (18.3)$$

где $T_{\alpha\beta}{}^\gamma$ и $R_{\alpha\beta\gamma}{}^\delta$ – тензоры кривизны и кручения аффинной связности Γ .

Доказательство. Необходимость. Пусть $\mathbb{M}$ – аффинное локально симметрическое пространство. Так как s_p – аффинное преобразование, оно не меняет связности и, следовательно, не меняет тензор кручения и кривизны. С другой стороны, тензор кручения является тензором третьего ранга, а ковариантная производная $\nabla_\epsilon R_{\alpha\beta\gamma}{}^\delta$ – тензором ранга 5. Поэтому из предложения (18.1.1) следуют равенства (18.3).

Доказательство утверждения теоремы в обратную сторону сложнее. Оно приведено, например, в [48], глава XI, теорема 1.1. $\square$

Напомним, что равенство нулю ковариантных производных тензора кривизны означает, что он инвариантен относительно параллельных переносов вдоль произвольных кривых на $\mathbb{M}$. Поэтому аффинные симметрические пространства называются также *пространствами постоянной кривизны*.

Пример 18.1.1. Для полных аффинных связностей с нулевыми кручением и кривизной пара $(\mathbb{M}, \Gamma)$ является аффинным симметрическим пространством. Эти пространства являются локально (псевдо)евклидовыми. При этом, если база $\mathbb{M}$ не является односвязным многообразием, то связность Γ может иметь нетривиальную группу голономии и в этом смысле являться нетривиальной. $\square$

Доказательство следующих четырех теорем содержится в [48], глава XI, §1.

Теорема 18.1.2. Каждое аффинное локально симметрическое пространство есть вещественно аналитическое многообразие с вещественно аналитической связностью по отношению к атласу, состоящему из нормальных координатных систем.

Теорема 18.1.3. Полное односвязное аффинное локально симметрическое пространство является аффинным симметрическим.

Теорема 18.1.4. Каждое аффинное симметрическое пространство полно.

Доказательство. Пусть $\gamma = x(t)$, $0 \leq t \leq a$, – геодезическая из точки p в точку q . Используя симметрию s_p , геодезическую можно продолжить за точку q следующим образом:

$$x(a+t) := s_q(x(a-t)), \quad \forall \quad 0 \leq t \leq a. \quad \square$$

Теорема 18.1.5. *На каждом связном аффинном симметрическом пространстве группа аффинных преобразований транзитивна.*

Из последней теоремы следует, что каждое аффинное симметрическое пространство является однородным пространством (см. раздел 9.1). Из теоремы 9.1.2 и предложения 9.1.4 вытекает, что $\mathbb{M} = \mathbb{G}/\mathbb{H}$, где $\mathbb{G}$ связная группа Ли и $\mathbb{H} \subset \mathbb{G}$ – ее подгруппа изотропии в некоторой точке $p \in \mathbb{M}$. Пусть s_p – симметрия в точке p . Определим автоморфизм группы Ли $\mathbb{G}$ следующим образом

$$\sigma(a) := s_p \circ a \circ s_p^{-1}, \quad \forall a \in \mathbb{G}. \quad (18.4)$$

Ясно, что этот автоморфизм инволютивный: $\sigma^2(a) = a$.

Теорема 18.1.6. *Пусть $\mathbb{G}$ – наибольшая связная группа аффинных преобразований аффинного симметрического пространства $\mathbb{M}$, и $\mathbb{H} \subset \mathbb{G}$ – подгруппа изотропии в фиксированной точке $p \in \mathbb{M}$, так что $\mathbb{M} = \mathbb{G}/\mathbb{H}$. Пусть s_p – симметрия для $\mathbb{M}$ в точке p и σ – автоморфизм $\mathbb{G}$ (18.4). Тогда, если $\mathbb{G}^\sigma$ – замкнутая подгруппа в $\mathbb{G}$, состоящая из всех элементов, неподвижных при действии σ , то $\mathbb{H}$ лежит между $\mathbb{G}^\sigma$ и связной компоненты единицы $\mathbb{G}_0^\sigma$ группы $\mathbb{G}^\sigma$, т.е. $\mathbb{G}_0^\sigma \subset \mathbb{H} \subset \mathbb{G}^\sigma$.*

Доказательство. Пусть $h \in \mathbb{H}$, и рассмотрим $\sigma(h) := s_p \circ h \circ s_p^{-1}$. Так как дифференциал отображения s_p в точке p равен $-\text{id}_p$, то дифференциал отображения $\sigma(h)$ совпадает с дифференциалом отображения h в точке p . Можно доказать, что два аффинных преобразования с одинаковыми дифференциалами в одной точке совпадают друг с другом (см. [130], глава X, §1). Следовательно $\mathbb{H} \subset \mathbb{G}^\sigma$.

С другой стороны. Пусть a_t – произвольная однопараметрическая подгруппа в $\mathbb{G}^\sigma$. Тогда справедливы равенства:

$$s_p \circ a_t(p) = a_t s_p(p) = a_t(p).$$

Отсюда следует, что орбита $a_t(p)$ поточечно неподвижна при действии симметрии s_p . Так как p есть изолированная неподвижная точка для s_p , то орбита $a_t(p)$ состоит из одной точки, т.е. $a_t(p) = p$. Отсюда следует, что $a_t \in \mathbb{H}$. Поскольку связная группа Ли порождается ее однопараметрическими подгруппами, то выполнено включение $\mathbb{G}_0^\sigma \subset \mathbb{H}$. $\square$

Доказанная теорема позволяет дать следующее

Определение. *Симметрическим пространством называется тройка $(\mathbb{G}, \mathbb{H}, \sigma)$, состоящая из связной группы Ли $\mathbb{G}$, ее замкнутой подгруппы $\mathbb{H}$ и инволютивного автоморфизма σ для $\mathbb{G}$ такого, что $\mathbb{G}_0^\sigma \subset \mathbb{H} \subset \mathbb{G}^\sigma$.* $\square$

Пусть задана тройка $(\mathbb{G}, \mathbb{H}, \sigma)$. Она определяет симметрию s_p в каждой точке $p \in \mathbb{G}/\mathbb{H}$ так, что p – изолированная неподвижная точка. Пусть s_p есть инволютивный диффеоморфизм из $\mathbb{G}/\mathbb{H}$ на себя, который порожден инволютивным диффеоморфизмом σ группы $\mathbb{G}$. Допустим, что $a(p)$ – неподвижная точка для s_p , где

$a \in \mathbb{G}$. Это значит, что $\sigma(a) \in a\mathbb{H}$. Положим $h := a^{-1}\sigma(a) \in \mathbb{H}$. Так как $\sigma(h) = h$, то справедливы равенства:

$$h^2 = h\sigma(h) = a^{-1}\sigma(a)\sigma(a^{-1}\sigma(a)) = a^{-1}\sigma(a)\sigma(a^{-1})a = \text{id},$$

т.е. h^2 – единичный элемент. Если a достаточно близко к единице, так, что h тоже близко к единице, то h само должно быть единицей. Отсюда следует равенство $\sigma(a) = a$. Поскольку a инвариантно относительно действия автоморфизма σ и близко к единице, то a лежит в компоненте единицы $\mathbb{G}_0^\sigma$ и потому в $\mathbb{H}$. Это влечет равенство $a(p) = p$, которое доказывает, что p – изолированная неподвижная точка. Для $x := a(p)$ положим $s_x := a \circ s_p \circ a^{-1}$. Тогда s_x не зависит от выбора a такого, что $x = a(p)$.

Каждое симметрическое пространство $(\mathbb{G}, \mathbb{H}, \sigma)$ порождает *симметрическую алгебру Ли* $(\mathfrak{g}, \mathfrak{h}, \sigma)$ естественным образом, где $\mathfrak{g}$ и $\mathfrak{h}$ – алгебры Ли для $\mathbb{G}$ и $\mathbb{H}$ соответственно, а инволютивный автоморфизм σ для $\mathfrak{g}$ порождается автоморфизмом σ в $\mathbb{G}$. Обратно. Если $(\mathfrak{g}, \mathfrak{h}, \sigma)$ – симметрическая алгебра Ли и если $\mathbb{G}$ – связная односвязная группа Ли (универсальная накрывающая) с алгеброй Ли $\mathfrak{g}$, то автоморфизм σ для $\mathfrak{g}$ индуцирует автоморфизм σ для $\mathbb{G}$ (см. [50], глава VI, §6, теорема 2), и для любой подгруппы $\mathbb{H}$, лежащей между $\mathbb{G}^\sigma$ и компонентой единицы $\mathbb{G}_0^\sigma$, тройка $(\mathbb{G}, \mathbb{H}, \sigma)$ есть симметрическое пространство. Отметим, что поскольку $\mathbb{G}^\sigma$ замкнута, то $\mathbb{H}$ также замкнута. Если группа Ли $\mathbb{G}$ неодносвязна, то автоморфизм σ алгебры Ли может и не индуцировать автоморфизм для $\mathbb{G}$.

Пусть $(\mathfrak{g}, \mathfrak{h}, \sigma)$ – симметрическая алгебра Ли. Так как отображение σ инволютивно, то его собственными значениями являются 1 и -1 , и $\mathfrak{h}$ – собственное подпространство для собственного значения 1. Обозначим собственное подпространство для -1 через $\mathfrak{m}$. Тогда разложение алгебры Ли в прямую сумму (как векторных пространств)

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$$

называется *каноническим разложением* для симметрической алгебры $(\mathfrak{g}, \mathfrak{h}, \sigma)$. Каноническое разложение всегда существует, что следует из разложения

$$X = \frac{1}{2}(1 + \sigma)X + \frac{1}{2}(1 - \sigma)X, \quad X \in \mathfrak{g},$$

где

$$\frac{1}{2}(1 + \sigma)X \in \mathfrak{h}, \quad \frac{1}{2}(1 - \sigma)X \in \mathfrak{m}.$$

Предложение 18.1.2. *Если $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ – каноническое разложение симметрической алгебры Ли $(\mathfrak{g}, \mathfrak{h}, \sigma)$, то выполнены включения:*

$$[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}, \quad [\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{m}, \quad [\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{h}. \quad (18.5)$$

Доказательство. Первое включение означает, что $\mathfrak{h}$ – подалгебра. Если $X \in \mathfrak{h}$ и $Y \in \mathfrak{m}$, то

$$\sigma([X, Y]) = [\sigma(X), \sigma(Y)] = [X, -Y] = -[X, Y] \in \mathfrak{m}.$$

Если $X, Y \in \mathfrak{m}$, то

$$\sigma([X, Y]) = [\sigma(X), \sigma(Y)] = [-X, -Y] = [X, Y] \in \mathfrak{h},$$

что доказывает третье включение. □

Таким образом, алгебра Ли $\mathfrak{g}$ является $\mathbb{Z}_2$ градуированной.

Предложение 18.1.3. Пусть $(\mathbb{G}, \mathbb{H}, \sigma)$ – симметрическое пространство, и $(\mathfrak{g}, \mathfrak{h}, \sigma)$ – его симметрическая алгебра. Если $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ – ее каноническое разложение, то

$$\mathrm{ad}(\mathbb{H})\mathfrak{m} \subset \mathfrak{m}.$$

Доказательство. Пусть $X \in \mathfrak{m}$ и $a \in \mathbb{H}$. Тогда

$$\sigma(\mathrm{ad} a \cdot X) = \mathrm{ad} \sigma(a) \cdot \sigma(X) = \mathrm{ad}(a) \cdot (-X) = -\mathrm{ad} a \cdot X \in \mathfrak{m}. \quad \square$$

Симметрические пространства обобщают понятие группы Ли в следующем смысле.

Пример 18.1.2. Пусть $\mathbb{G}$ – связная группа Ли, и $\Delta\mathbb{G}$ – ее диагональ, т.е. множество элементов вида

$$\Delta\mathbb{G} := \{(a, a) \in \mathbb{G} \times \mathbb{G} : a \in \mathbb{G}\}.$$

Определим автоморфизм

$$\sigma : \mathbb{G} \times \mathbb{G} \ni (a, b) \mapsto \sigma(a, b) := (b, a) \in \mathbb{G} \times \mathbb{G}.$$

Тогда тройка $(\mathbb{G} \times \mathbb{G}, \Delta\mathbb{G}, \sigma)$ является симметрическим пространством. Факторпространство $(\mathbb{G} \times \mathbb{G})/\Delta\mathbb{G}$ диффеоморфно $\mathbb{G}$. Этот диффеоморфизм индуцируется отображением

$$\mathbb{G} \ni a \mapsto (a, e) \in \mathbb{G} \times \mathbb{G},$$

где e – единица группы Ли $\mathbb{G}$.

Аналогично каждая алгебра Ли $\mathfrak{g}$ порождает симметрическую алгебру Ли $(\mathfrak{g} \oplus \mathfrak{g}, \Delta\mathfrak{g}, \sigma)$. $\square$

Таким образом, показано, что каждое аффинное симметрическое пространство $(\mathbb{M}, \Gamma)$ порождает симметрическое пространство $(\mathbb{G}, \mathbb{H}, \sigma)$. Верно также обратное утверждение.

Теорема 18.1.7. Пусть $(\mathbb{G}, \mathbb{H}, \sigma)$ – симметрическое пространство. Каноническая связность Γ на $\mathbb{M} := \mathbb{G}/\mathbb{H}$ есть единственная аффинная связность, которая инвариантна относительно симметрий для $\mathbb{M}$. Каноническая связность существует.

Доказательство. См., например, [48], глава XI, теорема 3.1. $\square$

Каноническая связность имеет следующие свойства.

Теорема 18.1.8. Пара $(\mathbb{M}, \Gamma)$, где Γ есть каноническая связность на $\mathbb{M} := \mathbb{G}/\mathbb{H}$ для симметрического пространства $(\mathbb{G}, \mathbb{H}, \sigma)$, представляет собой полное аффинное симметрическое пространство с симметриями $s_x := a \circ s_p \circ a^{-1}$, где $a \in \mathbb{G}$ и $x = a(p)$, обладающее следующими свойствами:

1) справедливы равенства:

$$T = 0, \quad \nabla R = 0, \quad R(X, Y)Z = -[X, Y], Z], \quad (18.6)$$

где, для краткости, опущены тензорные индексы и $X, Y, Z \in \mathfrak{m}$ (при этом $\mathfrak{m}$ отождествляется с касательным пространством $T_p(\mathbb{M})$ при помощи дифференциала проекции $\mathbb{G} \rightarrow \mathbb{G}/\mathbb{H}$);

- 2) для каждого $X \in \mathfrak{m}$ параллельный перенос вдоль кривой $(\exp tX)p$ совпадает с дифференциалом экспоненциального отображения $\exp tX$;
- 3) для каждого $X \in \mathfrak{m}$ кривая $(\exp tX)p$ есть геодезическая, исходящая из точки p , и, обратно, каждая геодезическая из p имеет такой вид;
- 4) каждое $\mathbb{G}$ -инвариантное тензорное поле на $\mathbb{M}$ параллельно;
- 5) алгебра Ли линейной группы голономии в опорной точке p порождается множеством

$$\{R(X, Y) = -\operatorname{ad}_{\mathfrak{m}}([X, Y]), \quad \forall X, Y \in \mathfrak{m}\}.$$

Доказательство. См., например, [48], глава XI, теорема 3.2. $\square$

Если тензорное поле K инвариантно относительно действия группы $\mathbb{G}$, то каждая симметрия s_x отображает K либо в K , либо в $-K$ в зависимости от того, четен ранг тензора или нечетен.

Таким образом, мы видим, что между (полными) аффинными симметрическими пространствами $(\mathbb{M}, \Gamma)$ и симметрическими пространствами $(\mathbb{G}, \mathbb{H}, \sigma)$ со связной и односвязной группой Ли $\mathbb{G}$ существует взаимно однозначное соответствие. Поэтому задача классификации аффинных симметрических пространств сводится к алгебраической задаче классификации троек $(\mathbb{G}, \mathbb{H}, \sigma)$ (универсальных накрывающих). Эта задача тесно примыкает к задаче классификации простых алгебр Ли, кратко описанной в разделе 8.16, и в настоящее время решена (см., например, [48], глава XI и [66], глава X).

По сути дела, симметрические и однородные пространства, рассмотренные в разделе 9.1, – это одно и то же. Разница сводится к следующему. При определении симметрического пространства исходным объектом является многообразие с заданной на нем связностью, $(\mathbb{M}, \Gamma)$, которая обладает симметрией. В то время как исходным объектом для определения однородного пространства была группа преобразований $(\mathbb{M}, \mathbb{G})$ с определенными свойствами. Соответствие между парами $(\mathbb{M}, \Gamma)$ и $(\mathbb{M}, \mathbb{G})$ устанавливается путем отождествления группового действия с параллельным переносом.

18.2 (Псевдо)римановы симметрические пространства

В приложениях важнейшую роль играет обширный класс симметрических пространств специального вида, включающий сферы и различные гиперboloиды.

Определение. (Псевдо)риманово пространство $(\mathbb{M}, g)$ есть (псевдо)риманово локально симметрическое пространство, если оно является аффинным локально симметричным пространством для аффинной связности $\tilde{\Gamma}$, построенной по метрике g (символов Кристоффеля). Пара $(\mathbb{M}, g)$ называется (псевдо)римановым симметрическим пространством, если оно является аффинным симметрическим для связности $\tilde{\Gamma}$. $\square$

Чтобы подчеркнуть, что аффинная связность $\tilde{\Gamma}$ построена по метрике g при нулевом кручении и тензоре неметричности, мы, как и в разделе 6.2, употребим знак тильды.

Описание этих пространств начнем с общего свойства.

Предложение 18.2.1. Пусть $(\mathbb{M}, g)$ – (псевдо)риманово локально симметрическое пространство. Тогда симметрия s_p для аффинной связности $\tilde{\Gamma}$ является изометрией при всех $p \in \mathbb{M}$.

Доказательство. См., например, [48], глава XI, предложение 6.1. $\square$

Теоремы для аффинных симметрических пространств прямо переносятся на (псевдо)римановы симметрические пространства. Приведем их без доказательств.

Теорема 18.2.1. *(Псевдо)риманово многообразие является (псевдо)римановым локально симметричным пространством тогда и только тогда, когда тензорное поле кривизны параллельно, т.е.*

$$\tilde{\nabla}_\epsilon \tilde{R}_{\alpha\beta\gamma}{}^\delta = 0. \quad (18.7)$$

Теорема 18.2.2. *Полное односвязное (псевдо)риманово локально симметрическое пространство является (псевдо)римановым симметрическим.*

Теорема 18.2.3. *(Псевдо)риманово симметрическое пространство полно.*

Теорема 18.2.4. *Пусть (M, g) – (псевдо)риманово симметрическое пространство, G – наибольшая связная группа изометрий для M , и H – подгруппа изотропии в точке $p \in M$. Пусть s_p – симметрия для $(M, \tilde{G})$ в точке p , и σ – инволютивный автоморфизм для G , определенный так:*

$$\sigma(a) := s_p \circ a \circ s_p^{-1}, \quad \forall a \in G.$$

Пусть G^σ – замкнутая подгруппа в G , состоящая из элементов, неподвижных относительно действия σ , со связной компонентой единицы G_0^σ . Тогда

- 1) группа G действует транзитивно на M , так что $M = G/H$;*
- 2) подгруппа H лежит между G^σ и ее связной компонентой единицы: $G_0^\sigma \subset H \subset G^\sigma$;*
- 3) для римановых симметрических пространств подгруппа H компактна.*

Доказательство. См., например, [48], глава XI, теорема 6.5. $\square$

Так же, как и для аффинных симметрических пространств, (псевдо)римановы симметрические пространства (M, g) порождают симметрические пространства (G, H, σ) . При этом справедливы следующие утверждения.

Теорема 18.2.5. *Пусть (G, H, σ) симметрическое пространство. Тогда G -инвариантная (псевдо)риманова метрика на $M := G/H$, если такая существует, индуцирует каноническую связность на M .*

Доказательство. См., например, [48], глава XI, теорема 3.3. $\square$

Существует широкий и важный класс симметрических пространств, которые допускают инвариантные (псевдо)римановы метрики.

Теорема 18.2.6. *Пусть (G, H, σ) – симметрическое пространство с полупростой группой G , и пусть $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ – каноническое разложение. Тогда сужение формы Киллинга–Картана с $\mathfrak{g}$ на $\mathfrak{m}$ определяет G -инвариантную (псевдо)риманову метрику на G/H .*

Доказательство. См., например, [48], глава XI, теорема 3.4. $\square$

Предложение 18.2.2. *Пусть (G, H, σ) – симметрическое пространство с полупростой группой G . Тогда каноническое разложение $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ ортогонально относительно метрики Киллинга–Картана.*

Доказательство. Пусть $X \in \mathfrak{h}$ и $Y \in \mathfrak{m}$. Тогда из коммутационных соотношений (18.5) вытекают включения:

$$\begin{aligned} \operatorname{ad} X(\mathfrak{h}) &\subset \mathfrak{h}, & \operatorname{ad} X(\mathfrak{m}) &\subset \mathfrak{m}, \\ \operatorname{ad} Y(\mathfrak{h}) &\subset \mathfrak{m}, & \operatorname{ad} Y(\mathfrak{m}) &\subset \mathfrak{h}. \end{aligned}$$

Отсюда следует, что $\operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y) = 0$. $\square$

Из данного предложения вытекает, что форма Киллинга–Картана в базисе, соответствующем каноническому разложению $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, является блочно диагональной

$$g = \begin{pmatrix} g_{\mathfrak{h}} & 0 \\ 0 & g_{\mathfrak{m}} \end{pmatrix},$$

где $g_{\mathfrak{h}}$ и $g_{\mathfrak{m}}$ – метрики на $\mathfrak{g}$ и $\mathfrak{m}$.

Касательное пространство $T_p(\mathbb{M})$ в точке p естественным образом отождествляется с векторным подпространством $\mathfrak{m} \subset \mathfrak{g}$. Действительно, определена каноническая проекция группы на свою факторгруппу $\operatorname{pr} : \mathbb{G} \rightarrow \mathbb{G}/\mathbb{H}$. Тогда дифференциал этого отображения действует из алгебры Ли $\operatorname{pr}_* : \mathfrak{g} \rightarrow \mathfrak{m}$. Поскольку $\mathbb{M} = \mathbb{G}/\mathbb{H}$, то возникает естественное отождествление $T_p(\mathbb{M}) \sim \mathfrak{m}$.

Симметрия метрики в точке p относительно действия группы изотропии $\mathbb{H}$ означает, что метрика инвариантна относительно действия присоединенного представления группы изотропии: $X \mapsto X_h := \operatorname{ad} h(X)$, где $X \in \mathfrak{m}$ и $h \in \mathbb{H}$. Условие инвариантности метрики означает сохранение скалярного произведения

$$(X, Y) = (X_h, Y_h). \quad (18.8)$$

В координатах это условие записывается следующим образом. Если $\eta_{\alpha\beta}$ – метрика в некотором фиксированном базисе $\mathfrak{m}$, то условие (18.8) эквивалентно уравнению

$$\eta_{\alpha\beta} = S_{\alpha}^{\gamma} S_{\beta}^{\delta} \eta_{\gamma\delta},$$

где $S_{\alpha}^{\gamma}(h)$ – матрица присоединенного представления. Для бесконечно малых преобразований $h = \exp(\epsilon A)$, $A \in \mathfrak{m}$ и $\epsilon \ll 1$, условие принимает вид

$$([A, X], Y) + (X, [A, Y]) = 0$$

или, в координатах,

$$f_{\alpha\beta}^{\delta} \eta_{\delta\gamma} + f_{\alpha\gamma}^{\delta} \eta_{\beta\delta} = 0,$$

где $f_{\alpha\beta}^{\gamma}$ – структурные константы подгруппы $\mathbb{H} \subset \mathbb{G}$.

Пример 18.2.1. Рассмотрим евклидово пространство $\mathbb{R}^n$. Декартова система координат x^{α} является нормальной, и в ней задана евклидова метрика $\delta_{\alpha\beta}$. Симметрией s_p является отражение декартовых координат $x^{\alpha} \mapsto -x^{\alpha}$. Это отображение является аффинным преобразованием и изометрией. Поэтому евклидово пространство является римановым симметрическим пространством нулевой кривизны. Группой изометрий евклидова пространства является неоднородная группа вращений $\mathbb{ISO}(n)$ (максимальная связная группа), которая представляет собой полупрямое произведение группы вращений $\mathbb{SO}(n)$ на группу трансляций $\mathbb{R}^n$. Подгруппа изотропии совпадает с группой вращений $\mathbb{SO}(n)$. Поэтому $\mathbb{R}^n = \mathbb{ISO}(n)/\mathbb{SO}(n)$. Алгебра Ли группы Ли $\mathbb{ISO}(n)$ разлагается в прямую сумму:

$$\mathfrak{iso}(n) = \mathfrak{so}(n) \oplus \mathfrak{t}^n,$$

где $\mathfrak{so}(n)$ – алгебра Ли группы вращений и $\mathfrak{t}^n = \mathbb{R}^n$ алгебра группы трансляций, которая, как многообразие, совпадает с евклидовым пространством. Алгебра прямого произведения является $\mathbb{Z}_2$ -градуированной (18.5), где $\mathfrak{h} = \mathfrak{so}(n)$ и $\mathfrak{m} = \mathfrak{t}^n$. При этом алгебра трансляций коммутативна $[\mathfrak{t}^n, \mathfrak{t}^n] = 0$. $\square$

Пример 18.2.2. Псевдоевклидово пространство $\mathbb{R}^{p,q}$, так же как и евклидово пространство, является псевдоримановым симметрическим пространством нулевой кривизны. Необходимо только заменить группу вращений $\mathrm{SO}(n)$ на связную компоненту единицы группы псевдовращений $\mathrm{SO}_0(p, q)$. Поэтому $\mathbb{R}^{p,q} = \mathrm{ISO}_0(p, q)/\mathrm{SO}_0(p, q)$. Для пространства Минковского $\mathbb{R}^{1,3}$ алгебра Ли группы Пуанкаре, которая является группой изометрий, рассмотрена в разделе 1.9.1, является $\mathbb{Z}_2$ -градуированной и имеет вид (18.5). $\square$

Равенство нулю всех ковариантных производных тензора кривизны (18.7) часто выбирают в качестве определения (псевдо)римановых пространств постоянной кривизны, предполагая, конечно, что кручение тождественно равно нулю. Не все уравнения (18.3) являются независимыми в силу тождеств Бианки (6.116). Тем не менее ковариантное постоянство тензора кривизны является сильным требованием, которое, как будет показано ниже, в частном случае приводит к однородности и изотропии пространства.

Из равенства (18.7) следует, что любое скалярное поле, построенное из метрики и тензора кривизны, равно константе, т.к. ковариантная производная функции совпадает с обычной частной производной. Например, скалярная кривизна и квадрат тензора кривизны,

$$\tilde{R} = g^{\alpha\beta} R_{\alpha\beta} = \text{const}, \quad \tilde{R}^{\alpha\beta\gamma\delta} \tilde{R}_{\alpha\beta\gamma\delta} = \text{const},$$

постоянны на пространствах постоянной кривизны.

В современной математической физике чаще всего рассматриваются пространства постоянной кривизны специального вида. А именно, пусть тензор кривизны (псевдо)риманова пространства определяется метрикой с точностью до постоянной:

$$\tilde{R}_{\alpha\beta\gamma\delta} = -K(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}), \quad K = \text{const}, \quad (18.9)$$

где в правой части комбинация компонент метрики подобрана таким образом, чтобы симметрии по индексам обеих частей равенства совпадали. Поскольку метрика ковариантно постоянна, т.е. $\tilde{\nabla}_\alpha g_{\beta\gamma} = 0$, то равенство (18.7) выполнено, и, следовательно, (псевдо)риманово пространство с тензором кривизны (18.9) является пространством постоянной кривизны. В дальнейшем под пространством постоянной кривизны, как правило, мы будем понимать (псевдо)риманово многообразие с тензором кривизны вида (18.9), не оговаривая этого явно. Соответствующий тензор Риччи и скалярная кривизна равны

$$\tilde{R}_{\alpha\beta} = -K(n-1)g_{\alpha\beta}, \quad \tilde{R} = -Kn(n-1).$$

В двумерном случае гауссова кривизна K (произведение главных кривизин поверхности, вложенной в трехмерное евклидово пространство) пропорциональна скалярной кривизне $\tilde{R}$ этой поверхности с индуцированной метрикой. Коэффициент пропорциональности можно фиксировать, рассмотрев сферу $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ единичного радиуса. Мы покажем, что для нее $\tilde{R} = -2K$.

В следующем разделе будет доказано, что n -мерная сфера $\mathbb{S}_a^n \hookrightarrow \mathbb{R}^{n+1}$ радиуса a является пространством постоянной кривизны вида (18.9) с $K = 1/a^2$. Поэтому мы

будем называть функцию $K(x)$, заданную на $\mathbb{M}$ равенством

$$\tilde{R} =: -Kn(n-1), \quad (18.10)$$

гауссовой кривизной. Она обладает тем свойством, что для сферы единичного радиуса $\mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1}$ она равна единице, $K = 1$, в произвольном числе измерений. Будем считать, что для пространств постоянной положительной кривизны $K > 0$, а для пространств отрицательной кривизны $K < 0$. Отметим, что деление пространств постоянной кривизны на пространства положительной и отрицательной кривизны является содержательным только в случае положительно определенных метрик.

Замечание. Пространства постоянной кривизны с метрикой вида (18.9) удовлетворяют вакуумным уравнениям Эйнштейна (23.5) с космологической постоянной $\Lambda = -\kappa K(n-1)$. В общей теории относительности эти решения называются решениями де Ситтера, которое представляет собой фактор пространство $\mathbb{SO}_0(1, 4)/\mathbb{SO}_0(1, 3)$, при $\Lambda > 0$ (раздел 35.4.2) и анти-де Ситтера, которое представляет собой фактор пространство $\mathbb{SO}_0(2, 3)/\mathbb{SO}_0(1, 3)$, при $\Lambda < 0$ (раздел 35.4.3). $\square$

Пример 18.2.3. Полупростую группу Ли $\mathbb{G}$, $\dim \mathbb{G} = N$, можно рассматривать, как (псевдо)риманово пространство с метрикой, определяемой формой Киллинга–Картана (см. раздел 8.5). Тогда ее тензор кривизны равен

$$\tilde{R}_{ABC}{}^D = \frac{1}{4} f_{AB}{}^E f_{CE}{}^D, \quad A, B, \dots = 1, \dots, N. \quad (18.11)$$

где $f_{AB}{}^C$ – структурные константы группы Ли, и ковариантная производная от него тождественно обращается в нуль. Отсюда следуют выражения для тензора Риччи и скалярной кривизны:

$$\tilde{R}_{AB} = -\frac{1}{4} \eta_{AB}, \quad \tilde{R} = -\frac{N}{4}.$$

С этой точки зрения полупростые группы Ли являются пространствами постоянной кривизны, т.к. выполнено равенство $\tilde{\nabla}_E \tilde{R}_{ABC}{}^D = 0$. Отметим, что тензор кривизны (18.11) не может быть сведен к виду (18.9). Отсюда следует, что в случае (псевдо)римановых пространств условие (18.9) является более сильным, чем равенство (18.3). $\square$

Предложение 18.2.3. Тензор Вейля (17.14) и тензор Коттона (17.49) для пространств постоянной кривизны вида (18.9) равны нулю.

Доказательство. Прямая проверка. $\square$

Согласно теореме 17.5.1 из этого утверждения следует, что любое пространство постоянной кривизны является конформно плоским. Соответствующая система координат построена в разделе 18.3 и называется стереографической.

Определение. Пусть на многообразии $\mathbb{M}$ задано двумерное направление простым отличным от нуля бивектором

$$V^{\alpha\beta} := \frac{1}{2}(X^\alpha Y^\beta - Y^\alpha X^\beta),$$

где векторные поля X и Y линейно независимы всюду на $\mathbb{M}$. Предположим, что двумерные направления не являются изотропными

$$V^{\alpha\beta} V_{\alpha\beta} = g_{(\alpha\beta)(\gamma\delta)} V^{\alpha\beta} V^{\gamma\delta} \neq 0, \quad g_{(\alpha\beta)(\gamma\delta)} := \frac{1}{2}(g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma}).$$

Тогда функция

$$S(x) := \frac{R_{\alpha\beta\gamma\delta} V^{\alpha\beta} V^{\gamma\delta}}{g_{(\alpha\beta)(\gamma\delta)} V^{\alpha\beta} V^{\gamma\delta}}, \quad (18.12)$$

называется *кривизной (псевдо)риманова пространства в данной точке $x \in \mathbb{M}$ в данном двумерном направлении $V^{\alpha\beta}(x)$* . $\square$

Если подставить в определение (18.12) тензор кривизны (18.9), то получим равенство

$$S = -2K, \quad (18.13)$$

т.е. кривизна пространства постоянной кривизны во всех точках и в каждом двумерном направлении постоянна и пропорциональна гауссовой кривизне K .

18.3 Псевдосферы

В настоящем разделе мы рассмотрим класс пространств постоянной кривизны, который очень часто встречается в приложениях. Эти многообразия будут построены в виде гиперповерхностей в (псевдо)евклидовом пространстве.

Рассмотрим $(n+1)$ -мерное (псевдо)евклидово пространство $\mathbb{R}^{p,q}$ с метрикой произвольной сигнатуры

$$\eta_{\alpha\beta} := \text{diag}(\underbrace{+\cdots+}_p \underbrace{-\cdots-}_q), \quad 1 \leq p \leq n, \quad q = n+1-p, \quad (18.14)$$

заданной в декартовой системе координат y^α , $\alpha = 0, 1, \dots, n$. Зададим в $\mathbb{R}^{p,q}$ гиперповерхность $\mathbb{H}_a^n$, размерности n , алгебраическим уравнением

$$\eta_{\alpha\beta} y^\alpha y^\beta = \pm a^2, \quad (18.15)$$

где $a > 0$ – произвольная постоянная. Эта гиперповерхность называется *псевдосферой*. Частным случаем псевдосфер являются обычные сферы $\mathbb{S}_a^n$, вложенные в евклидово пространство $\mathbb{R}^{n+1}$. Для них a – это радиус сферы, и в правой части уравнения (18.15) необходимо выбрать знак $+$, т.к. в противном случае уравнение не будет иметь решений. Если пространство псевдоевклидово, то знак правой части может быть любым, и разным знакам соответствуют различные гиперболоиды. При этом гиперболоиды могут состоять из нескольких компонент связности. Мы будем употреблять термин “псевдосфера”, имея в виду различные возможности, в том числе сферы. При $a = 0$ уравнение (18.15) выделяет в псевдоевклидовом пространстве конусы, которые представляют самостоятельный интерес, но здесь рассматриваться не будут, поскольку не являются многообразиями.

В последующих выкладках знак $\pm$ в правой части уравнения (18.15) мы опустим, но будем помнить, что все выкладки можно провести и для отрицательного знака правой части, если метрика не является положительно определенной. Кроме того, будем считать, что все подкоренные выражения, которые встретятся ниже, положительны, не описывая явно ту область, в которой определены координаты. В каждом конкретном случае это легко можно проделать.

Метрика (18.14) инвариантна относительно неоднородной группы вращений $\mathbb{IO}(p, q)$, состоящей из (псевдо)вращений, сдвигов и отражений. (Псевдо)вращениям и сдвигам соответствуют векторы Киллинга:

$$L_{\alpha\beta} = y_\beta \partial_\alpha - y_\alpha \partial_\beta, \quad P_\alpha = \partial_\alpha, \quad (18.16)$$

где $y_\alpha := y^\beta \eta_{\beta\alpha}$. В дальнейшем, для краткости, мы будем писать просто “вращения”.

Уравнение псевдосферы (18.15) инвариантно относительно вращений $\mathbb{SO}(p, q)$ и отражений.

Вложение псевдосферы $\mathbb{H}_a^n \hookrightarrow \mathbb{R}^{p,q}$, заданное уравнением (18.15), индуцирует на $\mathbb{H}_a^n$ метрику g . В результате мы получаем псевдориманово многообразие $(\mathbb{H}_a^n, g)$, обладающее высокой степенью симметрии. Поскольку и метрика (18.14), и уравнение (18.15) инвариантны относительно вращений и отражений, то эти преобразования являются изометриями псевдосфер.

При желании все отражения можно описать явно, но мы на этом останавливаться не будем.

Первым шагом является нахождение индуцированной метрики. В качестве координат на псевдосфере выберем n координат $\{y^\mu\} := \{y^1, \dots, y^n\}$, которые будем обозначать буквами из середины греческого алфавита $\mu, \nu, \dots$. Координата y^0 находится из уравнения (18.15)

$$y^0 = \pm \sqrt{a^2 - y^2}, \quad y^2 := \eta_{\mu\nu} y^\mu y^\nu, \quad (18.17)$$

где

$$\eta_{\mu\nu} := \text{diag} \left(\underbrace{+\dots+}_{p-1} \underbrace{-\dots-}_q \right) \quad (18.18)$$

– “пространственная” часть метрики (18.14). Псевдосфера $\mathbb{H}_a^n$ является нетривиальным многообразием и в общем случае покрывается несколькими картами, которым, в частности, соответствует разный знак перед квадратным корнем в равенстве (18.17). Область определения координат y^μ задается условием положительности подкоренного выражения. Она различна для разных значений a^2 и разных сигнатур метрики (18.18), и в каждом конкретном случае ее нетрудно описать.

Из выражения (18.17) следует выражение для дифференциала

$$dy^0 = \mp \frac{y_\mu dy^\mu}{\sqrt{a^2 - y^2}}.$$

Поэтому метрика, индуцированная на псевдосфере, в координатах y^μ принимает вид

$$ds^2 = dy_\mu dy^\mu + \frac{(y_\mu dy^\mu)^2}{a^2 - y^2}. \quad (18.19)$$

Введем сферические координаты в $\mathbb{R}^{p-1,q} \hookrightarrow \mathbb{R}^{p,q}$: $\{y^\mu\} \mapsto \{r, \chi^1, \dots, \chi^{n-1}\}$, где r – радиальная координата и χ обозначает угловые координаты в (псевдо)евклидовом пространстве $\mathbb{R}^{p-1,q}$. Тогда интервал и уравнение вложения псевдосферы (18.15) примут вид

$$ds^2 = (dy^0)^2 + dr^2 + r^2 d\Omega, \quad (18.20)$$

$$(y^0)^2 + r^2 = a^2, \quad (18.21)$$

где $d\Omega(\chi, d\chi)$ – угловая часть (псевдо)евклидовой метрики, явный вид которой не является существенным. Нам важно только то, что он существует. Уравнение (18.21) приводит к равенствам:

$$y^0 = \pm \sqrt{a^2 - r^2} \quad \Rightarrow \quad dy^0 = \mp \frac{r dr}{\sqrt{a^2 - r^2}}.$$

Подстановка dy^0 в выражение (18.20) дает индуцированную метрику

$$ds^2 = \frac{a^2 dr^2}{a^2 - r^2} + r^2 d\Omega. \quad (18.22)$$

Конечно, это та же метрика (18.19), только в сферической системе координат. Теперь совершим преобразование радиальной координаты $r \mapsto \rho$,

$$r := \frac{\rho}{1 + \frac{\rho^2}{4a^2}} \quad \Rightarrow \quad dr = \frac{1 - \frac{\rho^2}{4a^2}}{\left(1 + \frac{\rho^2}{4a^2}\right)^2} d\rho.$$

Тогда индуцированная метрика становится конформно (псевдо)евклидовой

$$ds^2 = \frac{d\rho^2 + \rho^2 d\Omega}{\left(1 + \frac{\rho^2}{4a^2}\right)^2}. \quad (18.23)$$

Возвращаясь к декартовым координатам $\{\rho, \chi^1, \dots, \chi^{n-2}\} \mapsto \{x^\mu\}$, получим метрику

$$ds^2 = \frac{\eta_{\mu\nu} dx^\mu dx^\nu}{\left(1 + \frac{x^2}{4a^2}\right)^2}, \quad (18.24)$$

где $x^2 := \eta_{\mu\nu} x^\mu x^\nu = \rho^2$.

Определение. Координаты $\{\rho, \chi^1, \dots, \chi^{n-1}\}$ или $\{x^\mu\}$, $\mu = 1, \dots, n$, в которых метрика псевдосферы $\mathbb{H}_a^n$ является конформно псевдоевклидовой (18.23), (18.24), называются *стереографическими*. $\square$

Приступим к вычислению геометрических объектов на псевдосфере. Вычисления, пожалуй, проще всего выглядят в стереографических координатах (18.24). Метрика и ее обратная являются диагональными:

$$g_{\mu\nu} = \frac{1}{(1 + bx^2)^2} \eta_{\mu\nu}, \quad g^{\mu\nu} = (1 + bx^2)^2 \eta^{\mu\nu}, \quad b := \frac{1}{4a^2}, \quad (18.25)$$

где $\eta^{\mu\rho} \eta_{\rho\nu} = \delta_\nu^\mu$.

Соответствующие символы Кристоффеля равны

$$\Gamma_{\mu\nu\rho} = -\frac{2b}{(1 + bx^2)^3} (x_\mu \eta_{\nu\rho} + x_\nu \eta_{\mu\rho} - x_\rho \eta_{\mu\nu}),$$

$$\Gamma_{\mu\nu}{}^\rho = -\frac{2b}{1 + bx^2} (x_\mu \delta_\nu^\rho + x_\nu \delta_\mu^\rho - \eta_{\mu\nu} x^\rho),$$

где в правых частях формул подъем и опускание индексов производится с помощью плоской метрики: $x_\mu := \eta_{\mu\nu} x^\nu$. Теперь по формуле (6.88) нетрудно вычислить полный тензор кривизны

$$R_{\mu\nu\rho\sigma} = -\frac{4b}{(1 + bx^2)^4} (\eta_{\mu\rho} \eta_{\nu\sigma} - \eta_{\mu\sigma} \eta_{\nu\rho}) = -4b (g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho}). \quad (18.26)$$

Свертка тензора кривизны с обратной метрикой дает тензор Риччи и скалярную кривизну:

$$R_{\mu\nu} = -4b(n-1)g_{\mu\nu}, \quad R = -4bn(n-1) = -\frac{1}{a^2}n(n-1). \quad (18.27)$$

Это выражение при $b = K/4$ совпадает с уравнением (18.9).

Таким образом, каждая псевдосфера является пространством постоянной кривизны. Двумерные и трехмерные псевдосферы построены и проанализированы в разделах 29 и 35.2. Среди псевдосфер есть связные и односвязные многообразия, например, сферы S_a^n . Есть многообразия состоящие из нескольких компонент связности, например, двуполостный гиперболоид, описанный в разделе 29.2. Он состоит из двух компонент связности, каждая из которых является связным и односвязным многообразием. Часть псевдосфер являются связными, но не односвязными многообразиями, например, однополостный гиперболоид из раздела 29.3. Для таких псевдосфер нетрудно построить универсальное накрывающее пространство. В этом смысле псевдосферы дают все универсальные накрывающие пространства постоянной кривизны вида (18.9). Тогда все остальные пространства постоянной кривизны указанного типа получаются из псевдосфер как фактор пространства по группе преобразований, действующей свободно и собственнo разрывно. Поэтому задача классификации пространств постоянной кривизны вида (18.9) сводится к описанию групп преобразований псевдосфер.

Таким образом доказана следующая

Теорема 18.3.1. *На любом пространстве постоянной кривизны (18.9) в некоторой окрестности произвольной точки существует система координат (стереографическая), в которой метрика имеет вид*

$$ds^2 = \frac{\eta_{\mu\nu} dx^\mu dx^\nu}{(1 + bx^2)^2}, \quad x^2 := x^\mu x^\nu \eta_{\mu\nu}, \quad (18.28)$$

где $\eta_{\mu\nu}$ – диагональная матрица (плоская метрика) (18.18), и $K = 4b = \text{const}$.

Доказательство. Любое пространство постоянной кривизны является фактор пространством своей универсальной накрывающей по группе преобразований, действующей свободно и собственнo разрывно. Выше мы доказали, что стереографические координаты существуют для всех универсальных накрывающих. Таким образом, теорема доказана. Доказательство существования стереографических координат было также доказано путем явного построения (15.54). $\square$

Стереографические координаты для пространств постоянной кривизны определены там, где выполнено равенство $(1 + bx^2) > 0$. В этой области все компоненты метрики являются вещественно аналитическими.

Следствие. Все пространства постоянной кривизны вида (18.9) с одинаковой сигнатурой метрики и скалярной кривизной локально изометричны. $\square$

Обсудим область определения метрики пространства постоянной кривизны (18.28). Для пространств нулевой кривизны $b = 0$, и стереографические координаты совпадают с декартовыми. Таким образом, пространство нулевой кривизны – это (псевдо)евклидово пространство в зависимости от сигнатуры метрики (см. примеры 18.2.1 и 18.2.2). Оно полно, связно и односвязно, т.е. является универсальным накрывающим пространством для всех пространств нулевой кривизны, которые являются фактор пространствами евклидова пространства по группам преобразований, действующих собственнo разрывно и свободно (см. раздел 9.1). Это все возможные цилиндры и торы (см. раздел 1.3.4).

Для сфер $b > 0$ и $x^2 > 0$. Поэтому метрика (18.28) также определена во всем евклидовом пространстве $\{x^\mu\} \in \mathbb{R}^n$. Следовательно, многообразие связно и односвязно. При $|\mathbf{x}| := \sqrt{(x^1)^2 + \dots + (x^n)^2} \rightarrow \infty$, метрика быстро убывает. Это приводит к тому, как можно проверить, что многообразие является геодезически неполным при $|\mathbf{x}| \rightarrow \infty$. Для того, чтобы получить сферу необходимо компактифицировать евклидово пространство, добавив бесконечно удаленную точку.

Область определения координат для гиперболоидов $\mathbb{H}_a^n$ зависит от сигнатуры метрики и знака правой части определяющего уравнения (18.15). В каждом конкретном случае это легко проделать.

Теперь обсудим псевдосферы $\mathbb{H}_a^n$ с точки зрения общей теории, развитой в предыдущих разделах. Симметрия s_0 относительно начала координат задается простым отражением, $s_0 : (x^\mu) \mapsto (-x^\mu)$. Максимальной связной группой изометрий является связная компонента единицы группы вращений $\mathbb{SO}_0(p+1, q)$. Подгруппой изотропии начала координат является группа $\mathbb{SO}_0(p, q)$. Следовательно, псевдосфера является факторпространством

$$\mathbb{H}_a^n \approx \frac{\mathbb{SO}_0(p+1, q)}{\mathbb{SO}_0(p, q)}. \quad (18.29)$$

Точнее, псевдосфера диффеоморфна факторпространству при любом фиксированном a .

Размерность группы изометрий псевдосферы $\mathbb{H}_a^n$ равна

$$\dim \mathbb{SO}_0(p+1, q) = \frac{n(n+1)}{2}.$$

Согласно теореме 15.3.2 эта размерность максимальна, и псевдосфера является однородным и изотропным пространством.

Глава 19

Симплектические и пуассоновы многообразия

Симплектические и пуассоновы многообразия играют важную роль в дифференциальной геометрии в связи с применениями, в первую очередь, к гамильтоновой динамике, рассмотренной в главе 21.

19.1 Симплектические группы

Симплектические группы играют в симплектической геометрии ту же роль, что и группы вращений в евклидовой геометрии. Поскольку эти группы устроены намного сложнее, чем ортогональные и унитарные матричные группы, то мы посвятим симплектическим группам и их свойствам целый раздел. Более подробное изложение содержится в [131].

Определение. Рассмотрим антисимметричную $2n \times 2n$ матрицу (*каноническую симплектическую форму*),

$$\varpi = (\varpi_{\alpha\beta}) := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad (19.1)$$

где $\alpha, \beta = 1, \dots, 2n$ и $\mathbb{1}$ – единичная $n \times n$ матрица. Эта матрица определяет билинейную квадратичную форму в евклидовом пространстве $\mathbb{R}^{2n}$ с базисом e_α , рассматриваемом, как векторное пространство,

$$\varpi : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \ni X, Y \mapsto \varpi(X, Y) := X^\alpha Y^\beta \varpi_{\alpha\beta} \in \mathbb{R},$$

где X^α и Y^β – компоненты векторов в выбранной системе координат. Квадратные матрицы A размера $2n \times 2n$ с вещественными элементами, оставляющие каноническую симплектическую форму ϖ инвариантной,

$$A^T \varpi A = \varpi, \quad (19.2)$$

образуют группу Ли $\mathrm{SP}(n, \mathbb{R})$, которая называется *вещественной симплектической группой*. $\square$

Взятие определителя от обеих частей определения (19.2) приводит к равенству $\det A = \pm 1$, поскольку $\det \varpi = 1$. Отсюда следует, что для любой матрицы A существует обратная. Нетрудно убедиться, что обратная матрица также является симплектической, а также в том, что произведение двух симплектических матриц снова дает симплектическую матрицу. Тем самым все групповые аксиомы выполнены.

Каноническая симплектическая форма универсальна в следующем смысле. Если ω – произвольная невырожденная антисимметричная матрица, то из курса линейной алгебры известно, что за счет линейного преобразования базиса в $\mathbb{R}^{2n}$ ее всегда можно преобразовать к каноническому виду (19.1).

Предложение 19.1.1. *Алгебра Ли $\mathfrak{sp}(n, \mathbb{R})$ группы $\mathbb{SP}(n, \mathbb{R})$ состоит из матриц вида*

$$\begin{pmatrix} B & C \\ D & -B^T \end{pmatrix} \in \mathfrak{sp}(n, \mathbb{R}), \quad (19.3)$$

где B – произвольная вещественная $n \times n$ матрица, а вещественные $n \times n$ матрицы C и D симметричны.

Доказательство. Вблизи единицы группы симплектическая матрица представима в виде

$$A = e^{tM} = \mathbb{1} + tM + \dots, \quad t \ll 1,$$

где $M \in \mathfrak{sp}(n, \mathbb{R})$ – элемент алгебры Ли. Подставляя это разложение в (19.2) в линейном по t порядке получаем равенство

$$M^T \varpi + \varpi M = 0. \quad (19.4)$$

Представим элемент алгебры в блочном виде

$$M = \begin{pmatrix} B & C \\ D & E \end{pmatrix}.$$

Тогда из (19.4) следуют равенства:

$$C = C^T, \quad D = D^T, \quad E = -B^T \quad \square$$

Следствие. Размерность симплектической группы $\mathbb{SP}(n, \mathbb{R})$ равна

$$\dim \mathbb{SP}(n, \mathbb{R}) = n^2 + 2 \frac{n(n+1)}{2} = n(2n+1). \quad \square$$

Выше было отмечено, что $\det A = \pm 1$. Справедливо более сильное утверждение.

Предложение 19.1.2. *Определитель любой симплектической матрицы равен единице*

$$\det A = 1, \quad \forall A \in \mathbb{SP}(n, \mathbb{R}). \quad (19.5)$$

Доказательство. Любую симплектическую матрицу $A = (A_\alpha^\beta) \in \mathbb{SP}(n, \mathbb{R})$, можно рассматривать, как невырожденное линейное преобразование $2n$ -мерного евклидова пространства $\mathbb{R}^{2n}$, которое оставляет каноническую 2-форму ϖ инвариантной. Следовательно, симплектическое преобразование сохраняет инвариантной и любую внешнюю степень формы ϖ . В частности, симплектическое преобразование сохраняет $2n$ -форму объема

$$\varpi^n := \underbrace{\varpi \wedge \dots \wedge \varpi}_n.$$

Известно, что при преобразовании координат форма объема умножается на якобиан преобразования координат. Таким образом, для симплектических преобразований имеем равенство

$$\varpi^n = \varpi^n \det A.$$

Отсюда следует равенство (19.5). $\square$

Симплектические группы $\mathbb{SP}(n, \mathbb{R})$ существенно отличаются от групп вращений евклидова пространства. В частности, они являются некомпактными. Продемонстрируем это на примере простейшей группы $\mathbb{SP}(1, \mathbb{R})$.

Теорема 19.1.1. *Группа $\mathbb{SP}(1, \mathbb{R})$ изоморфна группе 2×2 матриц $\mathbb{SL}(2, \mathbb{R})$. С топологической точки зрения эта группа трехмерна некомпактна и гомеоморфна прямому произведению окружности на двумерную плоскость, $\mathbb{SP}(1, \mathbb{R}) \approx \mathbb{S}^1 \times \mathbb{R}^2$. Она неодносвязна, и ее фундаментальная группа изоморфна группе целых чисел*

$$\pi_1(\mathbb{SP}(1, \mathbb{R})) \approx \mathbb{Z}.$$

Доказательство. В двумерном случае каноническая симплектическая форма с точностью до знака совпадает с полностью антисимметричным тензором $\varpi_{\alpha\beta} = -\varepsilon_{\alpha\beta}$ (37.69). Рассмотрим матрицу $A \in \mathbb{SP}(1, \mathbb{R})$, как линейное преобразование двумерной евклидовой плоскости $\mathbb{R}^2$. Тогда уравнение (19.2) запишется в виде

$$A_\alpha{}^\gamma A_\beta{}^\delta \varepsilon_{\gamma\delta} = \varepsilon_{\alpha\beta}.$$

Поскольку левая часть уравнения антисимметрична по индексам α и β , то она равна $\det A \varepsilon_{\alpha\beta}$. Отсюда следует, что уравнение (19.2) эквивалентно одному уравнению $\det A = 1$. Таким образом, группа $\mathbb{SP}(1, \mathbb{R})$ изоморфна группе вещественных 2×2 матриц с единичным определителем, т.е. группе $\mathbb{SL}(2, \mathbb{R})$.

Из курса линейной алгебры известно, что любое линейное преобразование плоскости, для которого $\det A = 1$, можно однозначно представить в виде композиции двух преобразований: ортогонального поворота плоскости (группа $\mathbb{SO}(2) \approx \mathbb{U}(1) \approx \mathbb{S}^1$) и преобразования, задающегося верхнетреугольной матрицей вида

$$\begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix}, \quad a > 0.$$

Вещественные числа a и b можно рассматривать в качестве координат на полуплоскости $a > 0$, которая гомеоморфна всей евклидовой плоскости. Таким образом мы получаем топологическое разложение группы $\mathbb{SP}(1, \mathbb{R})$ на прямое произведение окружности и двумерной плоскости. Прямое произведение $\mathbb{S}^1 \times \mathbb{R}^2$ гомотопически эквивалентно окружности (стягивается к окружности), и поэтому фундаментальная группа $\mathbb{SP}(1, \mathbb{R})$ изоморфна фундаментальной группой окружности (теорема 10.3.4). $\square$

Симплектические группы $\mathbb{SP}(n, \mathbb{R})$ при $n > 1$ имеют более сложную структуру, и их описание выходит за рамки настоящей монографии. Отметим лишь, что все группы $\mathbb{SP}(n, \mathbb{R})$ некомпактны. В дальнейшем нам понадобятся другой класс групп Ли, который обозначается $\mathbb{SP}(n)$ и состоит из компактных вещественных симплектических групп. Они существуют и строятся, как подгруппы в комплексных симплектических группах $\mathbb{SP}(n) \subset \mathbb{SP}(n, \mathbb{C})$.

Определение. Рассмотрим $2n$ -мерное комплексное пространство $\mathbb{C}^{2n}$ с координатами z^α , $\alpha = 1, \dots, 2n$. Каноническая симплектическая форма ϖ задает на $\mathbb{C}^{2n}$ симплектическую структуру (билинейную квадратичную форму), т.е. двум векторам X и Y ставится в соответствие число

$$\varpi : \mathbb{C}^{2n} \times \mathbb{C}^{2n} \ni X, Y \mapsto \varpi(X, Y) := X^\alpha Y^\beta \varpi_{\alpha\beta} \in \mathbb{C}.$$

По определению, симплектическая структура антисимметрична $\varpi(X, Y) = -\varpi(Y, X)$. Невырожденное комплексное линейное преобразование координат называется *симплектическим*, если оно сохраняет каноническую симплектическую структуру. Это

преобразование задается комплексной $2n \times 2n$ матрицей $A \in \mathbb{SP}(2n, \mathbb{C})$, которая удовлетворяет тому же равенству (19.2), что и в вещественном случае. Единственное отличие состоит в том, что элементами матрицы A теперь являются комплексные числа. Эти матрицы образуют *комплексную симплектическую группу* $\mathbb{SP}(n, \mathbb{C})$. $\square$

Ясно, что группа $\mathbb{SP}(n, \mathbb{R})$ содержится в группе $\mathbb{SP}(n, \mathbb{C})$, как подгруппа вещественных симплектических преобразований, $\mathbb{SP}(n, \mathbb{R}) \subset \mathbb{SP}(n, \mathbb{C})$.

Представление элементов алгебры Ли $\mathfrak{sp}(n, \mathbb{C})$ в виде (19.3) справедливо и для комплексных симплектических групп, только матрицы B, C и D будут комплексными. Поэтому размерность группы $\mathbb{SP}(n, \mathbb{C})$ в два раза больше размерности вещественной группы $\mathbb{SP}(n, \mathbb{R})$ и равна $2n(2n + 1)$. Так же, как и в вещественном случае, группа $\mathbb{SP}(n, \mathbb{C})$ при всех n является некомпактной.

Предложение 19.1.3. *Симплектические группы $\mathbb{SP}(n, \mathbb{R})$ и $\mathbb{SP}(n, \mathbb{C})$ являются связными, т.е. состоят из одной компоненты.*

Предложение 19.1.4. *Характеристический полином*

$$f(\lambda) := \det(\lambda \mathbb{1} - A) = \sum_{k=1}^{2n} a_k \lambda^k$$

симплектического вещественного преобразования $A \in \mathbb{SP}(n, \mathbb{R})$ обладает свойством

$$f(\lambda) = \lambda^{2n} f(1/\lambda),$$

что означает симметричность его коэффициентов: $a_k = a_{2n-k}$. В частности, если λ собственное число симплектического преобразования, то $1/\lambda$ также собственное число.

Доказательство. Из определения симплектического преобразования (19.2) следует, что $A = -\varpi A^{-1\tau} \varpi$, т.к. $\varpi^2 = -\mathbb{1}$. Отсюда вытекает цепочка равенств

$$\begin{aligned} f(\lambda) &= \det(\lambda \mathbb{1} + \varpi A^{-1\tau} \varpi) = \det(-\lambda \mathbb{1} + A^{-1\tau}) = \\ &= \det(\mathbb{1} - \lambda A) = \lambda^{2n} \det\left(\frac{1}{\lambda} \mathbb{1} - A\right), \end{aligned}$$

где мы воспользовались равенством $\det A = \det A^\tau = 1$. $\square$

Отметим, что у характеристического полинома не может быть нулевого собственного значения, т.к. $\det A = 1$. Поскольку характеристический полином является вещественным, то, если λ – комплексное собственное число, то $\bar{\lambda}$ – также собственное число. Таким образом, в случае общего положения собственные числа вещественного симплектического преобразования разбиваются на четверки $\lambda, \bar{\lambda}, 1/\lambda, 1/\bar{\lambda}$, т.е. собственные числа расположены на комплексной плоскости симметрично относительно вещественной оси и единичной окружности.

По построению, комплексная симплектическая группа $\mathbb{SP}(n, \mathbb{C})$ содержит некомпактную вещественную подгруппу $\mathbb{SP}(n, \mathbb{R})$. Оказывается, что группа $\mathbb{SP}(n, \mathbb{C})$ содержит также компактную подгруппу, которая называется компактной симплектической группой и обозначается через $\mathbb{SP}(n)$. Эту подгруппу удобно определить с использованием алгебры кватернионов $\mathbb{H}$ (см. приложение 37.3).

Определение. Рассмотрим n -мерное кватернионное пространство $\mathbb{H}^n$ с базисом e_A , $A = 1, \dots, n$. Каждый вектор $q \in \mathbb{H}^n$ однозначно представим в виде $q = q^A e_A$, где каждая координата является кватернионом: $q^A \in \mathbb{H}$. Каждый кватернион разлагается по базису $\{1, i, j, k\}$:

$$q^A = a^A + b^A i + c^A j + d^A k, \quad a^A, b^A, c^A, d^A \in \mathbb{R}.$$

Вещественная размерность $\mathbb{H}^n$ равна $4n$. Очевидно, что $\mathbb{H}^1 = \mathbb{H}$.

Рассмотрим в $\mathbb{H}^n$ симметричное вещественнозначное скалярное произведение

$$(q_1, q_2) := \operatorname{re} q_1 \bar{q}_2 = \sum_{A=1}^n (a_1^A a_2^A + b_1^A b_2^A + c_1^A c_2^A + d_1^A d_2^A). \quad (19.6)$$

Множество всех линейных кватернионных преобразований $A \in \operatorname{GL}(n, \mathbb{H})$ пространства $\mathbb{H}^n$, не меняющих начало координат и сохраняющих скалярное произведение

$$(Aq_1, Aq_2) = (q_1, q_2),$$

называется *симплектической компактной группой* $\operatorname{SP}(n)$. $\square$

Кватернионное пространство $\mathbb{H}^n$ естественным образом отождествляется с евклидовым пространством $\mathbb{R}^{4n}$. Тогда скалярное произведение (19.6) совпадает с обычным евклидовым скалярным произведением в $\mathbb{R}^{4n}$. Это значит, что группа $\operatorname{SP}(n)$ является подгруппой в $\operatorname{O}(4n, \mathbb{R})$. Поскольку группа вращений $\operatorname{O}(4n, \mathbb{R})$ компактна, то и симплектическая группа $\operatorname{SP}(n)$ также компактна.

В приложении 37.3 показано, что алгебра кватернионов $\mathbb{H}$ естественным образом отождествляется с двумерным комплексным пространством $\mathbb{C}^2$. Выполняя эту операцию вдоль каждой из кватернионных координат, мы получим отождествление $\mathbb{H}^n$ с $\mathbb{C}^{2n}$. При отождествлении

$$q = z_1 + j\bar{z}_2, \quad z_1 := a + ib, \quad z_2 := c + id,$$

квадратичная форма двух кватернионов $q_1 = z_{11} + j\bar{z}_{12}$ и $q_2 = z_{21} + j\bar{z}_{22}$ перейдет в сумму комплексных квадратичных форм

$$(q_1, q_2)_{\mathbb{H}} := q_1 \bar{q}_2 = (q_1, q_2)_{\mathbb{C}} + \langle q_1, q_2 \rangle_{\mathbb{C}},$$

где

$$(q_1, q_2)_{\mathbb{C}} := \sum_A (z_{11}^A \bar{z}_{21}^A + z_{12}^A \bar{z}_{22}^A), \quad \langle q_1, q_2 \rangle_{\mathbb{C}} := \sum_A (z_{12}^A \bar{z}_{21}^A - z_{11}^A \bar{z}_{22}^A).$$

Квадратичная форма $(q_1, q_2)_{\mathbb{C}}$ эрмитова, т.е. $\overline{(q_1, q_2)_{\mathbb{C}}} = (q_2, q_1)_{\mathbb{C}}$, а форма $\langle q_1, q_2 \rangle_{\mathbb{C}}$ антисимметрична: $\langle q_1, q_2 \rangle_{\mathbb{C}} = -\langle q_2, q_1 \rangle_{\mathbb{C}}$. Отметим, что квадратичная форма $(q_1, q_2)_{\mathbb{H}}$ отличается от квадратичной формы (19.6) отсутствием знака реальной части.

Теорема 19.1.2. *Множество элементов $\operatorname{SP}(n)$ является связной компактной группой Ли вещественной размерности $n(2n+1)$. При отождествлении $\mathbb{H}^n$ с $\mathbb{C}^{2n}$ группа $\operatorname{SP}(n)$ вкладывается как подгруппа в унитарную группу $\operatorname{U}(2n)$. При этом вложении алгебра Ли $\mathfrak{sp}(n)$ группы $\operatorname{SP}(n)$ состоит из комплексных $2n \times 2n$ матриц вида*

$$\begin{pmatrix} Z_1 & Z_2 \\ -\bar{Z}_2 & \bar{Z}_1 \end{pmatrix} \quad (19.7)$$

где Z_1 – комплексная антиэрмитова $n \times n$ матрица, а Z_2 – комплексная симметричная $n \times n$ матрица. Если матрицы из унитарной группы $\mathbb{U}(2n)$ представить в блочном виде

$$\begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix}$$

где U_1, U_2, U_3 и U_4 – комплексные $n \times n$ матрицы, то подгруппа $\mathbb{SP}(n)$ в $\mathbb{U}(2n)$ состоит из унитарных матриц вида

$$\begin{pmatrix} U_1 & U_2 \\ -\bar{U}_2 & \bar{U}_1 \end{pmatrix} \quad (19.8)$$

При этом матрица (19.8) является унитарной тогда и только тогда, когда комплексные матрицы U_1 и U_2 удовлетворяют уравнениям $U_1 U_1^\dagger + U_2 U_2^\dagger = \mathbb{1}$ и $U_2 U_1^\top = U_1 U_2^\top$.

Доказательство. Проводится прямой проверкой [131], раздел 2.3, лемма 6. $\square$

Связь симплектических групп с другими матричными группами дается следующими двумя теоремами, доказательство которых дано, например, в [131] раздел 2.4, предложение 2.

Теорема 19.1.3. Рассмотрим стандартные вложения групп $\mathbb{O}(n) \hookrightarrow \mathbb{U}(n)$, $\mathbb{U}(n) \hookrightarrow \mathbb{SO}(2n)$, $\mathbb{SP}(n) \hookrightarrow \mathbb{U}(2n)$. Тогда имеют место следующие соотношения:

- 1) $\mathbb{SO}(2n) \cap \mathbb{SP}(n) = \mathbb{U}(n)$; здесь группы $\mathbb{SO}(2n)$ и $\mathbb{SP}(n)$ рассматриваются, как подгруппы в одной группе $\mathbb{U}(2n)$;
- 2) $\mathbb{SP}(n, \mathbb{C}) \cap \mathbb{U}(2n) = \mathbb{SP}(n)$;
- 3) $\mathbb{SP}(n, \mathbb{R}) \cap \mathbb{GL}(2n, \mathbb{C}) = \mathbb{U}(n)$;
- 4) $\mathbb{SP}(n, \mathbb{R}) \cap \mathbb{U}(2n) = \mathbb{U}(n)$.

Теорема 19.1.4. Для компактных симплектических групп $\mathbb{SP}(1)$ и $\mathbb{SP}(2)$ имеют место изоморфизмы:

$$\mathbb{SP}(1) \simeq \mathbb{SU}(2) \simeq \mathbb{SPIN}(3), \quad \mathbb{SP}(2) \simeq \mathbb{SPIN}(5),$$

где спинорные группы $\mathbb{SPIN}$ определены в разделе 26.7.

При больших $n > 2$ компактные симплектические группы $\mathbb{SP}(n)$ уже не сводятся к унитарным и ортогональным группам.

Теорема 19.1.5. Группа $\mathbb{SP}(n)$ является максимальной компактной подгруппой в комплексной симплектической группе $\mathbb{SP}(n, \mathbb{C})$.

Доказательство. См., например, [66], глава X, §6, пункт 1, таблица IV. $\square$

Из этой теоремы следует, что вещественная группа $\mathbb{SP}(n)$ сохраняет каноническую симплектическую форму (19.1), однако реализуется с помощью комплексных матриц.

Теорема 19.1.6. Все компактные симплектические группы $\mathbb{SP}(n)$, $n \geq 1$, связны и односвязны.

Доказательство. См., например, [67], лекция 1, пример 16; лекция 12, текст после теоремы 8. $\square$

19.2 Симплектические многообразия

Определение. Многообразие $\mathbb{M}$ четной размерности, $\dim \mathbb{M} = 2n$, называется *симплектическим*, если на нем задана достаточно гладкая 2-форма

$$\omega = \frac{1}{2} dx^\alpha \wedge dx^\beta \omega_{\alpha\beta} \in \Lambda_2(\mathbb{M}),$$

удовлетворяющая двум условиям:

- 1) $\det \omega_{\alpha\beta} \neq 0, \quad \forall x \in \mathbb{M}$ — невырожденность,
- 2) $d\omega = 0$ — замкнутость.

Форма ω называется *симплектической*. □

Пусть на многообразии $\mathbb{M}$ задано два векторных поля $X, Y \in \mathcal{X}(\mathbb{M})$. Тогда форма ω называется невырожденной, если из условия $\omega(X, Y) = 0$ для всех $Y \in \mathcal{X}(\mathbb{M})$ следует $X = 0$. Легко проверить, что данное инвариантное определение невырожденности 2-формы эквивалентно условию $\det \omega_{\alpha\beta} \neq 0$, которое должно выполняться во всех картах и для всех точек $x \in \mathbb{M}$.

По определению, компоненты симплектической формы антисимметричны относительно перестановки индексов. Это значит, что ее невырожденность возможна только на многообразиях четной размерности. Поэтому размерность многообразия включена в определение.

В координатах замкнутость симплектической формы ω записывается в виде дифференциального уравнения

$$\partial_\alpha \omega_{\beta\gamma} + \partial_\beta \omega_{\gamma\alpha} + \partial_\gamma \omega_{\alpha\beta} = 0. \quad (19.9)$$

Рассмотрим связь симплектических форм с ориентацией многообразий и формой объема.

Теорема 19.2.1. Пусть ω — замкнутая 2-форма на многообразии $\mathbb{M}$, $\dim \mathbb{M} = 2n$. Для того, чтобы эта форма была симплектической необходимо и достаточно, чтобы $2n$ -форма ω^n была формой объема на $\mathbb{M}$, т.е. нигде не обращалась в нуль.

Доказательство. Следствие предложения 3.4.3. □

Следствие. Любое симплектическое многообразие $(\mathbb{M}, \omega)$ ориентируемо. □

Доказательство. Отличие от нуля формы ω^n согласно теореме 10.4.3 достаточно для ориентируемости. □

Обычно симплектическое многообразие $(\mathbb{M}, \omega)$ ориентируют формой объема v с дополнительным множителем:

$$v := \frac{1}{n!} (-1)^{\frac{n(n+1)}{2}} \omega^n = \frac{1}{n!} (-1)^{\frac{n(n+1)}{2}} \underbrace{\omega \wedge \dots \wedge \omega}_n, \quad (19.10)$$

чтобы согласовать выбор канонических координат (см. ниже) с канонической ориентацией евклидова пространства.

Пример 19.2.1. Рассмотрим каноническую 2-форму (19.1) на евклидовом пространстве $\mathbb{R}^{2n}$

$$\varpi = \frac{1}{2} dx^\alpha \wedge dx^\beta \varpi_{\alpha\beta} = dx^{n+1} \wedge dx^1 + dx^{n+2} \wedge dx^2 + \dots + dx^{2n} \wedge dx^n. \quad (19.11)$$

Она невырождена, т.к. $\det \varpi_{\alpha\beta} = 1$, и замкнута, поскольку компоненты формы постоянны. Тем самым форма ϖ определяет симплектическую форму на евклидовом пространстве. Поскольку

$$\varpi^n = (-1)^{\frac{n(n+1)}{2}} n! dx^1 \wedge \dots \wedge dx^{2n},$$

то соответствующая форма объема

$$v := \frac{1}{n!} (-1)^{\frac{n(n+1)}{2}} \varpi^n = dx^1 \wedge \dots \wedge dx^{2n} \quad (19.12)$$

является канонической формой объема евклидова пространства (3.80). $\square$

Обратное утверждение о связи между наличием формы объема и симплектической формы в общем случае неверно. Исключение составляют двумерные многообразия: всякая ориентируемая поверхность допускает симплектическую структуру. Это просто полностью антисимметричный тензор второго ранга:

$$\omega = v = \frac{1}{2} dx^\alpha \wedge dx^\beta \varepsilon_{\alpha\beta} = dx^1 \wedge dx^2 \sqrt{|g|}.$$

Для ориентируемых многообразий более высокой размерности $\dim \mathbb{M} = 2n$, $n > 1$, симплектическая форма существует не всегда.

Следующая конструкция позволяет строить симплектические структуры на произвольном кокасательном расслоении $\mathbb{T}^*(\mathbb{M})$, $\dim \mathbb{M} = n$, которое имеет размерность $2n$.

Определение. Пусть $X_r \in \mathbb{T}_r(\mathbb{T}^*(\mathbb{M}))$ – произвольный касательный вектор к кокасательному расслоению в точке $r \in \mathbb{T}^*(\mathbb{M})$. Определим *каноническую линейную форму* Θ на $\mathbb{T}^*(\mathbb{M})$, которая называется *формой Лиувилля* или *относительным интегральным инвариантом Пуанкаре*, следующим соотношением

$$\Theta(X_r) = r(\pi_* X_r), \quad (19.13)$$

где π_* – дифференциал проекции кокасательного расслоения, $\pi : \mathbb{T}^*(\mathbb{M}) \rightarrow \mathbb{M}$. $\square$

Теорема 19.2.2. На кокасательном расслоении $\mathbb{T}^*(\mathbb{M})$ к произвольному многообразию $\mathbb{M}$ существует симплектическая структура.

Доказательство. Заметим, что r является линейной формой на $\mathbb{T}_{\pi(r)}(\mathbb{M})$ и $\pi_* X_r$ – касательный вектор в точке $\pi(r)$. Следовательно, форма Лиувилля (19.13) определена всюду на кокасательном расслоении $\mathbb{T}^*(\mathbb{M})$. Пусть q^α , $\alpha = 1, \dots, n$ – локальная система координат на $\mathbb{U} \subset \mathbb{M}$. Введем на $\pi^{-1}(\mathbb{U})$ координаты $(x^1, \dots, x^{2n}) = (q^1, \dots, q^n, p_1, \dots, p_n)$ следующим образом. Формы $dx^1, \dots, dx^n$ порождают пространство 1-форм на $\mathbb{U}$. Поэтому для всякой 1-формы $r \in \pi^{-1}(\mathbb{U})$ имеем равенство

$$r = dq^\alpha p_\alpha.$$

Разложим касательный вектор X_r по координатному базису (q, p) :

$$X_r = X_r^\alpha \frac{\partial}{\partial q^\alpha} + X_{r\alpha} \frac{\partial}{\partial p_\alpha}.$$

Тогда

$$\pi_* X_r = X_r^\alpha \frac{\partial}{\partial x^\alpha},$$

и условие (19.13) принимает вид

$$\Theta(X_r) = X_r^\alpha p_\alpha.$$

Таким образом получаем выражение для формы Лиувилля в локальных координатах:

$$\Theta = dq^\alpha p_\alpha. \quad (19.14)$$

Внешняя производная от формы Лиувилля имеет вид

$$d\Theta = dp_\alpha \wedge dq^\alpha = \frac{1}{2} dx^A \wedge dx^B \varpi_{AB}, \quad A, B = 1, \dots, 2n,$$

т.е. совпадает с канонической симплектической формой ϖ . Таким образом построена невырожденная замкнутая 2-форма $d\Theta$ на произвольном кокасательном расслоении. $\square$

Форма Лиувилля естественна в том смысле, что для любого сечения $\sigma : \mathbb{M} \rightarrow T^*(\mathbb{M})$ возврат отображения σ^* переводит 1-форму Θ в r : $\sigma^*(\Theta) = r$.

Следствие. Любое кокасательное расслоение $T^*(\mathbb{M})$ ориентируемо. $\square$

Замечание. В гамильтоновом подходе к описанию динамики точечных частиц (см. главу 21) многообразие $\mathbb{M}$ и его кокасательное расслоение $T^*(\mathbb{M})$ отождествляются, соответственно, с конфигурационным и фазовым пространствами. $\square$

Поскольку ранг симплектической 2-формы максимален и равен $2n$, то ее класс также равен $2n$. В этом случае теорема Дарбу (3.4.4) формулируется следующим образом.

Теорема 19.2.3 (Дарбу). Пусть задано симплектическое многообразие $(\mathbb{M}, \omega)$. Тогда у каждой точки $x \in \mathbb{M}$ существует такая координатная окрестность $\mathbb{U}_x \subset \mathbb{M}$, в которой симплектическая форма принимает канонический вид

$$\omega = \varpi.$$

Определение. Координаты, в которых симплектическая форма имеет канонический вид, называются *координатами Дарбу*. $\square$

Замечание. В главе 4 мы видели, что метрику g на многообразии $\mathbb{M}$ за счет выбора системы координат можно привести к каноническому виду только в фиксированной точке $x \in \mathbb{M}$. Лучшее, что можно сделать в общем случае, это привести метрику к каноническому виду вдоль произвольной кривой $\gamma \in \mathbb{M}$. В этом отношении симплектические многообразия проще. Согласно теореме Дарбу симплектическую форму можно привести к каноническому виду не только в данной точке многообразия, но и в некоторой окрестности этой точки. $\square$

Поскольку симплектическая структура $\omega_{\alpha\beta}(x)$ является невырожденной, то в каждой точке $x \in \mathbb{M}$ существует обратная матрица $\omega^{-1\alpha\beta}(x)$, которая также антисимметрична:

$$\omega^{-1\alpha\beta}\omega_{\beta\gamma} = \omega_{\gamma\beta}\omega^{-1\beta\alpha} = \delta_{\gamma}^{\alpha}, \quad \omega^{-1\alpha\beta} = -\omega^{-1\beta\alpha}.$$

Замечание. Если на симплектическом многообразии $(\mathbb{M}, \omega)$ задана также метрика g , то в общем случае

$$\omega^{-1\alpha\beta} \neq \omega^{\alpha\beta} := g^{\alpha\gamma}g^{\beta\delta}\omega_{\gamma\delta}. \quad \square$$

Пример 19.2.2. Для канонической симплектической структуры (19.1)

$$\varpi^{-1} = (\varpi^{-1\alpha\beta}) = -\varpi = \varpi^T = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}. \quad \square$$

Симплектическая структура устанавливает взаимно однозначное соответствие сечений касательных и кокасательных расслоений. В компонентах данное соответствие задается простым правилом:

$$A_{\alpha} = X^{\beta}\omega_{\beta\alpha}, \quad X^{\alpha} = A_{\beta}\omega^{-1\beta\alpha}, \quad X \in \mathcal{X}(\mathbb{M}), \quad A \in \Lambda_1(\mathbb{M}). \quad (19.15)$$

Замкнутость симплектической формы (19.9) для обратного тензора $\omega^{-1\alpha\beta}(x)$ можно переписать в виде

$$\omega^{-1\alpha\delta}\partial_{\delta}\omega^{-1\beta\gamma} + \omega^{-1\beta\delta}\partial_{\delta}\omega^{-1\gamma\alpha} + \omega^{-1\gamma\delta}\partial_{\delta}\omega^{-1\alpha\beta} = 0.$$

Для этого достаточно свернуть равенство (19.9) с обратными матрицами $\omega^{-1\alpha\beta}$.

Используя внутреннее умножение (3.24), формулы (19.15) можно переписать в инвариантном виде:

$$A = i_{X_A}\omega = X_A \rfloor \omega,$$

где X_A – векторное поле, соответствующее 1-форме A .

Поскольку симплектическая структура на $\mathbb{M}$ устанавливает взаимно однозначное соответствие между множеством векторных полей $\mathcal{X}(\mathbb{M})$ и пространством 1-форм $\Lambda_1(\mathbb{M})$, то структуру алгебры Ли на $\mathcal{X}(\mathbb{M})$ можно перенести на $\Lambda_1(\mathbb{M})$. Для двух произвольных 1-форм $A, B \in \Lambda_1(\mathbb{M})$ положим

$$[A, B] := [X_A, X_B] \rfloor \omega. \quad (19.16)$$

Тем самым линейное пространство $\Lambda_1(\mathbb{M})$ приобретает структуру алгебры Ли. Таким образом коммутатору векторных полей ставится в соответствие *скобка Пуассона* 1-форм (19.16).

В компонентах $A = dx^{\alpha}A_{\alpha}$, $B = dx^{\alpha}B_{\alpha}$ и

$$[A, B] = dx^{\alpha} [(A_{\beta}\partial_{\gamma}B_{\alpha} + B_{\gamma}\partial_{\beta}A_{\alpha})\omega^{-1\beta\gamma} + A_{\beta}B_{\gamma}\partial_{\alpha}\omega^{-1\beta\gamma}]. \quad (19.17)$$

Эта формула получается прямыми вычислениями.

Скобку Пуассона двух 1-форм можно выразить через производную Ли:

$$\begin{aligned} \mathcal{L}_{X_A}B &= \mathcal{L}_{X_A}(X_B \rfloor \omega) = \mathcal{L}_{X_A}X_B \rfloor \omega + X_B \rfloor (\mathcal{L}_{X_A}\omega) = \\ &= [X_A, X_B] \rfloor \omega + X_B \rfloor (X_A \rfloor d\omega + d(X_A \rfloor \omega)) = [A, B] + X_B \rfloor dA, \end{aligned}$$

где мы воспользовались линейностью производной Ли, основной формулой гомотопии (3.43) и замкнутостью симплектической формы. Отсюда следует выражение для скобки Пуассона через производную Ли:

$$[A, B] = \mathcal{L}_{X_A}B - X_B \rfloor dA. \quad (19.18)$$

Если 1-формы A и B замкнуты, то полученное выражение упрощается:

$$[A, B] = L_{X_A} B = -L_{X_B} A.$$

Предложение 19.2.1. *Скобка Пуассона двух замкнутых форм является точной формой.*

Доказательство. Из основной формулы гомотопии (3.43) следует, что для любой замкнутой формы A справедливо равенство $L_X A = d(X \lrcorner A)$. $\square$

В компонентах скобка Пуассона двух замкнутых 1-форм (19.17) принимает простой вид

$$[A, B] = dx^\alpha \partial_\alpha (\omega^{-1\beta\gamma} A_\beta B_\gamma).$$

Скобку Пуассона можно также определить в алгебре функций на $\mathbb{M}$. А именно, каждая симплектическая структура определяет скобку Пуассона двух дифференцируемых функций $f(x)$ и $g(x)$ класса $\mathcal{C}^2$:

$$[f, g] := \omega^{-1\alpha\beta} \frac{\partial f}{\partial x^\alpha} \frac{\partial g}{\partial x^\beta}. \quad (19.19)$$

Дважды непрерывная дифференцируемость функций необходима для выполнения тождеств Якоби.

Поля $\omega^{-1\alpha\beta}(x)$ являются компонентами антисимметричного контравариантного тензора второго ранга. Они являются частным случаем пуассоновой структуры, рассмотренной в следующем разделе.

Пусть на симплектическом многообразии $(\mathbb{M}, \omega)$ задана также аффинная геометрия, т.е. метрика g и аффинная связность Γ . Назовем аффинную связность *согласованной с симплектической структурой*, если ковариантная производная симплектической формы равна нулю:

$$\nabla_\alpha \omega_{\beta\gamma} = \partial_\alpha \omega_{\beta\gamma} - \Gamma_{\alpha\beta}^\delta \omega_{\delta\gamma} - \Gamma_{\alpha\gamma}^\delta \omega_{\beta\delta} = 0. \quad (19.20)$$

Рассмотрим линейную комбинацию ковариантных производных, в которой слагаемые отличаются циклической перестановкой индексов:

$$\nabla_\alpha \omega_{\beta\gamma} + \nabla_\beta \omega_{\gamma\alpha} + \nabla_\gamma \omega_{\alpha\beta}.$$

Слагаемые с производными от симплектической формы сокращаются ввиду замкнутости ω , и мы получаем уравнение на тензор кручения

$$T_{\alpha\beta}^\delta \omega_{\delta\gamma} + T_{\beta\gamma}^\delta \omega_{\delta\alpha} + T_{\gamma\alpha}^\delta \omega_{\delta\beta} = 0. \quad (19.21)$$

Это необходимое условие для того, чтобы аффинная связность была согласована с симплектической структурой, но не достаточное.

19.3 Пуассоновы многообразия

Рассмотрим множество скалярных полей $f, g, \dots \in \mathcal{C}^\infty(\mathbb{M})$ (алгебру гладких функций) на многообразии $\mathbb{M}$ произвольной размерности n .

Определение. Билинейное отображение

$$J : C^\infty(\mathbb{M}) \times C^\infty(\mathbb{M}) \ni f, g \mapsto [f, g] \in C^\infty(\mathbb{M})$$

называется *пуассоновой структурой* или *скобкой Пуассона*, если выполнены следующие четыре условия:

- 1) $[af + bg, h] = a[f, h] + b[g, h], \quad a, b \in \mathbb{R}$ — линейность,
- 2) $[f, g] = -[g, f]$ — антисимметрия,
- 3) $[f, gh] = [f, g]h + g[f, h]$ — правило Лейбница,
- 4) $[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0$ — тождество Якоби.

Многообразие с заданной скобкой Пуассона называется *пуассоновым*. $\square$

Из линейности скобки Пуассона и правила Лейбница следует, что скобка Пуассона постоянной функции $f = C = \text{const}$ с произвольной функцией равна нулю:

$$[C, g] = 0.$$

Для доказательства достаточно положить $g = C$ в условии 3. Используя это свойство, получим явное выражение для скобки Пуассона в локальной системе координат x^α . С этой целью разложим функции f и g в ряды Тейлора в окрестности произвольной точки x_0 :

$$\begin{aligned} f(x) &= f_0 + (x^\alpha - x_0^\alpha) \left. \frac{\partial f}{\partial x^\alpha} \right|_{x=x_0} + \dots, \\ g(x) &= g_0 + (x^\alpha - x_0^\alpha) \left. \frac{\partial g}{\partial x^\alpha} \right|_{x=x_0} + \dots \end{aligned}$$

Тогда

$$\begin{aligned} [f, g] &= [x^\alpha - x_0^\alpha, x^\beta - x_0^\beta] \left. \frac{\partial f}{\partial x^\alpha} \frac{\partial g}{\partial x^\beta} \right|_{x=x_0} + \mathbf{o}(x - x_0) = \\ &= [x^\alpha, x^\beta] \left. \frac{\partial f}{\partial x^\alpha} \frac{\partial g}{\partial x^\beta} \right|_{x=x_0} + \mathbf{o}(x - x_0). \end{aligned}$$

Таким образом, в точке x_0 скобка определяется первым слагаемым. Поскольку точка x_0 произвольна, то получаем явное выражение для скобки Пуассона в координатах

$$[f, g] = J^{\alpha\beta} \frac{\partial f}{\partial x^\alpha} \frac{\partial g}{\partial x^\beta}, \quad (19.22)$$

где введены *структурные функции* пуассонова многообразия:

$$J^{\alpha\beta} := [x^\alpha, x^\beta] \quad (19.23)$$

Напомним, что каждую координату можно рассматривать, как функцию на многообразии $x^\alpha : \mathbb{M} \ni x \mapsto x^\alpha(x) \in \mathbb{R}$, поэтому для координатных функций скобка Пуассона (19.23) определена.

Скобка Пуассона (19.22) совпадает со скобкой Пуассона (19.19), введенной для симплектической структуры, при $J^{\alpha\beta} = \omega^{-1\alpha\beta}$.

По построению, структурные функции $J^{\alpha\beta}(x)$ представляют собой компоненты антисимметричного тензора второго ранга, и выражение в правой части равенства (19.22) является функцией. Из определения скобки Пуассона следует, что структурные функции антисимметричны,

$$J^{\alpha\beta} = -J^{\beta\alpha},$$

и удовлетворяют тождеству Якоби:

$$J^{\alpha\delta}\partial_\delta J^{\beta\gamma} + J^{\beta\delta}\partial_\delta J^{\gamma\alpha} + J^{\gamma\delta}\partial_\delta J^{\alpha\beta} = 0, \quad (19.24)$$

где слагаемые отличаются циклической перестановкой индексов α, β, γ . Легко проверяется, что произвольный антисимметричный тензор второго ранга, удовлетворяющий тождеству Якоби (19.24), взаимно однозначно определяет скобку Пуассона (19.22).

Пример 19.3.1. Произвольная постоянная антисимметричная матрица определяет пуассонову структуру в некоторой системе координат. Действительно, тождества Якоби в этом случае выполняются. $\square$

Пуассонова структура является билинейным дифференциальным оператором и записывается также в виде

$$J(f, g) := [f, g] = J^{\alpha\beta}\partial_\alpha f \partial_\beta g.$$

Поэтому для нее применяют иногда обозначение

$$J = \frac{1}{2} J^{\alpha\beta} \partial_\alpha \wedge \partial_\beta,$$

где знак $\wedge$ обозначает внешнее умножение (см. раздел 3).

Если на многообразии помимо пуассоновой структуры J задана также аффинная связность Γ , то тождество Якоби (19.24) можно переписать в явно ковариантном виде

$$\begin{aligned} J^{\alpha\delta}\nabla_\delta J^{\beta\gamma} + J^{\beta\delta}\nabla_\delta J^{\gamma\alpha} + J^{\gamma\delta}\nabla_\delta J^{\alpha\beta} - \\ - J^{\alpha\delta}T_{\delta\epsilon}^{\beta\gamma} J^{\epsilon\gamma} - J^{\beta\delta}T_{\delta\epsilon}^{\gamma\alpha} J^{\epsilon\alpha} - J^{\gamma\delta}T_{\delta\epsilon}^{\alpha\beta} J^{\epsilon\beta} = 0, \end{aligned}$$

где $T_{\alpha\beta}^{\gamma}$ – тензор кручения аффинной связности. Это доказывает, что тождества Якоби ковариантны, и их выполнение в одной системе координат влечет за собой их выполнение во всех других системах. Впрочем, это можно было бы и не доказывать, т.к. определение пуассоновой структуры было дано в инвариантном виде.

Если матрица структурных функций невырождена, $\det J^{\alpha\beta} \neq 0$, то она имеет обратную $J_{\alpha\beta}^{-1}$:

$$J^{\alpha\beta} J_{\beta\gamma}^{-1} = \delta_\gamma^\alpha,$$

которая также антисимметрична. Это значит, что контравариантный тензор $J^{\alpha\beta}$ взаимно однозначно определяет 2-форму $\frac{1}{2} dx^\alpha \wedge dx^\beta J_{\alpha\beta}^{-1}$ на многообразии. Из тождеств Якоби (19.24) следует, что 2-форма $J_{\alpha\beta}^{-1}$ замкнута,

$$\partial_\alpha J_{\beta\gamma}^{-1} + \partial_\beta J_{\gamma\alpha}^{-1} + \partial_\gamma J_{\alpha\beta}^{-1} = 0,$$

т.е. является симплектической (см. раздел 19.2). Легко проверяется и обратное утверждение: произвольная симплектическая форма $\omega_{\alpha\beta} = J_{\alpha\beta}^{-1}$ определяет пуассонову

структуру на многообразии с невырожденными структурными функциями. При этом тождества Якоби следуют из условия замкнутости формы. Таким образом, пуассоновы структуры с невырожденными структурными функциями находятся во взаимно однозначном соответствии с симплектическими формами на многообразии.

В общем случае структурные функции пуассоновой структуры могут быть вырождены. Это значит, в частности, что, в отличие от симплектической структуры, скобка Пуассона может быть также определена на многообразии нечетной размерности. Рангом пуассоновой структуры называется ранг матрицы $J^{\alpha\beta}$ структурных функций. Ввиду антисимметрии $J^{\alpha\beta}$ ранг пуассоновой структуры может быть только четным. В общем случае ранг пуассоновой структуры может меняться от точки к точке.

Определение. Если пуассонова структура вырождена, $\det J^{\alpha\beta} = 0$, то существуют функции Казимира $c \in C^\infty(\mathbb{U})$, определенные, возможно, только в некоторой области $\mathbb{U} \subset \mathbb{M}$. Они определяются следующим равенством

$$[c, f] = 0, \quad \forall f \in C^\infty(\mathbb{U}). \quad \square \quad (19.25)$$

В координатах это равенство сводится к системе дифференциальных уравнений

$$J^{\alpha\beta} \partial_\beta c = 0. \quad (19.26)$$

Пусть ранг матрицы $J^{\alpha\beta}$ постоянен на $\mathbb{M}$ и равен $\text{rank } J^{\alpha\beta} = 2m < n$. Это значит, что уравнение $J^{\alpha\beta} h_\beta = 0$ относительно h_β имеет $n - 2m$ функционально независимых решений:

$$J^{\alpha\beta} h_\beta^A = 0, \quad A = 1, \dots, n - 2m. \quad (19.27)$$

Продифференцируем данное уравнение по x^γ ,

$$J^{\alpha\beta} \partial_\gamma h_\beta^A + \partial_\gamma J^{\alpha\beta} h_\beta^A = 0,$$

умножим на $J^{\delta\gamma}$ и антисимметризуем по индексам α и δ . В результате получим уравнение

$$\partial_\gamma h_\beta^A - \partial_\beta h_\gamma^A = 0,$$

где мы использовали тождество Якоби. Это значит, что система уравнений

$$\partial_\alpha c^A = h_\alpha^A,$$

локально разрешима для всех значений индекса A . Таким образом, количество функционально независимых функций Казимира равно $n - 2m$, т.е. разности между размерностью многообразия n и рангом пуассоновой структуры $2m$. Если все функции Казимира известны и определены на всем $\mathbb{M}$, то условия $c^A = \text{const}$ выделяют в $\mathbb{M}$ подмногообразие $\mathbb{U}$ размерности $2m$. В разделе 19.5 мы увидим, что сужение J на это подмногообразие приводит к невырожденной пуассоновой структуре на $\mathbb{U}$.

Очевидно, что любая функция Казимира определена с точностью до аддитивной постоянной, а также произвольная линейная комбинация функций Казимира тоже является функцией Казимира.

Из определения следует, в частности, что скобка Пуассона двух функций Казимира равна нулю:

$$[c^A, c^B] = J^{\alpha\beta} \partial_\alpha c^A \partial_\beta c^B = 0. \quad (19.28)$$

Более того, из правила Лейбница вытекает равенство

$$[c^A c^B, f] = c^A [c^B, f] + c^B [c^A, f] = 0.$$

То есть произведение функций Казимира снова дает функцию Казимира. Поскольку существует только $n - 2m$ функционально независимых функций Казимира, то любая функция Казимира имеет вид $d(c(x))$, где d – достаточно гладкая функция переменных c^A . Действительно,

$$[d, f] = J^{\alpha\beta} \partial_A d \partial_\alpha c^A \partial_\beta f = 0, \quad \forall f \in \mathcal{C}^\infty(\mathbb{M}).$$

Пусть задано два набора функционально независимых функций Казимира $d = \{d^A\}$ и $c = \{c^A\}$. Тогда $d^A = d^A(c)$ и $\partial_\alpha d^A = \partial_B d^A \partial_\alpha c^B$. Так как $\text{rank } \partial_\alpha d^A = \text{rank } \partial_\alpha c^A$, то преобразование между наборами задается $n - 2m$ функциями с невырожденным Якобианом.

Пусть задано пуассоново многообразие $(\mathbb{M}, J)$. Наличие пуассоновой структуры задает билинейное отображение на алгебре функций $\mathcal{C}^\infty(\mathbb{M})$, которое превращает множество функций в алгебру Ли. Эта алгебра Ли гомоморфно отображается в алгебру Ли векторных полей на $\mathbb{M}$ следующим образом. Скобка Пуассона координатных функций x^α с произвольной функцией f определяет компоненты векторного поля $X_f = X_f^\alpha \partial_\alpha$, где

$$X_f^\alpha := [f, x^\alpha] = -J^{\alpha\beta} \partial_\beta f.$$

Таким образом, мы имеем отображение множества гладких функций в алгебру Ли векторных полей

$$\mathcal{C}^\infty(\mathbb{M}) \ni f \mapsto X_f \in \mathcal{X}(\mathbb{M}). \quad (19.29)$$

Предложение 19.3.1. Скобка Пуассона двух функций $f, g \in \mathcal{C}^\infty(\mathbb{M})$ связана с коммутатором соответствующих векторных полей соотношением

$$[X_f, X_g] = X_{[f, g]}. \quad (19.30)$$

То есть отображение (19.29) является гомоморфизмом.

Доказательство. Простое следствие тождеств Якоби. $\square$

Если пуассонова структура невырождена, то гомоморфизм (19.29) имеет нетривиальное ядро, состоящее из функций, постоянных на $\mathbb{M}$. Если пуассонова структура вырождена, то ядро отображения (19.30) включает также линейную оболочку функций Казимира.

Предложение 19.3.2. Векторное поле X_f , порожденное произвольной функцией f , сохраняет скобку Пуассона. А именно, производная Ли вдоль векторного поля X_f от структурных функций равна нулю:

$$\mathcal{L}_{X_f} J = 0.$$

Доказательство. Простая проверка. $\square$

Обсудим связь пуассоновых структур с гамильтоновой динамикой точечных частиц. Рассмотрим произвольную функцию $H \in \mathcal{C}^\infty(\mathbb{M})$ и соответствующее векторное поле X_H . Тогда экспоненциальное отображение, определяемое системой уравнений

$$\dot{x}^\alpha = -X_H^\alpha = [x^\alpha, H] = J^{\alpha\beta} \partial_\beta H,$$

является ни чем иным, как уравнениями движения механической системы точечных частиц в канонической (гамильтоновой) форме. При этом точка обозначает дифференцирование по времени, и $H = H(x)$ – гамильтониан системы. Векторное поле

$$X_H = -J^{\alpha\beta} \partial_\beta H \partial_\alpha \quad (19.31)$$

называется *гамильтоновым*. Другими словами, каждую функцию на пуассоновом многообразии можно рассматривать, как гамильтониан некоторой механической системы. При этом пуассоново многообразие отождествляется с фазовым пространством, и каждому гамильтониану соответствует свое гамильтоново векторное поле.

Легко проверить, что для каждого гамильтонова векторного поля справедливо равенство

$$X_H f = [f, H]. \quad (19.32)$$

Пример 19.3.2. Механическая система, состоящая из N точечных частиц, в фазовом пространстве описывается координатами q^i и импульсами p_i , $i = 1, \dots, N$, со скобками Пуассона

$$[q^i, p_j] = \delta_j^i, \quad [q^i, q^j] = 0, \quad [p_i, p_j] = 0.$$

Фазовое пространство представляет собой $2N$ -мерное пуассоново многообразие. В координатах $\{x^\alpha\} = \{q^1, \dots, q^N, p_1, \dots, p_N\}$ структурные функции имеют вид

$$J^{\alpha\beta} = \varpi^{-1\alpha\beta} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}. \quad (19.33)$$

Эта пуассонова структура невырождена и называется *канонической*. Как матрица каноническая пуассонова структура с точностью до знака совпадает с канонической симплектической формой (19.1), однако ее индексы расположены сверху, что соответствует контравариантному тензору второго ранга.

Координаты q^i и сопряженные импульсы p_i являются системой локальных координат фазового пространства $\mathbb{T}^*(\mathbb{M})$, которое является кокасательным расслоением к конфигурационному пространству $\mathbb{M}$ с координатами q^i .

Канонические уравнения движения механической системы

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}$$

записываются в виде

$$\dot{x}^\alpha = [x^\alpha, H],$$

где $H = H(q, p)$ – гамильтониан системы. Гамильтоново векторное поле на фазовом пространстве $\mathbb{T}^*(\mathbb{M})$ имеет вид

$$X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i}. \quad \square$$

Замечание. С точки зрения динамики функции Казимира представляют собой первые интегралы движения механической системы, которые существуют для широкого класса функций Гамильтона. В этом смысле они носят кинематический характер. $\square$

19.4 Структура Ли–Пуассона

Теперь рассмотрим важный класс пуассоновых структур, связанных с алгебрами Ли.

Определение. Пусть $\mathfrak{g}$ – алгебра Ли размерности N . Это – линейное пространство, каждая точка которого $y = y^A L_A \in \mathfrak{g}$, $A = 1, \dots, N$, разлагается по базису L_A с коммутационными соотношениями

$$[L_A, L_B] = f_{AB}^C L_C,$$

где f_{AB}^C – структурные константы алгебры Ли $\mathfrak{g}$. Обозначим дуальное пространство к алгебре Ли через $\mathfrak{g}^*$ (множество линейных функционалов на $\mathfrak{g}$). Пусть ω^A – базис в $\mathfrak{g}^*$, который дуален к L_A , т.е. $\omega^A(L_B) = \delta_B^A$. Тогда точка дуального пространства представима в виде $x = \omega^A x_A \in \mathfrak{g}^*$. Значение функционала x на векторе y равно

$$x(y) = y^A x_A.$$

Рассмотрим функцию $f(x) \in C^\infty(\mathfrak{g}^*)$ на дуальном пространстве $\mathfrak{g}^*$. Ее дифференциал равен

$$df = dx_A \partial^A f.$$

Мы пишем индекс у частной производной ∂^A вверху, т.к. координаты дуального пространства $\mathfrak{g}^*$ нумеруются нижним индексом. Каждый дифференциал естественным образом отождествляется с элементом алгебры Ли $\nabla f := \partial^A f L_A \in \mathfrak{g}$. Определим скобку Пуассона для функций на дуальном пространстве $f, g \in C^\infty(\mathfrak{g}^*)$ в произвольной точке $x \in \mathfrak{g}^*$:

$$[f, g](x) := x([\nabla f, \nabla g]), \quad (19.34)$$

где $[\nabla f, \nabla g]$ – коммутатор в алгебре Ли $\mathfrak{g}$. В компонентах данная скобка Пуассона записывается в виде

$$[f, g] = J_{AB} \partial^A f \partial^B g, \quad (19.35)$$

где структурные функции

$$J_{AB} := f_{AB}^C x_C \quad (19.36)$$

линейны по координатам x_C . Нетрудно проверить, что тождества Якоби для структурных функций (19.36) совпадают с тождествами Якоби для структурных констант алгебры Ли $\mathfrak{g}$:

$$f_{AB}^D f_{CD}^E + f_{BC}^D f_{AD}^E + f_{CA}^D f_{BD}^E = 0.$$

Скобка Пуассона (19.34) называется скобкой Ли–Пуассона. □

Таким образом, скобка Ли–Пуассона является частным случаем скобки Пуассона с линейными структурными функциями (19.36). Ранг этой скобки в нуле всегда равен нулю, $J_{AB}|_{x=0} = 0$.

Пример 19.4.1. (Вращение твердого тела.) В трехмерном евклидовом пространстве с декартовыми координатами $x_i \in \mathbb{R}^3$, $i = 1, 2, 3$, которое мы отождествим с дуальной алгеброй Ли трехмерных вращений $\mathfrak{so}(3)^*$, скобка Ли–Пуассона равна

$$[x_i, x_j] = -\varepsilon_{ijk} x^k, \quad (19.37)$$

где ε_{ijk} – полностью антисимметричный тензор третьего ранга, а подъем и опускание индексов производится с помощью евклидовой метрики δ_{ij} . Эта структура вырождена, ее ранг равен двум всюду, кроме начала координат, где он равен нулю. Для пуассоновой структуры (19.37) существует единственная функция Казимира

$$c = x^i x_i.$$

Скобку Ли–Пуассона (19.37) можно сузить на сферу произвольного радиуса, соответствующую постоянному значению функции Казимира $c = \text{const}$. Соответствующая пуассонова структура на сфере $x^i x_i = \text{const} > 0$ невырождена. В качестве координат Дарбу, которые будут определены в следующем разделе, на сфере можно выбрать ось $z = x^3$ и полярный угол $\varphi = \text{arctg}(x^2/x^1)$. Используя определение (19.37), нетрудно проверить, что

$$[z, \varphi] = 1, \quad [\varphi, \varphi] = 0, \quad [z, z] = 0.$$

Таким образом, симплектическими слоями, соответствующими постоянной функции Казимира, являются сферы, а координатами Дарбу – цилиндрические координаты.

Скобке Ли–Пуассона (19.37) соответствует хорошо известный пример из механики твердого тела. Рассмотрим вращающееся твердое тело с покоящимся центром масс в декартовой системе координат, оси которой направлены по главным осям инерции. Гамильтониан системы в этом случае имеет вид

$$H(x) = \frac{x_1^2}{2I_1} + \frac{x_2^2}{2I_2} + \frac{x_3^2}{2I_3},$$

где x_i , $i = 1, 2, 3$ – моменты количества движения и $I_{1,2,3}$ – моменты инерции твердого тела. Если для координат x_i определить скобку Пуассона (19.37), то гамильтоновы уравнения движения примут вид

$$\begin{aligned} \dot{x}_1 &= \frac{I_2 - I_3}{I_2 I_3} x_2 x_3, \\ \dot{x}_2 &= \frac{I_3 - I_1}{I_1 I_3} x_1 x_3, \\ \dot{x}_3 &= \frac{I_1 - I_2}{I_1 I_2} x_1 x_2. \end{aligned}$$

Это есть *уравнения Эйлера* вращения твердого тела (см., например, [132], §36). $\square$

19.5 Отображения пуассоновых многообразий

Определение. Пусть $\varphi : \mathbb{M} \rightarrow \mathbb{N}$ – гладкое отображение двух пуассоновых многообразий. Это отображение называется *пуассоновым*, если оно сохраняет скобку Пуассона:

$$[f \circ \varphi, g \circ \varphi]_{\mathbb{M}} = [f, g]_{\mathbb{N}} \circ \varphi, \quad (19.38)$$

где f и g – произвольные функции на многообразии $\mathbb{N}$ и, следовательно, $f \circ \varphi$ и $g \circ \varphi$ – функции на $\mathbb{M}$. Соответственно, подмногообразие $\mathbb{M} \subset \mathbb{N}$ называется *пуассоновым*, если вложение $\mathbb{M} \hookrightarrow \mathbb{N}$ является пуассоновым отображением. В этом случае мы говорим, что пуассонова структура на $\mathbb{N}$ *сужена* до пуассоновой структуры на подмногообразии $\mathbb{M} \subset \mathbb{N}$.

С другой стороны, если пуассонова структура задана на многообразии $\mathbb{M}$, то ее всегда можно отобразить на образ $\varphi(\mathbb{M}) \subseteq \mathbb{N}$ с помощью дифференциала отображения. Эту пуассонову структуру на образе $\varphi(\mathbb{M}) \subset \mathbb{N}$ будем называть *индуцированной*.

В классической механике пуассоново отображение фазового пространства на себя называется *каноническим преобразованием*. $\square$

Выпишем структурные функции для индуцированной пуассоновой структуры. Пусть отображение $\varphi : \mathbb{M} \rightarrow \mathbb{N}$ в координатах задается функциями $y^A(x)$, где y^A , $A = 1, \dots, \dim \mathbb{N}$ и x^α , $\alpha = 1, \dots, \dim \mathbb{M}$, – координаты, соответственно, на $\mathbb{N}$ и $\mathbb{M}$. Тогда индуцированная пуассонова структура на образе $\varphi(\mathbb{M}) \subseteq \mathbb{N}$ всегда определена и имеет следующие структурные функции:

$$J^{AB} = J^{\alpha\beta} \partial_\alpha y^A \partial_\beta y^B.$$

Из свойств умножения матриц следует, что если отображение φ – вложение, $\text{rank } \partial_\alpha y^A = \dim \mathbb{M}$, то ранг индуцированной пуассоновой структуры J^{AB} равен рангу пуассоновой структуры на подмногообразии $\mathbb{M}$.

Рассмотренный ниже пример показывает, что пуассонову структуру на $\mathbb{N}$ в общем случае нельзя сузить на произвольное подмногообразие $\mathbb{M} \subset \mathbb{N}$.

Пример 19.5.1. Приведем простой пример непуассонова вложения, которое, на первый взгляд, должно быть таковым. Пусть многообразие $\mathbb{N}$ является прямым произведением двух многообразий, $\mathbb{N} := \mathbb{M} \times \mathbb{K}$. Обозначим координаты и пуассонову структуру на $\mathbb{M}$ через x^α , $\alpha = 1, \dots, \dim \mathbb{M}$, и $J^{\alpha\beta}$. Пусть на $\mathbb{K}$ также заданы координаты y^μ , $\mu = 1, \dots, \dim \mathbb{K}$, и пуассонова структура $J^{\mu\nu}$. Тогда матрица

$$J = \begin{pmatrix} J^{\alpha\beta} & 0 \\ 0 & J^{\mu\nu} \end{pmatrix}$$

определяет пуассонову структуру на $\mathbb{N}$ в координатах $\{x, y\}$. Это – пуассонова структура индуцирована двумя естественными вложениями $\mathbb{M} \hookrightarrow \mathbb{M} \times \mathbb{K}$ и $\mathbb{K} \hookrightarrow \mathbb{M} \times \mathbb{K}$.

Рассмотрим естественное вложение

$$\varphi : \mathbb{M} \ni x \mapsto (x, y_0) \in \mathbb{N} = \mathbb{M} \times \mathbb{K},$$

где $y_0 \in \mathbb{K}$ – произвольная фиксированная точка. Пусть на $\mathbb{N}$ заданы две произвольные функции $f(x, y)$ и $g(x, y)$. Вычислим левую и правую части равенства (19.38):

$$\begin{aligned} [f \circ \varphi, g \circ \varphi]_{\mathbb{M}} &= J^{\alpha\beta} \partial_\alpha f \partial_\beta g \big|_{y=y_0}, \\ [f, g]_{\mathbb{N}} \circ \varphi &= (J^{\alpha\beta} \partial_\alpha f \partial_\beta g + J^{\mu\nu} \partial_\mu f \partial_\nu g) \big|_{y=y_0}. \end{aligned}$$

Мы видим, что в общем случае равенство (19.38) не выполняется. Отсюда следует, что вложение φ не является пуассоновым. $\square$

Предложение 19.5.1. Пусть $(\mathbb{M}, J)$ – пуассоново многообразие и X_H – гамильтоново векторное поле. Тогда поток векторного поля

$$s_t := \exp(tX_H) : \mathbb{M} \ni x(0) \mapsto x(t) \in \mathbb{M},$$

где $x(t)$ – интегральная кривая для X_H , определяет пуассоново отображение.

Доказательство. Пусть $f, g \in \mathcal{C}^\infty(\mathbb{M})$. При малых t поток векторного поля действует на x следующим образом $s_t : x^\alpha \mapsto x^\alpha + tX_H^\alpha$. Поэтому

$$f \circ s_t = f(x^\alpha + tX_H^\alpha) \approx f(x) + tX_H f.$$

Продифференцируем по t условие пуассоновости (19.38) и положим $t = 0$. Тогда условие пуассоновости примет вид

$$[X_H f, g] + [f, X_H g] = X_H[f, g].$$

Из формулы (19.32) следует, что полученное равенство совпадает с тождествами Якоби. При $t = 0$ экспоненциальное отображение является тождественным и, следовательно, пуассоново. Поэтому интегрирование условия (19.38) вдоль X_H доказывает его выполнение при всех t , для которых определены интегральные кривые. $\square$

Пример 19.5.2. Пусть на фазовой плоскости $q, p \in \mathbb{R}^2$ задана каноническая пуассонова структура (19.33). Рассмотрим гармонический осциллятор с гамильтонианом

$$H = \frac{1}{2}(q^2 + p^2).$$

Соответствующее гамильтоново векторное поле имеет вид

$$X_H = q\partial_p - p\partial_q$$

и задает группу вращений плоскости. Таким образом, каждое вращение фазовой плоскости задает пуассоново преобразование для гармонического осциллятора. Оно является каноническим, т.к. сохраняет скобку Пуассона. $\square$

Поскольку любой гамильтонов поток сохраняет скобку Пуассона, то он, в частности сохраняет ее ранг. Поэтому справедливо

Следствие. Для любого гамильтонова векторного поля X_H на пуассоновом многообразии (M, J) ранг пуассоновой структуры J постоянен вдоль произвольной интегральной кривой для X_H . $\square$

Это следствие является, по существу, переформулировкой предложения 19.3.2.

Предложение 19.5.2. Если ранг пуассоновой структуры в какой либо точке $x \in M$ пуассонова многообразия (M, J) равен нулю, то эта точка является неподвижной для любой гамильтоновой системы H на M .

Доказательство. Из курса линейной алгебры известно, что если матрица антисимметрична и ее ранг равен нулю, то она сама равна нулю. Если в точке $x \in M$ ранг пуассоновой структуры равен нулю, то гамильтоново векторное поле X_H в этой точке для произвольного гамильтониана H обращается в нуль. Следовательно, точка x является неподвижной. $\square$

Пример 19.5.3. Пуассонова структура в начале координат в примере 19.4.1 имеет нулевой ранг. Она остается неподвижной для любого гамильтонова потока. $\square$

Пусть (M, J) – пуассоново многообразие. Тогда для каждой точки $x \in M$ существует единственное линейное отображение кокасательного пространства в соответствующее касательное

$$J : T_x^*(M) \rightarrow T_x(M), \quad (19.39)$$

такое, что для любой функции $f(x)$ справедливо равенство

$$J(df) = [f, x^\alpha] \partial_\alpha \in T_x(M).$$

Это есть рассмотренное ранее отображение (19.29). Для произвольной 1-формы $A = dx^\alpha A_\alpha$ отображение J в компонентах задается матрицей структурных функций:

$$X^\alpha = -J^{\alpha\beta} A_\beta.$$

Для симплектических многообразий, для которых матрица $J^{\alpha\beta}$ невырождена, отображение J является взаимно однозначным.

Ясно, что ядром отображения (19.39) является линейная оболочка дифференциалов функций Казимира dc^A .

Обозначим образ отображения (19.39) в точке $x \in \mathbb{M}$ через

$$\mathbb{J}_x(\mathbb{M}) := \{J(A) \in T_x(\mathbb{M}) : \forall A \in T_x^*(\mathbb{M})\}.$$

Размерность векторного подпространства $\mathbb{J}_x(\mathbb{M}) \subset T_x(\mathbb{M})$ равна рангу пуассоновой структуры $\dim \mathbb{J}_x(\mathbb{M}) = \text{rank } J^{\alpha\beta}(x)$. Если ранг пуассоновой структуры на $\mathbb{M}$ является постоянным и равен $2m$, то совокупность подпространств $\mathbb{J}_x(\mathbb{M})$ для всех точек x задает на $\mathbb{M}$ распределение векторных полей $\mathcal{J}_{2m}(\mathbb{M})$ размерности $2m$, которое было определено в разделе 2.11. При этом, если пуассонова структура дифференцируема, то соответствующее распределение векторных полей также дифференцируемо.

В общем случае ранг пуассоновой структуры может меняться от точки к точке. Из определения отображения J следует, что образ $\mathbb{J}_x(\mathbb{M})$ является линейной оболочкой всех гамильтоновых векторных полей на $\mathbb{M}$ в точке x .

Введенные понятия позволяют сформулировать следующее утверждение.

Теорема 19.5.1. *Подмногообразие $\mathbb{M}$ пуассонова многообразия $(\mathbb{N}, J)$ является пуассоновым тогда и только тогда, когда $\mathbb{J}_x(\mathbb{M}) := \mathbb{J}_x(\mathbb{N})|_{\mathbb{M}} \subset T_x(\mathbb{M})$ для всех $x \in \mathbb{M}$, т.е. каждое гамильтоново векторное поле на $\mathbb{N}$ всюду касается $\mathbb{M}$. В частности, если $\mathbb{J}_x(\mathbb{M}) = T_x(\mathbb{M})$ для всех $x \in \mathbb{M}$, то $\mathbb{M}$ является симплектическим подмногообразием в $\mathbb{N}$.*

Доказательство. См., например, [133], предложение 6.19. □

Из этой теоремы следует, что, поскольку размерность пространства $\mathbb{J}_x(\mathbb{M})$ совпадает с рангом пуассоновой структуры в данной точке, то размерность пуассонова подмногообразия не может быть меньше ранга, $\dim \mathbb{M} \geq \text{rank } J$.

Пусть $\mathbb{M} \subset \mathbb{N}$ – пуассоново подмногообразие. Как было отмечено в теореме 19.5.1, любое гамильтоново векторное поле X_H на $\mathbb{N}$ касается подмногообразия $\mathbb{M}$. Это означает, что его сужение на $\mathbb{M}$ может быть получено из сужения гамильтониана на $\mathbb{M}$:

$$X_H|_{\mathbb{M}} = X_{\tilde{H}}, \quad \text{где} \quad \tilde{H} := H|_{\mathbb{M}}.$$

Допустим, что нас интересуют траектории движения гамильтоновой системы с некоторым гамильтонианом H , которые начинаются в некоторой точке пуассонова подмногообразия $x \in \mathbb{M} \subset \mathbb{N}$. В этом случае можно ограничиться движением точки в подмногообразии $\mathbb{M}$, которое порождается суженным гамильтонианом $\tilde{H}$, тем самым понизив порядок гамильтоновой системы. Можно поставить вопрос, каково минимальное пуассоново подмногообразие для данных начальных условий. Ответ дает следующая

Теорема 19.5.2. *Пусть $(\mathbb{N}, J)$ – пуассоново многообразие. Тогда соответствующее распределение гамильтоновых векторных полей $\mathcal{J}(\mathbb{N})$ на $\mathbb{N}$ интегрируемо, т.е. через каждую точку $x \in \mathbb{N}$ проходит интегральное подмногообразие $\mathbb{M}$ распределения $\mathcal{J}(\mathbb{N})$, для которого $T_y(\mathbb{M}) = \mathbb{J}_y(\mathbb{M})$ в любой точке $y \in \mathbb{M}$. Всякое интегральное подмногообразие является симплектическим подмногообразием в $\mathbb{N}$, и в совокупности эти подмногообразия определяют симплектическое слоение пуассонова многообразия $\mathbb{N}$. Кроме того, если $H : \mathbb{N} \rightarrow \mathbb{R}$ – произвольный гамильтониан и $x(t)$ – соответствующая траектория системы, проходящая через точку $x_0 \in \mathbb{N}$, то $x(t)$ остается в одном и том же интегральном подмногообразии $\mathbb{M}$ при всех t .*

Доказательство. Как уже отмечалось, распределение $\mathcal{J}(\mathbb{M})$ является линейной оболочкой всех гамильтоновых векторных полей на $\mathbb{M}$. Поскольку скобка Пуассона гамильтоновых векторных полей является гамильтоновым векторным полем (19.30), то распределение $\mathcal{J}(\mathbb{M})$ находится в инволюции. Отсюда по теореме Фробениуса следует существование интегрального подмногообразия. Остальные утверждения теоремы следуют из теоремы 19.5.1 и инвариантности ранга пуассоновой структуры вдоль гамильтоновых векторных полей. $\square$

Проиллюстрируем данную теорему для пуассонова многообразия $(\mathbb{N}, J)$, $\dim \mathbb{N} = n$, с пуассоновой структурой постоянного ранга $\text{rank } J^{\alpha\beta} = 2m < n$ путем построения специальной системы координат. В некоторой окрестности $\mathbb{U} \subset \mathbb{N}$ существует $(n - 2m)$ функционально независимых функций Казимира $c^A(x)$, $A = 1, \dots, n - 2m$. Поверхности уровня функций Казимира $c^A = \text{const}$ определяют $2m$ -мерное подмногообразие $\mathbb{M} \subset \mathbb{U}$. Дополним функции Казимира $2m$ функционально независимыми скалярными полями g^M , $M = 1, \dots, 2m$, (координаты на подмногообразии $\mathbb{M}$) таким образом, чтобы выполнялись условия

$$\text{rank } [g^M, g^N]_{\mathbb{M}} = 2m.$$

Это всегда можно сделать, т.к. ранг пуассоновой структуры равен $2m$. Совокупность функций $\{c^A, g^M\}$ выберем в качестве новой системы координат на $\mathbb{U}$. Пуассонова структура в этой системе координат примет вид

$$J = \begin{pmatrix} J^{AB} = [c^A, c^B] & J^{AN} = [c^A, g^N] \\ J^{MB} = [g^M, c^B] & J^{MN} = [g^M, g^N] \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & J^{MN} \end{pmatrix},$$

т.к. скобка Пуассона функции Казимира с любой функцией на $\mathbb{N}$ равна нулю. Таким образом, сужение Пуассоновой структуры на подмногообразии $\mathbb{M}$ определено и невырождено. То есть поверхности уровней функций Казимира представляют собой симплектические многообразия. Пусть на $\mathbb{N}$ задан гамильтониан. Соответствующие гамильтоновы уравнения

$$\dot{c}^A = J^{AB} \partial_B H + J^{AN} \partial_N H = 0, \quad (19.40)$$

$$\dot{g}^M = J^{MB} \partial_B H + J^{MN} \partial_N H = J^{MN} \partial_N H \quad (19.41)$$

определяют траектории системы. Поскольку уравнение (19.40) имеет решение $c^A = \text{const}$, то траектория гамильтоновой системы, проходящей через точку $x_0 \in \mathbb{N}$, принадлежит соответствующему симплектическому подмногообразию $\mathbb{M}$:

$$\dot{g}^M = \tilde{J}^{MN} \partial_N \tilde{H},$$

где

$$\tilde{J}^{MN} = J^{MN}|_{\mathbb{M}} \quad \text{и} \quad \tilde{H} = H|_{\mathbb{M}}.$$

Таким образом, всякое пуассоново многообразие $(\mathbb{N}, J)$ расщепляется на симплектические подмногообразия – слои симплектического слоения. Размерность любого такого слоя $\mathbb{M}$ равна рангу пуассоновой структуры в произвольной точке $x \in \mathbb{M}$. Это означает, что, если ранг пуассоновой структуры на $\mathbb{N}$ непостоянен, то размерность симплектических слоев будет различной.

Теорема 19.5.3 (Дарбу). Пусть $(\mathbb{M}, J)$ – пуассоново многообразие размерности n . Если пуассонова структура на многообразии имеет постоянный ранг, $\text{rank } J =$

$2m \leq n$, то в некоторой окрестности $\mathbb{U}$ произвольной точки $x \in \mathbb{M}$ существует такая система координат

$$\{x^\alpha\} = \{q^m, p_m, c^A\} = (q^1, \dots, q^m, p_1, \dots, p_m, c^1, \dots, c^{n-2m}),$$

в которой скобка Пуассона двух произвольных функций $f, g \in C^\infty(\mathbb{U})$ имеет вид

$$[f, g] = \frac{\partial f}{\partial q^m} \frac{\partial g}{\partial p_m} - \frac{\partial f}{\partial p_m} \frac{\partial g}{\partial q^m}. \quad (19.42)$$

Координаты c^A являются функциями Казимира пуассоновой структуры J , и их постоянные значения, $c_A = \text{const}$, определяет симплектическое слоение пуассонова многообразия $(\mathbb{U}, J)$.

Доказательство. Если ранг пуассоновой структуры равен нулю, то $J^{\alpha\beta} = 0$ и доказывать нечего. Любая система координат на $\mathbb{U}$ удовлетворяет утверждению теоремы.

Допустим, что ранг отличен от нуля. Зафиксируем точку $x_0 \in \mathbb{M}$ и выберем в ее окрестности $\mathbb{U}$ две такие функции $Q, f \in C^\infty(\mathbb{U})$, что

$$[Q, f] = X_Q f \neq 0.$$

В частности $X_Q|_{x_0} \neq 0$. Согласно теореме 2.6.3, возможно, в меньшей окрестности существует такая система координат $Q, x^2, \dots, x^n$, что $X_Q = \partial_Q$. Тогда существует такая функция $P \in C^\infty(\mathbb{U})$, что выполнено условие

$$X_Q P = [Q, P] = 1.$$

Из предложения 19.3.1 следует равенство

$$[X_Q, X_P] = X_{[Q, P]} = 0.$$

Кроме того, в силу антисимметрии $[X_Q, X_Q] = 0$ и $[X_P, X_P] = 0$. Таким образом, X_Q, X_P образуют пару коммутирующих векторных полей. Теорема Фробениуса позволяет дополнить пару функций $q^1 := Q, p_1 := P$ до локальной системы координат $q^1, p_1, y^3, \dots, y^n$ в некоторой, возможно, меньшей окрестности точки x_0 . Поскольку $J^{\alpha\beta} = [x^\alpha, x^\beta]$ и выполнены равенства $[q^1, y^i] = 0, [p_1, y^i] = 0, i = 3, \dots, n$, то структурные функции в данной системе координат принимают вид

$$J = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & J^{ij} \end{pmatrix},$$

где $J^{ij} = [y^i, y^j]$. Покажем, что матрица J^{ij} не зависит от координат q^1 и p_1 . Действительно,

$$\frac{\partial J^{ij}}{\partial q^1} = [q^1, J^{ij}] = [q^1, [y^i, y^j]] = 0,$$

где мы воспользовались тождествами Якоби. Аналогично доказывается, что матрица J^{ij} не зависит от p_1 . Таким образом, матрица J^{ij} задает пуассонову структуру на подмногообразии $q^1 = \text{const}, p_1 = \text{const}$. Ранг этой структуры на два меньше исходного, $\text{rank } J^{ij} = 2m - 2$. Поэтому, если $m > 1$, то этот процесс можно продолжить. $\square$

Глава 20

Принцип наименьшего действия

Можно с уверенностью сказать, что в основе построения моделей современной математической физики лежит принцип наименьшего действия. Этот принцип требует стационарности некоторого функционала – действия – относительно вариаций полей, описывающих данную модель. В результате мы получаем систему уравнений Эйлера–Лагранжа, которая принимается в качестве уравнений движения, уравнений равновесия и т.д. для данной модели. При этом инвариантность действия относительно некоторой группы преобразований приводит к ковариантным уравнениям движения и к законам сохранения, которые играют важнейшую роль в физике.

20.1 Постановка вариационных задач

Начнем с постановки вариационной задачи в евклидовом пространстве. Предположим, для простоты, что $\mathbb{M} \subset \mathbb{R}^n$ – ограниченная область евклидова пространства с достаточно гладкой границей $\partial\mathbb{M}$. Пусть в этой области задан некоторый набор дважды непрерывно дифференцируемых функций вплоть до границы, $\varphi = \{\varphi^a\} \in \mathcal{C}^2(\overline{\mathbb{M}})$, $a = 1, \dots, N$. Это значит, что все функции и их производные до второго порядка непрерывны и ограничены в $\mathbb{M}$ и имеют конечный предел на границе. Если x^α , $\alpha = 1, \dots, n$, – система координат на $\mathbb{M}$, то обозначим, для краткости, все первые производные полей через $\partial\varphi = \{\partial_\alpha \varphi^a\}$. Предположим, что на $\mathbb{M}$ определен *функционал действия* или, просто, *действие*

$$S[\varphi] = \int_{\mathbb{M}} dx L(x, \varphi, \partial\varphi), \quad (20.1)$$

где L – некоторая функция от n переменных x^α , N переменных φ^a и nN переменных $\partial_\alpha \varphi^a$. Она предполагается дважды непрерывно дифференцируемой функцией переменных $x \in \mathbb{M}$ и всех остальных переменных для всех конечных значений φ и $\partial\varphi$. Назовем функцию $L(x, \varphi, \partial\varphi)$ *лагранжевой плотностью* или *лагранжианом* данной модели, которая описывается набором полей φ .

Для определения функционала действия мы ограничили себя классом дважды непрерывно дифференцируемых функций $\mathcal{C}^2(\overline{\mathbb{M}})$. Тогда функционал действия задает отображение

$$S : [\mathcal{C}^2(\overline{\mathbb{M}})]^N \ni \varphi \mapsto S[\varphi] \in \mathbb{R} \quad (20.2)$$

бесконечномерного функционального пространства в поле вещественных чисел. Для постановки вариационной задачи нам важно, что функциональное пространство

$[\mathcal{C}^2(\overline{\mathbb{M}})]^N$ снабжено структурой линейного пространства с обычным поточечным сложением и умножением на вещественные числа. Для того, чтобы говорить о непрерывности и вариационных производных отображения S , введем на $[\mathcal{C}^2(\overline{\mathbb{M}})]^N$ норму:

$$\|\varphi\| := \sum_{a=1}^N \left(\sup_{x \in \overline{\mathbb{M}}} |\varphi^a| + \sum_{\alpha=1}^n \sup_{x \in \overline{\mathbb{M}}} |\partial_\alpha \varphi^a| + \sum_{\alpha, \beta=1}^n \sup_{x \in \overline{\mathbb{M}}} |\partial_\alpha \partial_\beta \varphi^a| \right). \quad (20.3)$$

Тем самым класс рассматриваемых функций $[\mathcal{C}^2(\overline{\mathbb{M}})]^N$ становится нормированным линейным функциональным пространством. По данной норме строится метрика и определяется топология, относительно которой отображение (20.2) непрерывно.

Мы потребовали от функций φ^a дважды непрерывную дифференцируемость, чтобы уравнения Эйлера–Лагранжа, которые будут получены позже, были определены в классическом смысле. Это пространство функций полно по норме (20.3).

Назовем *вариацией* функции φ^a разность двух представителей из класса рассматриваемых функций: $\delta\varphi^a := \varphi'^a - \varphi^a$, где φ'^a – произвольная функция из $\mathcal{C}^2(\overline{\mathbb{M}})$. Ясно, что вариация функции принадлежит тому же классу, что и сама функция. Для малых вариаций функций $\epsilon\delta\varphi^a$, где $\epsilon > 0$ – малая величина, вариация (главная линейная часть) функционала действия, если она существует, равна

$$\begin{aligned} \delta S &= S[\varphi + \epsilon\delta\varphi] - S[\varphi] = \\ &= \int_{\mathbb{M}} dx \left[\frac{\partial L}{\partial \varphi^a} - \partial_\alpha \left(\frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \right) \right] \epsilon\delta\varphi^a + \int_{\partial\mathbb{M}} ds_\alpha \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \epsilon\delta\varphi^a + o(\epsilon), \end{aligned} \quad (20.4)$$

где второе слагаемое возникло при интегрировании по частям, и ds_α обозначает ориентированный элемент объема границы $\partial\mathbb{M}$. Отметим, что при вычислении вариации действия (20.1) область интегрирования $\mathbb{M}$ считалась неизменной.

В рассматриваемом классе функций φ и лагранжианов L вариация функционала всегда существует.

Назовем набор функций φ *стационарной*, или *критической* точкой, или *экстремалью* функционала $S[\varphi]$, если в этой точке линейная часть вариации действия равна нулю, $\delta S = o(\epsilon)$. Такие точки соответствуют либо локальному минимуму, либо локальному максимуму, либо седловой точке функционала S , если “матрица” вторых производных (гессиан) невырождена. Это можно проверить после нахождения экстремали функционала, рассмотрев члены более высокого порядка по ϵ .

Для действия (20.1) можно поставить различные вариационные задачи. Рассмотрим задачи, которые наиболее часто встречаются в физике.

20.1.1 Задача с заданными граничными условиями

Вариационная задача с заданными граничными условиями является наиболее распространенной и самой простой с точки зрения постановки. Рассмотрим класс функций в $\mathcal{C}^2(\overline{\mathbb{M}})$ с заданными граничными условиями

$$\varphi|_{\partial\mathbb{M}} = \varphi_0. \quad (20.5)$$

Поскольку граничные условия для всех функций при фиксированном индексе a одни и те же, то вариации полей обращаются в нуль на границе:

$$\delta\varphi|_{\partial\mathbb{M}} = 0. \quad (20.6)$$

Тогда интеграл по границе области в вариации действия (20.4) обращается в нуль в силу граничных условий (20.6). В рассматриваемом случае существует предел

$$\lim_{\epsilon \rightarrow 0} \frac{S[\varphi + \epsilon \delta \varphi] - S[\varphi]}{\epsilon} = \int_{\mathbb{M}} dx \frac{\delta S}{\delta \varphi^a} \delta \varphi^a. \quad (20.7)$$

Функция $\delta S / \delta \varphi^a$, стоящая под знаком интеграла, называется *вариационной производной* функционала S по полю φ^a и обозначается также запятой:

$$S,{}_a := \frac{\delta S}{\delta \varphi^a}. \quad (20.8)$$

Из вида вариации (20.4) получаем явное выражение для вариационной производной

$$S,{}_a = \frac{\partial L}{\partial \varphi^a} - \partial_\alpha \frac{\partial L}{\partial (\partial_\alpha \varphi^a)}. \quad (20.9)$$

Вариационная производная $S,{}_a$ является ядром линейного оператора, который есть производная по Фреше и, следовательно, по Гато отображения (20.2).

Формула (20.7) позволяет дать другое определение вариационной производной. А именно, вариационной производной функционала S в точке $\varphi \in [\mathcal{C}^2(\overline{\mathbb{M}})]^N$ назовем производную

$$\frac{\delta S}{\delta \varphi^a} := \left. \frac{\partial S(\varphi + \lambda \psi)}{\partial \lambda} \right|_{\lambda=0}, \quad (20.10)$$

где $\psi \in [\mathcal{C}^2(\overline{\mathbb{M}})]^N$. Ясно, что если главная линейная часть приращения функционала существует, то она равна производной (20.10). Обратное утверждение в общем случае неверно. Существуют примеры функционалов (менее гладких, чем мы рассматриваем), для которых производная (20.10) определена, однако из их приращения нельзя выделить главную линейную часть. Поэтому определение вариационной производной (20.10) является более общим. В рассматриваемом нами классе функций и лагранжианов данные выше определения эквивалентны.

Определение вариационной производной (20.10) просто обобщается на случай вариационных производных более высокого порядка. Вторые вариационные производные (16.35) были рассмотрены для экстремалей функционала длины кривой.

Из условия стационарности действия $\delta S = o(\epsilon)$ в силу произвольности вариации $\delta \varphi^a$ и основной леммы вариационного исчисления следует

Теорема 20.1.1. *Набор функций φ при заданных граничных условиях является стационарной точкой действия $S[\varphi]$ тогда и только тогда, когда он удовлетворяет системе уравнений Эйлера–Лагранжа*

$$S,{}_a = \frac{\partial L}{\partial \varphi^a} - \partial_\alpha \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} = 0. \quad (20.11)$$

В общем случае уравнения Эйлера–Лагранжа представляют собой систему нелинейных дифференциальных уравнений в частных производных второго порядка, число которых N равно числу функций, от которых зависит функционал действия. Их надо решать при заданных граничных условиях (20.5). Решение поставленной вариационной задачи может не существовать, а если оно существует, то может быть не единственно. Это зависит от вида лагранжиана и области $\mathbb{M}$.

Замечание. В общем случае уравнения Эйлера–Лагранжа для заданного лагранжиана могут приводить к противоречию. Например, пусть лагранжиан зависит от одной функции φ и имеет вид $L := \varphi$. Тогда уравнения Эйлера–Лагранжа приводят к противоречию $1 = 0$. Таким образом, для того, чтобы уравнения Эйлера–Лагранжа имели решение, лагранжиан не может быть произвольной функцией полей и их производных. В дальнейшем мы предполагаем, что лагранжиан выбран таким образом, что соответствующие уравнения Эйлера–Лагранжа непротиворечивы. $\square$

Действие (20.1) для заданных уравнений Эйлера–Лагранжа при постановке задачи с фиксированными граничными условиями определено неоднозначно. Действительно, рассмотрим новый лагранжиан, $\tilde{L}_\alpha = L + \partial_\alpha F$, который отличается от исходного на частную производную от некоторой достаточно гладкой функции $F(x, \varphi, \partial\varphi)$. Тогда действие получит дополнительный вклад, сводящийся к интегралу по границе. Вариация дополнительного слагаемого равна нулю, т.к. вариации всех полей на границе равны нулю. Отсюда следует, что уравнения Эйлера–Лагранжа не изменятся при добавлении к лагранжиану частных производных $\partial_\alpha F$ от произвольной функции.

Рассмотренная вариационная задача наиболее часто рассматривается в моделях математической физики. При этом уравнения Эйлера–Лагранжа приводят к уравнениям движения, равновесия и т.д.

Замечание. Поверхностный интеграл в вариации действия (20.4) в задаче с заданными граничными условиями равен нулю, поскольку вариации полей обращаются в нуль на границе (20.6). Этого достаточно, если область $\mathbb{M}$ ограничена. Однако для действия (20.1), рассматриваемого во всем евклидовом пространстве $\mathbb{R}^n$, ситуация усложняется. В этом случае интеграл по границе является несобственным, т.к. площадь бесконечно удаленной границы стремится к бесконечности. Тогда важно не только граничное условие на вариации, но и их асимптотическое поведение. Так, в общей теории относительности в асимптотически плоском пространстве-времени поверхностный интеграл отличен от нуля, т.к. метрика недостаточно быстро стремится к метрике Минковского на пространственной бесконечности. Анализ граничного поведения полей важен и в общем случае сложен, поскольку зависит от рассматриваемой задачи. На данном этапе мы пренебрежем граничными слагаемыми. $\square$

Замечание. Во многих моделях математической физики, например, в электродинамике, в качестве области $\mathbb{M}$ выбирается все пространство Минковского $\mathbb{R}^{1,3}$. При этом функционал действия не ограничен ни снизу, ни сверху. Кроме того, для многих решений уравнений Эйлера–Лагранжа он расходится. Поэтому говорить о принципе наименьшего действия в строгом смысле не приходится. Тем не менее уравнения Эйлера–Лагранжа имеют смысл, поскольку являются локальным объектом. Поэтому для действия часто пишут формальное выражение, не заботясь о сходимости интеграла. Этот способ очень удобен для получения уравнений с заданными свойствами симметрии. $\square$

20.1.2 Задача со свободными граничными условиями

Для действия (20.1) можно поставить другую вариационную задачу, расширив класс рассматриваемых функций. Пусть, по-прежнему, все функции дважды непрерывно дифференцируемы $\varphi^a \in C^2(\overline{\mathbb{M}})$ для всех значений $a = 1, \dots, N$, но теперь снимем ограничения, накладываемые граничными условиями (20.5). В этом случае вариации

полей $\delta\varphi^a$ также ничем не ограничены на границе $\partial\mathbb{M}$, и из явного вида вариации действия (20.4) следует

Теорема 20.1.2. *Набор функций φ является стационарной точкой действия $S[\varphi]$ тогда и только тогда, когда он удовлетворяет системе уравнений Эйлера–Лагранжа (20.11) и граничным условиям*

$$\left. \frac{\partial L}{\partial(\partial_\alpha \varphi^a)} \right|_{\partial\mathbb{M}} = 0. \quad (20.12)$$

Таким образом, в задаче со свободными граничными условиями граничные условия все таки возникают из условия стационарности действия.

Пример 20.1.1. Вариационная задача со свободными граничными условиями рассматривается в теории открытых бозонных струн, для которых из вариационного принципа вытекают граничные условия Неймана. $\square$

Постановка вариационной задачи со свободными граничными условиями, зависит от добавления к действию граничных слагаемых. Иногда появления граничных условий (20.12) можно избежать, если из исходного действия вычесть все граничные вклады, которые возникают при интегрировании по частям. В этом случае условие (20.12) тождественно выполняется.

Замечание. В общем случае можно рассматривать смешанные вариационные задачи, когда граничные условия ставятся только для части полей. Именно такого типа задача естественным образом возникает в теории гравитации, где можно считать заданными на границе только физические поля. Нефизические поля находятся, как решение уравнений связей и калибровочных условий, и для них граничные значения не могут быть фиксированы произвольным образом. $\square$

20.1.3 Задача с подвижной границей

Возможна также более общая постановка вариационной задачи для функционала (20.1), когда рассматриваются не только вариации полей, но и вариации самой области $\mathbb{M}$. Сначала уточним постановку задачи. Предположим, что координаты и поля преобразуются следующим образом:

$$x^\alpha \mapsto x'^\alpha(x, \varphi, \partial\varphi, \epsilon), \quad (20.13)$$

$$\varphi^a \mapsto \varphi'^a(x, \varphi, \partial\varphi, \epsilon), \quad (20.14)$$

где $\epsilon \in \mathbb{R}$ – параметр преобразования. Если поля φ являются заданными функциями от координат, $\varphi = \varphi(x)$, то из уравнения (20.13) можно выразить старые координаты через новые: $x = x(x', \epsilon)$. Тогда подстановка найденных функций в формулу (20.14) позволяет рассматривать новые поля, как функции от новых координат, $\varphi' = \varphi'(x', \epsilon)$.

Мы считаем, что до и после преобразования координаты определены, соответственно, на ограниченных областях $\mathbb{M}$ и $\mathbb{M}'$ евклидова пространства $\mathbb{R}^n$, т.е. $x \in \mathbb{M}$ и $x' \in \mathbb{M}'$. Тогда под вариацией функционала подразумевается разность

$$\delta S := S[\varphi'(x')] - S[\varphi(x)] = \int_{\mathbb{M}'} dx' L(x', \varphi', \partial' \varphi') - \int_{\mathbb{M}} dx L(x, \varphi, \partial \varphi).$$

Нам требуется найти эту вариацию в линейном по ϵ приближении. При этом мы считаем, что при $\epsilon = 0$ преобразование координат и полей является тождественным. Разлагая формулы преобразования в ряд Тейлора при малых ϵ , получаем

$$\begin{aligned} x^\alpha &\mapsto x'^\alpha = x^\alpha + \delta x^\alpha + \mathbf{o}(\epsilon), \\ \varphi^a(x) &\mapsto \varphi'^a(x') = \varphi^a(x) + \bar{\delta}\varphi^a(x) + \mathbf{o}(\epsilon), \end{aligned} \quad (20.15)$$

где независимые вариации

$$\delta x^\alpha = \epsilon R^\alpha(x, \varphi, \partial\varphi), \quad (20.16)$$

$$\bar{\delta}\varphi^a = \varphi'^a(x') - \varphi^a(x) = \epsilon R^a(x, \varphi, \partial\varphi) \quad (20.17)$$

определяются некоторыми достаточно гладкими функциями $R^\alpha(x, \varphi, \partial\varphi)$ и $R^a(x, \varphi, \partial\varphi)$. Отметим, что вариация полей (20.17) определена, как разность значений полей в различных точках и не совпадает с вариацией формы поля, рассмотренной в разделе 2.13. Напомним, что под вариацией формы функции мы понимаем разность значений этой функции после и до преобразования в одной и той же точке:

$$\delta\varphi^a(x) := \varphi'^a(x) - \varphi^a(x).$$

При этом преобразования (20.15) рассматриваются, как активные (см. раздел 1.5). Тогда для вариации формы функции справедливо равенство

$$\delta\varphi^a(x) = \bar{\delta}\varphi^a - \delta x^\alpha \partial_\alpha \varphi^a = \epsilon(R^a - R^\alpha \partial_\alpha \varphi^a). \quad (20.18)$$

По построению, для постоянного параметра преобразования вариация формы функции δ перестановочна с операцией частного дифференцирования ∂_α . Вариация действия относительно преобразований (20.15) имеет вид

$$\delta S = \int_{\mathbb{M}'} dx' L(x') - \int_{\mathbb{M}} dx L(x) = \int_{\mathbb{M}'} dx' (L(x) + \bar{\delta}L(x)) - \int_{\mathbb{M}} dx L(x),$$

где $L(x) = L(x, \varphi(x), \partial_\alpha \varphi(x))$. Здесь и в дальнейшем мы будем рассматривать вариацию действия только с точностью до слагаемых, линейных по ϵ , и не будем это указывать. Вариации лагранжиана и элемента объема равны, соответственно,

$$\bar{\delta}L(x) = \frac{\partial L}{\partial x^\alpha} \delta x^\alpha + \frac{\partial L}{\partial \varphi^a} \bar{\delta}\varphi^a + \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \bar{\delta}(\partial_\alpha \varphi^a) = \delta L + \frac{dL}{dx^\alpha} \delta x^\alpha, \quad (20.19)$$

$$dx' \approx dx + dx \frac{\partial(\delta x^\alpha)}{\partial x^\alpha}, \quad (20.20)$$

где

$$\frac{dL}{dx^\alpha} = \frac{\partial L}{\partial x^\alpha} + \frac{\partial L}{\partial \varphi^a} \partial_\alpha \varphi^a + \frac{\partial L}{\partial (\partial_\beta \varphi^a)} \partial_{\alpha\beta}^2 \varphi^a.$$

Используя преобразование (20.20), перейдем от интегрирования по $\mathbb{M}'$ к интегрированию по исходной области $\mathbb{M}$ и запишем вариацию действия в виде

$$\delta S = \int_{\mathbb{M}} dx \left(\delta L + \frac{dL}{dx^\alpha} \delta x^\alpha + L \frac{\partial(\delta x^\alpha)}{\partial x^\alpha} \right) = \int_{\mathbb{M}} dx (\delta L + \partial_\alpha (L \delta x^\alpha)). \quad (20.21)$$

Теперь перепишем вариацию формы лагранжиана:

$$\begin{aligned} \delta L &= \frac{\partial L}{\partial \varphi^a} \delta\varphi^a + \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \delta(\partial_\alpha \varphi^a) = \\ &= \left(\frac{\partial L}{\partial \varphi^a} - \partial_\alpha \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \right) \delta\varphi^a + \partial_\alpha \left(\frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \delta\varphi^a \right). \end{aligned} \quad (20.22)$$

При этом существенно, что параметр преобразования постоянен, $\epsilon = \text{const}$, и, следовательно,

$$\delta(\partial_\alpha \varphi^a) = \partial_\alpha(\delta \varphi^a).$$

Воспользовавшись формулой Стокса, дивергентные слагаемые можно переписать в виде поверхностного интеграла. Таким образом для вариации действия получаем окончательное выражение

$$\delta S = \int_{\mathbb{M}} dx \left(\frac{\partial L}{\partial \varphi^a} - \partial_\alpha \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \right) \delta \varphi^a + \int_{\partial \mathbb{M}} ds_\alpha \left(\frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \delta \varphi^a + L \delta x^\alpha \right). \quad (20.23)$$

Поверхностный интеграл перепишем с учетом выражения для вариации формы функции (20.18):

$$\int_{\partial \mathbb{M}} ds_\alpha \left(\frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \bar{\delta} \varphi^a + \left(L \delta_\beta^\alpha - \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \partial_\beta \varphi^a \right) \delta x^\beta \right).$$

поскольку вариации координат δx^α и функций $\bar{\delta} \varphi^a$ произвольны и независимы, то из полученного выражения для вариации действия вытекает

Теорема 20.1.3. *Набор функций φ является стационарной точкой действия $S[\varphi]$ в вариационной задаче с подвижной границей тогда и только тогда, когда он удовлетворяет системе уравнений Эйлера–Лагранжа (20.11) и граничным условиям:*

$$\begin{aligned} \left. \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \right|_{\partial \mathbb{M}} &= 0, \\ L \delta_\beta^\alpha - \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} \partial_\beta \varphi^a \Big|_{\partial \mathbb{M}} &= 0. \end{aligned} \quad (20.24)$$

В последней вариационной задаче с подвижной границей мы имеем $n(n + N)$ граничных условий, число которых быстро растет с увеличением размерности пространства. Соответствующая вариационная задача далеко не всегда имеет решение. Чтобы уменьшить число независимых граничных условий, предположим, что в окрестности границы $\partial \mathbb{M}$ поля принимают наперед заданные значения:

$$\varphi^a(x) = \Phi^a(x),$$

где достаточно гладкие функции Φ^a заданы в некоторой окрестности $\partial \mathbb{M}$. Тогда вариации функций и координат связаны соотношением

$$\bar{\delta} \varphi^a \Big|_{\partial \mathbb{M}} = \delta x^\alpha \partial_\alpha \Phi^a \Big|_{\partial \mathbb{M}}.$$

В этом случае вместо $n(n + N)$ граничных условий (20.24) остаются n^2 условий:

$$\left[L \delta_\beta^\alpha + \frac{\partial L}{\partial (\partial_\alpha \varphi^a)} (\partial_\beta \Phi^a - \partial_\beta \varphi^a) \right]_{\partial \mathbb{M}} = 0. \quad (20.25)$$

Эти граничные условия называются *условиями трансверсальности*.

Стационарные точки действия для задачи с подвижной границей являются также стационарными точками для двух задач с фиксированной областью $\mathbb{M}$, рассмотренных ранее. Выполнение уравнений Эйлера–Лагранжа является необходимым условием во всех трех рассмотренных вариационных задачах. В задачах со свободными граничными условиями и с подвижной границей к уравнениям Эйлера–Лагранжа добавляются граничные условия.

20.1.4 Задача на условную стационарную точку

В настоящем разделе мы ограничимся рассмотрением вариационных задач на конечном отрезке $[x_1, x_2] \subset \mathbb{R}$ с заданными граничными условиями. Пусть требуется найти стационарную точку действия (20.1) в классе функций $\varphi^a \in \mathcal{C}^2([x_1, x_2])$, $a = 1, \dots, N$ при наличии $M < N$ независимых дополнительных условий, которые называются *связями*

$$G_A(x, \varphi, \partial\varphi) = 0, \quad A = 1, \dots, M < N, \quad (20.26)$$

где G_A – достаточно гладкие функции своих аргументов. Мы предполагаем, что связи G_A не противоречат граничным условиям и функционально независимы, т.е. ни одна из связей не является следствием остальных. В частности, ни одна из связей не выполняется тождественно для всех функций φ . Функциональная независимость связей означает, что матрица производных

$$\frac{\partial G_A}{\partial(\varphi^a, \partial_x \varphi^b)}$$

имеет постоянный ранг M . Отсюда следует, что локально связи можно разрешить относительно M функций или их первых производных, рассматривая остальные $2(N - M)$ функций и их производные как независимые.

В общем случае связи являются дифференциальными уравнениями, и их решения содержат произвольные постоянные. Мы предполагаем, что этот произвол устранен, например, наложением граничных условий, либо каким либо иным образом.

В частном случае связи могут быть алгебраическими уравнениями на неизвестные функции $G_A(x, \varphi) = 0$. В этом случае они называются *голономными*. В общем случае связи (20.26) называются *неголономными*.

При наличии связей вариации функций не являются независимыми, и выполнение уравнений Эйлера–Лагранжа для исходного действия (20.1) не является необходимым условием. Прямым способом решения задачи на условный экстремум является явное разрешение связей относительно M функций, подстановка полученного решения в исходное действие и исследование нового действия от $N - M$ функций на безусловный экстремум.

Задачи на условную стационарную точку часто встречаются в математической физике. В частности, к ним приводят все калибровочные модели, инвариантные относительно локальных преобразований полей. В связи с этим введем удобную терминологию, которая часто используется в физике. А именно, назовем поля *нефизическими*, если связи разрешаются относительно этих полей. Остальные поля, относительно которых после исключения нефизических полей возникает задача на безусловную стационарную точку, называются *физическими*. Деление полей на физические и нефизические условно, т.к. связи можно разрешать относительно различных переменных. В то же время число физических ($N - M$) и нефизических (M) полей, по предположению, постоянно.

Прямой способ исключения нефизических полей неприменим, если связи не решаются явно. Кроме этого исключение части полей может нарушить симметрию задачи, например, лоренц-инвариантность, что часто приводит к существенному усложнению вычислений. Поэтому используют метод неопределенных множителей Лагранжа. А именно, строят полное (total) действие

$$S_T := \int_{x_1}^{x_2} dx (L - \lambda^A G_A), \quad (20.27)$$

где $\lambda(x) \in \mathcal{C}^1([x_1, x_2])$ – новые функции, которые называются *множителями Лагранжа*. Это действие исследуется на безусловный экстремум. Вариация действия (20.27) по полям φ^a и множителям Лагранжа λ^A приводит к $N + M$ уравнениям Эйлера–Лагранжа, M из которых, возникших при вариации по множителям Лагранжа, совпадают с уравнениями связей (20.26). При этом вариации множителей Лагранжа на границе не обязаны быть равными нулю, т.к. они входят в действие без производных и никаких дополнительных граничных условий не возникает. Решение новой задачи на безусловный экстремум дает решение исходной задачи на условный экстремум, что является содержанием следующего утверждения.

Теорема 20.1.4. *Для функций φ , на которых функционал (20.1) имеет стационарное значение при выполнении уравнений связей (20.26), существует такой набор множителей Лагранжа, что они вместе с полями φ удовлетворяют уравнениям Эйлера–Лагранжа для действия (20.27):*

$$\frac{\delta S_T}{\delta \varphi^a} = 0, \quad \frac{\delta S_T}{\delta \lambda^A} = G_A = 0.$$

Доказательство. См., например, [134], глава 9, §1. □

Сформулированная теорема позволяет свести вариационную задачу на условный экстремум к вариационной задаче на безусловный экстремум, но для действия, зависящего от большего числа функций. В теории поля вариационные задачи рассматриваются не на прямой $\mathbb{R}$, а в евклидовом пространстве $\mathbb{R}^n$. Тогда связи представляют собой в общем случае дифференциальные уравнения в частных производных. Для того чтобы доказать аналог теоремы о множителях Лагранжа необходимо зафиксировать каким либо образом класс рассматриваемых связей, что является сложной задачей. На практике метод неопределенных множителей Лагранжа часто используют, не заботясь о его применимости. В таком случае применимость метода необходимо доказывать в каждом конкретном случае.

20.1.5 Другие задачи и терминология

В общем случае функционал действия может зависеть от частных производных полей φ^a любого порядка, вплоть до бесконечного:

$$S[\varphi] = \int_{\mathbb{M}} dx L(x, \varphi, \partial\varphi, \partial^2\varphi, \dots). \quad (20.28)$$

Тогда уравнения Эйлера–Лагранжа примут вид

$$S_{,a} = \frac{\partial L}{\partial(\varphi^a)} - \partial_\alpha \frac{\partial L}{\partial(\partial_\alpha \varphi^a)} + \partial_{\alpha\beta}^2 \frac{\partial L}{\partial(\partial_{\alpha\beta}^2 \varphi^a)} - \dots = 0. \quad (20.29)$$

Для простоты в настоящем разделе мы не будем обсуждать возможные граничные слагаемые.

Определение. Модели теории поля, лагранжиан которых зависит от производных бесконечного порядка, называются *нелокальными*. Для *локальных* моделей, порядок производных ограничен и ряд (20.29) обрывается. Модели, для которых уравнения Эйлера–Лагранжа содержат производные третьего или более высокого, но конечного порядка называются *моделями с высшими производными*. □

Хорошо известно, что порядок уравнений можно понизить, рассматривая частные производные в качестве новых независимых переменных. В этом смысле любую теорию с высшими производными можно свести к модели без высших производных.

Замечание. С физической точки зрения теории с высшими производными представляют определенные трудности, т.к. наличие векторных лоренцевых индексов у полей, как правило, приводит к каноническому гамильтониану, неограниченному снизу за счет вклада временных компонент. После квантования такие теории приводят к гильбертову пространству с индефинитной метрикой, которая не допускает вероятностной интерпретации квантовой теории. Эти трудности можно избежать за счет выбора лагранжиана специального вида или налагая условие калибровочной инвариантности, которое позволяет исключить вклад временных компонент в канонический гамильтониан для физических степеней свободы. В квантовой теории поля модели с высшими производными принято считать неудовлетворительными до тех пор, пока не доказана положительная определенность канонического гамильтониана для физических степеней свободы.

Нелокальные теории поля представляют собой еще большие трудности для физической интерпретации, т.к. помимо проблем с индефинитной метрикой гильбертова пространства, в общем случае они нарушают причинность. Это следует из того, что значение функции, разложимой в ряд Тейлора в точке x , в точке $y \neq x$ выражаются через ее значения и значения ее производных в точке x в виде бесконечного ряда. Следовательно, значение функции в некоторой точке может зависеть от ее значений в конусе будущего. $\square$

Допустим, что задана система уравнений Эйлера–Лагранжа. Функционал действия, приводящий к этим или эквивалентным уравнениям Эйлера–Лагранжа, определен неоднозначно. Во-первых, как уже отмечалось при рассмотрении вариационной задачи с заданными граничными условиями, уравнения не изменятся, если к лагранжиану добавить частную производную от функции, зависящей произвольным образом от полей и их частных производных. В частности, можно добавлять члены, имеющие вид дивергенции. Отметим, что если сам лагранжиан равен полной дивергенции, то он не приводит ни к каким уравнениям Эйлера–Лагранжа (получается тождество $0 = 0$). Во-вторых, вместо одного набора полей φ^a можно выбрать другой: $\varphi'^a = \varphi'^a(\varphi)$. Если это преобразование полей невырождено, то новая система уравнений Эйлера–Лагранжа будет эквивалентна старой. Пример дают канонические преобразования в гамильтоновом формализме. Иногда можно изменить даже число новых переменных в лагранжиане, и доказать биективность пространств возникающих решений уравнений Эйлера–Лагранжа. В частности, лагранжев и гамильтоновы способы описания динамики точечных частиц, рассмотренные в следующей главе, приводят к эквивалентным системам уравнений, но для разного числа переменных.

Замечание. Выбор независимых переменных в действии, по которым проводится варьирование, чрезвычайно важен, поскольку может привести к существенному упрощению возникающих уравнений движения, особенно в нелинейных теориях. С другой стороны, квантования моделей теории поля, основанные на различном выборе динамических переменных, может привести к различным квантовым теориям. В последнем случае теоретическим критерием выбора способа квантования является простота и самосогласованность конечной квантовой теории. Этот вопрос актуален для построения самосогласованной квантовой теории гравитации, которая в настоящее время отсутствует. $\square$

В релятивистских моделях математической физики, т.е. в моделях, инвариантных относительно преобразований из группы Пуанкаре, действие записывается в виде интеграла по всему пространству Минковского $\mathbb{R}^{1,3}$. В этом случае уравнения Эйлера–Лагранжа (20.11) называются также *уравнениями движения*, поскольку описывают эволюцию системы во времени. При этом для уравнений движения часто ставится не краевая задача, а задача Коши.

При рассмотрении моделей теории поля в пространстве Минковского $\mathbb{R}^{1,3}$ действие, как правило, расходится. Например, действие в электродинамике для электромагнитных волн расходится. Это связано с бесконечной областью интегрирования. Тем не менее с действием проводятся формальные выкладки, которые приводят к уравнениям Эйлера–Лагранжа, которые локальны и хорошо определены. При рассмотрении законов сохранения, связанных с первой теоремой Нетер, мы рассматриваем либо поля в конечном объеме, либо достаточно быстро убывающие на бесконечности.

Для корректной постановки вариационной задачи необходим глубокий анализ уравнений Эйлера–Лагранжа. Во многих важных случаях эти уравнения настолько сложны, что корректность постановки вариационной задачи доказать не удается. Поэтому в теоретической физике выбор действия означает, как правило, просто удобный способ задания уравнений движения для модели с заданными свойствами симметрии, что, конечно, чрезвычайно важно.

20.2 Первая теорема Нетер

В наиболее содержательных моделях математической физики функционал действия инвариантен относительно глобальных или локальных преобразований симметрии. С каждым преобразованием симметрии связан закон сохранения, что было установлено Эмми Нетер [135] в первой и второй теореме, соответственно, для глобальных и локальных преобразований.

Пусть функционал действия (20.1) инвариантен относительно бесконечно малых преобразований

$$x^\alpha \mapsto x'^\alpha = x^\alpha + \delta x^\alpha, \quad (20.30)$$

$$\varphi^a(x) \mapsto \varphi'^a(x') = \varphi^a(x) + \bar{\delta}\varphi^a(x). \quad (20.31)$$

Рассмотрим независимые вариации координат и полей:

$$\delta x^\alpha = \epsilon^A R_A^\alpha(x, \varphi, \partial\varphi), \quad (20.32)$$

$$\bar{\delta}\varphi^a = \varphi'^a(x') - \varphi^a(x) = \epsilon^A R_A^a(x, \varphi, \partial\varphi), \quad (20.33)$$

где $R_A^\alpha(x, \varphi, \partial\varphi)$ и $R_A^a(x, \varphi, \partial\varphi)$ – некоторые достаточно гладкие функции своих аргументов, которые называются *генераторами* преобразований симметрии, а $\epsilon^A(x)$, $A = 1, 2, \dots, K$, – постоянные или локальные параметры преобразований, число которых зависит от рассматриваемой модели. Мы говорим, что каждому значению индекса A соответствует одно преобразование симметрии.

Преобразования (20.30) и (20.31) уже рассматривались нами при обсуждении вариационной задачи с подвижной границей. Разница заключается в том, что сейчас у нас не один, а K параметров преобразования, которые, вдобавок, могут зависеть от точки $x \in \mathbb{M}$.

Начнем с доказательства первой теоремы Нетер, т.е. будем считать параметры преобразований постоянными, $\epsilon^A = \text{const}$. Под инвариантностью функционала действия мы понимаем следующее равенство

$$S = \int_{\mathbb{M}} dx L(x, \varphi, \partial\varphi) = \int_{\mathbb{M}'} dx' L(x', \varphi', \partial'\varphi'), \quad (20.34)$$

где интегрирование производится по ограниченной области $\mathbb{M} \subset \mathbb{R}^n$, которая отображается в $\mathbb{M}'$ при преобразовании (20.30).

Преобразования (20.32), (20.33) нетривиально действуют как на поля, так и на координаты. В дальнейшем нам понадобится вариация формы функции в данной точке $x \in \mathbb{M}$ (см., раздел 2.13):

$$\delta\varphi^a(x) := \varphi'^a(x) - \varphi^a(x),$$

которая определяется разностью значений полей после и до преобразования в точке x . Она связана с вариацией (20.31) следующим соотношением

$$\delta\varphi^a(x) = \bar{\delta}\varphi^a - \delta x^\alpha \partial_\alpha \varphi^a = \epsilon^A (R_A^a - R_A^\alpha \partial_\alpha \varphi^a). \quad (20.35)$$

По построению, для постоянных параметров преобразований вариация формы функции δ перестановочна с операцией частного дифференцирования ∂_α . Вариация действия относительно преобразований (20.30), (20.31) была вычислена ранее (20.23). Если выполнены уравнения Эйлера–Лагранжа, то вариацию действия при постоянных параметрах преобразований симметрии запишем в виде

$$\delta S = - \int_{\mathbb{M}} dx \epsilon^A \partial_\alpha J_A^\alpha, \quad (20.36)$$

где

$$J_A^\alpha := - \frac{\partial L}{\partial(\partial_\alpha \varphi^a)} (R_A^a - R_A^\beta \partial_\beta \varphi^a) - L R_A^\alpha. \quad (20.37)$$

Совокупность величин J_A^α , $\alpha = 1, \dots, n$ можно рассматривать, как компоненты некоторого вектора (точнее, векторной плотности) J_A , который называется *сохраняющимся током* для каждого преобразования симметрии с параметром ϵ^A . Из полученного выражения следует

Теорема 20.2.1 (Первая теорема Нетер). *Если действие (20.1) инвариантно относительно преобразований (20.30)–(20.33) с постоянными параметрами ϵ^A , то для каждого преобразования симметрии и любого решения уравнений Эйлера–Лагранжа токи сохраняются:*

$$\partial_\alpha J_A^\alpha = 0, \quad A = 1, \dots, K. \quad (20.38)$$

Заметим, что для сохранения тока достаточно глобальной инвариантности, когда параметр преобразования не зависит от точек пространства-времени.

Поскольку лагранжиан не содержит производных выше первого порядка, то компоненты токов в общем случае зависят только от координат, полей и их первых производных.

Закон сохранения (20.38) не нарушится, если к току (20.37) добавить слагаемое

$$J_A'^\alpha = J_A^\alpha + \partial_\beta f_A^{\beta\alpha}, \quad (20.39)$$

где $f_A^{\beta\alpha} = -f_A^{\alpha\beta}$ – произвольная антисимметричная по индексам α, β функция. Чтобы не менять структуры тока (20.37), будем считать, что она зависит только от координат x^α , полей φ и их первых производных $\partial\varphi$. Это преобразование часто используется, чтобы упростить выражения для токов.

Перепишем закон сохранения (20.38) в интегральной форме и используем формулу Стокса

$$\int_{\mathbb{M}} dx \partial_\alpha J_A^\alpha = \int_{\partial\mathbb{M}} ds_\alpha J_A^\alpha = 0,$$

где интегрирование ведется по многообразию $\mathbb{M}$ и его краю $\partial\mathbb{M}$. Пусть на $\mathbb{M}$ задана (псевдо)риманова геометрия, т.е. метрика $g_{\alpha\beta}$ и связность Леви–Чивиты $\tilde{\Gamma}_{\alpha\beta}^\gamma$. Тогда, если индекс A не преобразуется при преобразовании координат, то первый интеграл по (псевдо)риманову многообразию, можно переписать в ковариантной форме:

$$\int_{\mathbb{M}} dx \partial_\alpha J_A^\alpha = \int_{\mathbb{M}} dx \sqrt{|g|} \tilde{\nabla}_\alpha \left(\frac{1}{\sqrt{|g|}} J_A^\alpha \right),$$

где

$$\tilde{\nabla}_\alpha \left(\frac{1}{\sqrt{|g|}} J_A^\alpha \right) = \partial_\alpha \left(\frac{1}{\sqrt{|g|}} J_A^\alpha \right) + \tilde{\Gamma}_{\alpha\beta}^\alpha \frac{1}{\sqrt{|g|}} J_A^\beta$$

– ковариантная производная от вектора тока, и мы воспользовались формулой для дивергенции (6.63).

Рассмотрим действие (20.1) в пространстве Минковского $\mathbb{R}^{1,n-1}$. Пусть $\{x^\alpha\} = \{x^0, \mathbf{x}\}$ – декартова система координат, и все поля достаточно быстро убывают на пространственной бесконечности:

$$\lim_{|\mathbf{x}| \rightarrow \infty} \varphi^a = 0$$

для всех моментов времени. Тогда, интегрируя уравнение (20.38) по области пространства Минковского, которая ограничена двумя пространственноподобными сечениями $x_1^0 = \text{const}$ и $x_2^0 = \text{const}$, получим закон сохранения

$$Q_A = \int_{\mathbb{S}} d\mathbf{x} J_A^0 = \text{const}, \quad (20.40)$$

где $\mathbb{S}$ – произвольное сечение $x^0 = \text{const}$. Это означает, что каждому преобразованию симметрии соответствует закон сохранения: для любого решения уравнений движения, достаточно быстро убывающего на пространственной бесконечности, интеграл (20.40) не зависит от времени. Интеграл (20.40) называется *сохраняющимся зарядом*, соответствующим току J_A^α . Если для уравнений движения поставлена задача Коши, то значение заряда Q_A однозначно определяется начальными условиями.

20.2.1 Тензор энергии-импульса

Предположим, что некоторая модель описывается набором полей φ^a в пространстве Минковского $\mathbb{R}^{1,n-1}$ с декартовыми координатами x^α . При этом метрика $\eta_{\alpha\beta} = \text{diag}(+ \dots -)$ является заданной функцией в действии, по которой варьирование не проводится. Пусть действие инвариантно относительно трансляций

$$\delta x^\alpha = \epsilon^\alpha = \text{const}, \quad (20.41)$$

$$\bar{\delta} \varphi^a = 0. \quad (20.42)$$

Для этого достаточно, чтобы лагранжиан модели $L(\varphi, \partial\varphi)$ не зависел явно от координат. Для трансляций индекс A в (20.32) пробегает те же значения, что и α , генератор трансляций совпадает с символом Кронекера, $R_A^\alpha \mapsto \delta_\beta^\alpha$, и $R_A^a = 0$. В этом случае выражение для тока (20.37) имеет вид

$$T_\alpha{}^\beta = \partial_\alpha \varphi^a \frac{\partial L}{\partial(\partial_\beta \varphi^a)} - \delta_\alpha^\beta L. \quad (20.43)$$

Это выражение называется *тензором энергии-импульса* полей φ^a . В силу первой теоремы Нетер, он сохраняется:

$$\partial_\beta T_\alpha{}^\beta = 0. \quad (20.44)$$

Тензор энергии-импульса (20.43) будем называть *каноническим*.

Если лагранжиан модели является скалярным полем (функцией) относительно глобальных преобразований Лоренца $\mathbb{O}(1, n-1)$, то выражение (20.43) представляет собой тензор второго ранга типа $(1, 1)$. Ясно, что выражение для T_0^0 всегда совпадает с плотностью гамильтониана для полей φ^a , и это оправдывает название “канонический”.

Замечание. В общей теории относительности, основанной на псевдоримановой геометрии, постулируется, что тензор Эйнштейна пропорционален тензору энергии-импульса материи. При этом тензор энергии-импульса материи (23.11) определяется, как вариационная производная действия для полей материи по метрике. При таком определении тензор энергии-импульса всегда симметричен. Для скалярного поля вариационная производная действия по метрике является ковариантным обобщением тензора (20.43). В других случаях связь двух определений сложнее и будет обсуждаться в каждом конкретном случае. $\square$

Вообще говоря, тензор энергии-импульса с опущенным верхним индексом $T_{\alpha\beta}$ не является симметричным. Если это так, то в ряде случаев можно провести симметризацию, добавив соответствующую дивергенцию (20.39). Однако это не всегда возможно. Действительно, после добавления дивергенции получим новый тензор энергии-импульса

$$T'_{\alpha\beta} = T_{\alpha\beta} + \partial^\gamma f_{\alpha\gamma\beta}.$$

Из условия симметрии $T'_{\alpha\beta} - T'_{\beta\alpha} = 0$ следуют уравнения на неизвестную функцию $f_{[\alpha\gamma\beta]} = 0$, в которые входят только полностью антисимметричные компоненты,

$$\partial^\gamma f_{[\alpha\gamma\beta]} = -\frac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha}).$$

Таким образом, мы имеем $n(n-1)/2$ дифференциальных уравнений на $n(n-1)(n-2)/3!$ неизвестных компонент. При $n = 4$ возникает 6 уравнений на 4 неизвестные функции, которые не всегда имеют решения.

Введем стандартные 3-формы на координатных трехмерных гиперповерхностях в четырехмерном пространстве-времени $\mathbb{R}^{1,3}$:

$$ds_\alpha = \frac{1}{6} \varepsilon_{\alpha\beta\gamma\delta} dx^\beta \wedge dx^\gamma \wedge dx^\delta. \quad (20.45)$$

Определим сохраняющийся во времени *ковектор энергии-импульса* с помощью интеграла

$$P_\alpha := \int_{x^0 = \text{const}} ds_\beta T_\alpha{}^\beta, \quad (20.46)$$

где по индексу β производится суммирование. Полученное выражение (20.46), по построению, является ковектором относительно глобальных лоренцевых вращений. В предположении, что все поля достаточно быстро убывают на пространственной бесконечности ковектор энергии-импульса определяется одним интегралом по пространству,

$$P_\alpha = \int_{x^0 = \text{const}} d\mathbf{x} T_\alpha^0. \quad (20.47)$$

Выражение для нулевой компоненты P_0 совпадает с гамильтонианом системы полей φ^a , т.е. равно сохраняющейся полной энергии. Это оправдывает название ковектора энергии-импульса. Пространственные компоненты тензора энергии-импульса

$$T_i^0 = \frac{\partial L}{\partial(\partial_0 \varphi^a)} \partial_i \varphi^a, \quad i = 1, \dots, n-1, \quad (20.48)$$

определяют сохраняющийся полный импульс системы полей φ^a

$$P_i = \int_{x^0 = \text{const}} d\mathbf{x} T_i^0.$$

Полная энергия системы P_0 и каждая компонента полного импульса P_i относительно данной декартовой системы сохраняются во времени. В другой декартовой системе координат они тоже сохраняются, но имеют другие численные значения.

20.2.2 Тензор момента количества движения

Пусть действие $S[\varphi]$ в пространстве Минковского $\mathbb{R}^{1,n-1}$ инвариантно относительно лоренцевых вращений. Мы предполагаем, что набор полей φ^a преобразуется по некоторому, возможно, приводимому представлению группы Лоренца $\mathbb{SO}_0(1, n-1)$. Обозначим представление генераторов группы для полей через $L_{\gamma\delta b}^a = -L_{\delta\gamma b}^a$. Тогда в инфинитезимальной форме лоренцевы вращения примут вид

$$\delta x^\alpha = -x^\beta \omega_\beta^\alpha = \sum_{\gamma < \delta} \omega^{\gamma\delta} (x_\delta \delta_\gamma^\alpha - x_\gamma \delta_\delta^\alpha), \quad (20.49)$$

$$\bar{\delta} \varphi^a = \sum_{\gamma < \delta} \omega^{\gamma\delta} L_{\gamma\delta b}^a \varphi^b, \quad (20.50)$$

$\omega^{\gamma\delta} = -\omega^{\delta\gamma}$ — параметры преобразований, которые предполагаются постоянными. Для инвариантности действия достаточно, чтобы лагранжиан был скалярным полем (функцией) от координат, полей и их производных. Для лоренцевых вращений индекс $A \mapsto (\alpha\beta) = -(\beta\alpha)$ представляет собой пару антисимметричных векторных индексов.

Выражение для тока (20.37) приводит к следующему тензору *момента количества движения*, который мы представим в виде суммы двух слагаемых:

$$\begin{aligned} J_{\gamma\delta}^\alpha &:= -\frac{\partial L}{\partial(\partial_\alpha \varphi^a)} \left(L_{\gamma\delta b}^a \varphi^b - (x_\delta \delta_\gamma^\beta - x_\gamma \delta_\delta^\beta) \partial_\beta \varphi^a \right) - L(x_\delta \delta_\gamma^\alpha - x_\gamma \delta_\delta^\alpha) = \\ &= M_{\gamma\delta}^\alpha + S_{\gamma\delta}^\alpha, \end{aligned} \quad (20.51)$$

где введен *орбитальный* и *спиновый* моменты, соответственно,

$$M_{\alpha\beta}^\gamma := x_\beta T_\alpha^\gamma - x_\alpha T_\beta^\gamma, \quad (20.52)$$

$$S_{\alpha\beta}^\gamma := -\frac{\partial L}{\partial(\partial_\gamma \varphi^a)} L_{\alpha\beta b}^a \varphi^b. \quad (20.53)$$

Здесь T_α^β – канонический тензор энергии-импульса (20.43). Оба объекта являются тензорами третьего ранга относительно преобразований Лоренца. Обратим внимание, что орбитальный момент (20.52) не инвариантен относительно трансляций, т.к. явно зависит от координат. В противоположность этому спиновый момент инвариантен относительно трансляций.

Если все поля φ^a являются скалярами относительно лоренцевых вращений, то $L_{\alpha\beta b}^a = 0$ и спиновый момент равен нулю, $S_{\alpha\beta}^\gamma = 0$.

Допустим, что действие для некоторой системы полей инвариантно относительно трансляций и лоренцевых вращений (группы Пуанкаре), и спиновый момент равен нулю $S_{\alpha\beta}^\gamma = 0$, как для скалярных полей. Тогда закон сохранения момента количества движения принимает вид

$$\partial_\gamma M_{\alpha\beta}^\gamma = T_{\alpha\beta} - T_{\beta\alpha} + x_\beta \partial_\gamma T_\alpha^\gamma - x_\alpha \partial_\gamma T_\beta^\gamma.$$

С учетом закона сохранения тензора энергии-импульса (20.44) отсюда вытекает, что для такой системы ковариантный тензор энергии-импульса симметричен:

$$T_{\alpha\beta} = T_{\beta\alpha}. \quad (20.54)$$

Так же, как и для канонического тензора энергии-импульса, для тензора момента количества движения можно ввести полный момент системы. Для полей, достаточно быстро убывающих на пространственной бесконечности, он равен интегралу по пространству:

$$J_{\alpha\beta} := \int_{x^0 = \text{const}} d\mathbf{x} J_{\alpha\beta}^0. \quad (20.55)$$

Полный момент количества движения является антисимметричным тензором второго ранга относительно преобразований Лоренца.

Замечание. Требование инвариантности моделей математической физики относительно преобразований группы Пуанкаре имеет глубокий физический смысл и составляет основное содержание специальной теории относительности. Инвариантность действия относительно трансляций означает однородность пространства-времени. То есть все точки пространства-времени равноправны, и законы природы имеют одинаковый вид в декартовых координатах с произвольно выбранным началом. Инвариантность относительно преобразований Лоренца означает изотропность пространства-времени. То есть равноправие всех направлений и одинаковый вид законов природы в декартовых координатах с произвольной ориентацией осей. Законы сохранения энергии-импульса и момента количества движения к настоящему времени нашли многочисленные экспериментальные подтверждения в различных областях физики. Поэтому инвариантность фундаментальных моделей математической физики относительно действия группы Пуанкаре следует считать экспериментально установленным фактом.

Помимо этого требование инвариантности функционала действия относительно преобразований группы Пуанкаре в квантовой теории поля означает, что все элементарные частицы должны описываться полями, принадлежащими одному из неприводимых представлений группы Пуанкаре, которые характеризуются массой и спином. Использование этих понятий в экспериментальной физике элементарных частиц также чрезвычайно плодотворно. Это также можно рассматривать, как экспериментальное подтверждение инвариантности законов природы относительно преобразований группы Пуанкаре. $\square$

20.3 Вторая теорема Нетер

Рассмотрим действие (20.1), которое инвариантно относительно преобразований (20.30)–(20.33) с локальными параметрами $\epsilon^A(x)$, зависящим от точек пространства-времени. Мы допускаем, что эти преобразования могут зависеть от частных производных $\partial_\alpha \epsilon^A$ первого и более высокого порядка. Чтобы упростить формулы будем использовать обозначения Девитта [136], т.е. суммирование по индексу A в формулах (20.30), (20.31) подразумевает интегрирование, а генераторы локальных преобразований рассматриваются, как двухточечные функции, содержащие δ -функции и (или) их производные.

Пример 20.3.1. Калибровочное преобразование в электродинамике

$$\delta A_\alpha = \partial_\alpha \epsilon$$

будем записывать в виде

$$\delta A_\alpha = \epsilon R_\alpha = \partial_\alpha \int dx' \epsilon(x') \delta(x' - x), \quad (20.56)$$

где

$$R_\alpha(x', x) := \frac{\partial}{\partial x^\alpha} \delta(x' - x). \quad \square \quad (20.57)$$

Пример 20.3.2. Бесконечно малые общие преобразования координат для электромагнитного поля (2.101) можно записать в виде

$$\delta x^\alpha = \epsilon^\alpha = \epsilon^\beta R_\beta{}^\alpha, \quad (20.58)$$

$$\delta A_\alpha = -\partial_\alpha \epsilon^\beta A_\beta - \epsilon^\beta \partial_\beta A_\alpha = \epsilon^\beta N_{\beta\alpha}, \quad (20.59)$$

где

$$R_\beta{}^\alpha := \delta_\beta^\alpha \delta(x' - x), \quad (20.60)$$

$$N_{\beta\alpha} := F_{\alpha\beta}(x') \delta(x' - x) - A_\beta(x') \frac{\partial}{\partial x^\alpha} \delta(x' - x). \quad \square \quad (20.61)$$

Определение. Преобразования полей (20.33) с локальными параметрами $\epsilon^A(x)$ называются *калибровочными*. $\square$

Рассмотрим одну из вариационных задач. Будем считать, что параметры ϵ^A и их производные равны нулю на границе области. Тогда инвариантность действия относительно калибровочных преобразований можно записать в виде

$$\delta S = \int dx \delta \varphi^a S_{,a} = \int dx \epsilon^A (R_A{}^a - R_A{}^\alpha \partial_\alpha \varphi^a) S_{,a} = 0. \quad (20.62)$$

При этом были отброшены все граничные слагаемые. Отсюда следует

Теорема 20.3.1 (Вторая теорема Нетер). *Если функционал действия (20.1) инвариантен относительно калибровочных преобразований, которые параметризуются K произвольными функциями $\epsilon^A(x)$, $A = 1, \dots, K$, то уравнения Эйлера–Лагранжа удовлетворяют K тождествам:*

$$(R_A{}^a - R_A{}^\alpha \partial_\alpha \varphi^a) S_{,a} = 0. \quad (20.63)$$

Замечание. В формулировке теоремы мы отбросили предположение о том, что параметры преобразований и их производные равны нулю на границе. Если это не так, то зависимость уравнений Эйлера–Лагранжа все равно сохранится. В этом случае из требования инвариантности действия появятся дополнительные следствия для граничных условий, которые мы не рассматриваем. $\square$

Напомним, что в линейном соотношении между уравнениями движения (20.63) суммирование по индексу a предполагает интегрирование. Отсюда следует, что если калибровочные преобразования зависят от частных производных l -того порядка от параметра преобразования, то соотношения (20.63) представляют собой систему K линейных дифференциальных уравнений в частных производных l -того порядка относительно вариационных производных $S_{,a}$.

Вторая теорема Нетер утверждает, что в калибровочных моделях, а также моделях, инвариантных относительно общих преобразований координат, не все уравнения движения являются линейно независимыми. Это указывает на то, что в решениях задачи Коши будет содержаться функциональный произвол, т.к. количества уравнений недостаточно для однозначного определения решений по начальным данным.

Для доказательства теоремы существенно, что параметры преобразований $\epsilon^\alpha(x)$ является произвольными функциями, т.к. только в этом случае подынтегральное выражение в (20.62) согласно основной лемме вариационного исчисления должно обращаться в нуль.

Пример 20.3.3. Проведем аналогию с теорией функций многих переменных. Пусть $f = f(x)$ – функция n переменных $x = \{x^\alpha\}$. Аналогом вариационной производной действия в таком случае является обычная частная производная $\partial_\alpha f$. Допустим, что f инвариантна относительно калибровочных преобразований $\delta x^\alpha = \epsilon X^\alpha$, где $\epsilon = \epsilon(x)$ – параметр преобразования и X^α – векторное поле (генератор калибровочного преобразования), которое предполагается отличным от нуля. Тогда “зависимость уравнений движения” сводится к линейной зависимости частных производных $X^\alpha \partial_\alpha f = 0$. Поэтому функция f постоянна вдоль интегральной кривой $x(t)$ векторного поля X^α :

$$f(x(t)) = \text{const}, \quad \dot{x}^\alpha = X^\alpha.$$

Это значит, что локальный экстремум $\partial_\alpha f = 0$ достигается не в точке, а на интегральной кривой $x(t)$. $\square$

Из второй теоремы Нетер следует, что функционал действия для калибровочных моделей достигает экстремального значения не на отдельных функциях, а на классах функций, которые связанных между собой калибровочными преобразованиями.

Если некоторая модель инвариантна относительно локальных преобразований, то она, в частности, инвариантна относительно тех же преобразований с постоянными параметрами. Это значит, что токи (20.37) приводят к законам сохранения (20.40) и для локальных преобразований. Поэтому в моделях, инвариантных относительно локальных преобразований можно применить обе теоремы Нетер. При этом первая теорема дает выражения для сохраняющихся токов, а вторая – зависимость уравнений движения. В общем случае это не одно и то же.

Пример 20.3.4. Рассмотрим модели математической физики, инвариантные относительно общих преобразований координат. Пусть действие $S = S(g, \Gamma)$ зависит только от метрики $g_{\alpha\beta}$ и аффинной связности $\Gamma_{\alpha\beta}^\gamma$. Обозначим вариационные производные действия следующим образом:

$$\sqrt{|g|} S_{, \alpha\beta} := \frac{\delta S}{\delta g_{\alpha\beta}}, \quad \sqrt{|g|} S_{, \alpha\beta}{}^\gamma := \frac{\delta S}{\delta \Gamma_{\alpha\beta}^\gamma}. \quad (20.64)$$

Здесь мы явно ввели в качестве множителя определитель репера $\sqrt{|g|} = \det e_\alpha^a$, поскольку вариационные производные так же, как и лагранжиан, являются тензорными плотностями степени -1 . Инвариантность действия относительно общих преобразований координат означает равенство нулю вариации

$$\delta S = \int dx \sqrt{|g|} (S,^{\alpha\beta} \delta g_{\alpha\beta} + S,^{\alpha\beta}{}_\gamma \delta \Gamma_{\alpha\beta}^\gamma) = 0.$$

Подставляя сюда вариации метрики (2.104) и связности (2.107) и интегрируя по частям (6.71), получим тождества

$$\begin{aligned} & 2\tilde{\nabla}_\alpha S,^\alpha{}_\gamma - \nabla_\beta \nabla_\alpha S,^\alpha\beta{}_\gamma + \\ & + \nabla_\beta (T_\alpha + \tfrac{1}{2}Q_\alpha) S,^\alpha\beta{}_\gamma + (T_\alpha + \tfrac{1}{2}Q_\alpha) \nabla_\beta (S,^\alpha\beta{}_\gamma + S,^\beta\alpha{}_\gamma) - \nabla_\alpha S,^\alpha\beta{}_\delta T_{\beta\gamma}^\delta - \\ & - (T_\alpha + \tfrac{1}{2}Q_\alpha) (T_\beta + \tfrac{1}{2}Q_\beta) S,^\alpha\beta{}_\gamma + (T_\alpha + \tfrac{1}{2}Q_\alpha) T_{\beta\gamma}^\delta S,^\alpha\beta{}_\delta + S,^\alpha\beta{}_\delta R_{\alpha\gamma}^\delta = 0, \end{aligned} \quad (20.65)$$

где $\tilde{\nabla}_\alpha$ и ∇_α — ковариантные производные, соответственно, со связностью Леви-Чивиты $\tilde{\Gamma}_{\alpha\beta}^\gamma$ и аффинной связностью $\Gamma_{\alpha\beta}^\gamma$, а подъем и опускание индексов производится с помощью метрики $g_{\alpha\beta}$. Таким образом, в моделях, инвариантных относительно общих преобразований координат, уравнения движения удовлетворяют n линейным дифференциальным тождествам.

В (псевдо)римановой геометрии, когда гравитационная часть действия зависит только от метрики, эти тождества значительно упрощаются:

$$\tilde{\nabla}_\alpha S,^\alpha{}_\beta = 0. \quad (20.66)$$

В общей теории относительности для действия Гильберта–Эйнштейна справедливо равенство

$$\tilde{\nabla}_\alpha \left(\tilde{R}^\alpha{}_\beta - \tfrac{1}{2} \delta^\alpha_\beta \tilde{R} \right) = 0. \quad (20.67)$$

Это тождество совпадает со свернутыми тождествами Бианки (6.121). $\square$

Пример 20.3.5. Модели гравитации, построенные в рамках геометрии Римана–Картана, в переменных Картана инвариантны относительно общих преобразований координат и локальных преобразований Лоренца. Как следствие второй теоремы Нетер не все уравнения движения являются независимыми, поскольку удовлетворяют тождествам. Пусть действие $S = S(e, \omega)$ зависит только от репера e_α^a и лоренцевой связности ω_α^{ab} . Обозначим вариационные производные следующим образом:

$$\sqrt{|g|} S,^\alpha{}_a := \frac{\delta S}{\delta e_\alpha^a}, \quad \sqrt{|g|} S,^\alpha{}_{ab} := \frac{\delta S}{\delta \omega_\alpha^{ab}}. \quad (20.68)$$

Бесконечно малые преобразования Лоренца для репера и лоренцевой связности при локальных лоренцевых вращениях имеют вид

$$\delta e_\alpha^a = -e_\alpha^b \omega_b^a, \quad (20.69)$$

$$\delta \omega_{\alpha a}^b = \omega_a^c \omega_{\alpha c}^b - \omega_{\alpha a}^c \omega_c^b + \partial_\alpha \omega_a^b. \quad (20.70)$$

Отсюда вытекает следующая зависимость уравнений движения

$$\tilde{\nabla}_\alpha S,^\alpha{}_{ab} + \tfrac{1}{2} (S,_{ab} - S,_{ba}) = 0, \quad (20.71)$$

где переход от греческих индексов к латинским осуществляется с помощью репера. Полученная зависимость соответствует инвариантности действия относительно локальных лоренцевых вращений.

Вариации полей ϵ и ω при общих преобразованиях координат имеют вид (2.113) и (2.114). Поэтому инвариантность действия относительно общих преобразований координат приводит к тождествам:

$$\nabla_\alpha S,^\alpha_\beta + T_\alpha S,^\alpha_\beta + S,^\alpha_a T_{\alpha\beta}^a + S,^\alpha_{ab} R_{\alpha\beta}^{ab} = 0, \quad (20.72)$$

где мы учли полученное ранее тождество (20.71).

При добавлении к гравитационному действию слагаемых, зависящих от других полей (полей материи), тождества, которым удовлетворяют уравнения движения меняются, т.к. необходимо учитывать вариации всех полей. $\square$

20.4 Эффективное действие

При исследовании моделей математической физики, действие которых зависит от нескольких полей, иногда удается решить часть уравнений Эйлера–Лагранжа явно, выразив одни переменные через другие в общем виде. В этом случае вариационную задачу можно свести к новому эффективному действию, зависящему от меньшего числа переменных. В настоящем разделе мы докажем простую теорему, позволяющую строить эффективное действие в случае вариационной задачи с фиксированными граничными условиями. То есть будем пренебрегать всеми граничными слагаемыми. Обобщение на более сложные случаи будет ясно из дальнейшего рассмотрения.

Начнем с простейшего случая. Пусть на ограниченной области $\mathbb{M} \subset \mathbb{R}^n$ заданы два скалярных поля φ и ψ . Предположим, что функция $\psi = \psi(x, \varphi, \partial\varphi, \partial^2\varphi)$ в каждой точке $x \in \mathbb{M}$ задана как функция φ , ее первых и вторых частных производных: $\partial_\alpha\varphi$ и $\partial_\alpha\partial_\beta\varphi$. Представим значение функции $\psi(x) := \psi[x, \varphi(x), \partial\varphi(x), \partial^2\varphi(x)]$ в точке $x \in \mathbb{M}$ в виде функционала, используя δ -функцию,

$$\psi(x) = \int_{\mathbb{M}} dy \psi(y) \delta(y - x).$$

Вариация функционала $\psi(x)$, вызванная вариацией $\delta\varphi$, имеет вид

$$\delta\psi(x) = \int_{\mathbb{M}} dy \left(\frac{\partial\psi}{\partial\varphi} \delta\varphi + \frac{\partial\psi}{\partial(\partial_\alpha\varphi)} \delta(\partial_\alpha\varphi) + \frac{\partial\psi}{\partial(\partial_\alpha\partial_\beta\varphi)} \delta(\partial_\alpha\partial_\beta\varphi) \right) \delta(y - x).$$

Проинтегрировав второе и третье слагаемые по частям, получим выражение для вариационной производной

$$\frac{\delta\psi(x)}{\delta\varphi(y)} = \frac{\partial\psi}{\partial\varphi} \delta(y - x) - \frac{\partial}{\partial y^\alpha} \left(\frac{\partial\psi}{\partial(\partial_\alpha\varphi)} \delta(y - x) \right) + \partial_\alpha\partial_\beta \left(\frac{\partial\psi}{\partial(\partial_\alpha\partial_\beta\varphi)} \delta(y - x) \right), \quad (20.73)$$

где в правой части $\psi = \psi(y)$ и $\varphi = \varphi(y)$.

Теперь обсудим вариационную задачу. Пусть действие $S[\varphi, \psi]$ зависит от двух функций φ и ψ . Тогда из принципа наименьшего действия следуют два уравнения Эйлера–Лагранжа:

$$\frac{\delta S}{\delta\varphi} = 0, \quad (20.74)$$

$$\frac{\delta S}{\delta\psi} = 0. \quad (20.75)$$

Допустим, что второе уравнение Эйлера–Лагранжа допускает общее решение для ψ , как функции от φ и ее производных:

$$\psi = \psi(x, \varphi, \partial\varphi, \partial^2\varphi). \quad (20.76)$$

При этом мы предполагаем, что общее решение не имеет особенностей. Если действие зависит только от самих функций и их первых производных, то в общее решение будут входить производные от φ не выше второго порядка. Поскольку уравнение Эйлера–Лагранжа (20.75) является дифференциальным уравнением в частных производных, то общее решение зависит также от некоторого набора произвольных функций и постоянных. Часть этих произвольных функций и постоянных фиксируется, если это возможно, граничными условиями $\psi|_{\partial M} = \psi_0$ и $\varphi|_{\partial M} = \varphi_0$. Используем полученное решение для построения нового *эффективного* действия

$$S_{\text{eff}}[\varphi] := S[\varphi, \psi(\varphi)], \quad (20.77)$$

которое зависит только от одной функции φ . Тогда уравнение Эйлера–Лагранжа для φ связано со старыми уравнениями (20.74), (20.75) простым соотношением

$$\frac{\delta S_{\text{eff}}}{\delta\varphi(x)} = \frac{\delta S}{\delta\varphi(x)} + \left. \frac{\delta S}{\delta\psi(y)} \frac{\delta\psi(y)}{\delta\varphi(x)} \right|_{\psi=\psi(\varphi)} = 0, \quad (20.78)$$

где во втором слагаемом подразумевается интегрирование по аргументу поля $\psi(y)$, которое снимается δ -функцией в вариационной производной. Ясно, что второе слагаемое равно нулю, если выполнено уравнение Эйлера–Лагранжа для ψ (20.75).

Проведенные вычисления остаются в силе и в том случае, когда мы рассматриваем наборы полей $\varphi = \{\varphi^a\}$, $a = 1, \dots, N$ и ψ^A , $A = 1, \dots, m$. Отсюда следует

Теорема 20.4.1. *Пусть дано действие $S[\varphi, \psi]$, зависящее от двух наборов полей φ и ψ . Тогда множество решений уравнений Эйлера–Лагранжа для вариационной задачи с заданными граничными условиями совпадает с множеством решений уравнений Эйлера–Лагранжа для эффективного действия (20.77), дополненным выражением ψ через φ (20.76).*

Эта теорема важна, поскольку позволяет строить эффективное действие, которое зависит от меньшего числа полей, подставляя решение части системы уравнений Эйлера–Лагранжа непосредственно в исходное действие.

При доказательстве теоремы 20.4.1 мы предположили, что всеми граничными слагаемыми можно пренебречь. В полевых моделях математической физики это не всегда так. В разделе 22.9 будет построен пример, где подстановка решения части уравнений Эйлера–Лагранжа в действие не воспроизводит оставшиеся уравнения движения. Это связано с нетривиальной ролью граничных слагаемых в действии для полевых моделей.

20.5 Редуцированное действие

В настоящем разделе мы рассмотрим еще один способ сведения сложной вариационной задачи к более простой. Пусть задано действие $S[\varphi]$. Рассмотрим для него вариационную задачу с фиксированными граничными условиями. Как правило, уравнения

Эйлера–Лагранжа (уравнения движения) настолько сложны, что не позволяют найти все решения. В таких случаях для нахождения частных решений делают упрощающие предположения: решение уравнений движения ищется в определенном классе функций. Например, ищутся статические или сферически симметричные решения. Чтобы найти стационарную точку действия, упрощающую подстановку, которую часто называют *анзац* (от немецкого *ansatz* $\simeq$ исходное математическое выражение), следует производить в уравнения Эйлера–Лагранжа (20.11), а не в действие. В этом случае найденное точное решение уравнений движения действительно будет стационарной точкой исходного действия.

Однако в ряде случаев подстановки можно производить непосредственно в действие. Это означает следующее. Пусть в результате некоторых упрощающих предположений исходный набор функций φ^a , $a = 1, \dots, N$, будет выражен через меньшее число независимых функций ψ^A , $A = 1, \dots, m$, и координаты:

$$\varphi^a = \varphi^a(x, \psi). \quad (20.79)$$

При этом функции ψ могут зависеть от меньшего числа координат. В результате подстановки будет получено новое *редуцированное действие*

$$S_{\text{red}}[\psi] := S[\varphi(\psi)], \quad (20.80)$$

зависящее от меньшего числа независимых полей.

Допустим, что найдено решение уравнений движения для редуцированного действия

$$\frac{\delta S_{\text{red}}}{\delta \psi^A} = 0.$$

Тогда функции (20.79), как правило, не будут удовлетворять исходным уравнениям (20.11). Тем не менее в ряде случаев исходные уравнения все же будут удовлетворены. Это замечательные случаи, которые позволяют существенно упростить вычисления. Вместе с этим наличие редуцированного действия помогает в анализе свойств рассматриваемой модели.

Опишем достаточные условия для возможности подстановки (20.79) непосредственно в действие. Пусть на многообразии $\mathbb{M}$ действует группа Ли преобразований $\mathbb{G}$ справа:

$$\mathbb{M} \times \mathbb{G} \ni \quad x, g \mapsto xg \quad \in \mathbb{M}.$$

Предположим, что набор полей $\varphi^a(x)$ при этом преобразуются по некоторому представлению $T(g)_a{}^b$ группы Ли $\mathbb{G}$:

$$\varphi'^a(xg) = \varphi^b(x)T(g)_b{}^a, \quad (20.81)$$

где штрихом обозначены новые поля на $\mathbb{M}$, полученные в результате действия группы преобразований $\mathbb{G}$. Поля $\varphi^a(x)$ называются $\mathbb{G}$ -инвариантными, если в равенстве (20.81) можно убрать штрих в левой части. То есть выполнено условие

$$\varphi^a(xg) = \varphi^b(x)T(g)_b{}^a,$$

или

$$\varphi^a(x) = \varphi^b(xg^{-1})T(g)_b{}^a. \quad (20.82)$$

Эти условия представляют собой уравнения, определяющие $\mathbb{G}$ -инвариантные поля на многообразии $\mathbb{M}$.

Теорема 20.5.1 (Принцип Коулмена). Пусть исходный функционал действия $S[\varphi]$ инвариантен относительно действия группы Ли преобразований (20.81),

$$S[\varphi'] = S[\varphi].$$

Допустим, что множество всех $\mathbb{G}$ -инвариантных полей на M параметризуется некоторым набором функций ψ^A , которые могут зависеть от меньшего числа координат и не определяются из уравнений (20.82). В результате $\mathbb{G}$ -инвариантные функции будут представлены в виде (20.79). Тогда поля (20.79), построенные для стационарных точек редуцированного действия (20.80), будут удовлетворять уравнениям Эйлера–Лагранжа исходного действия.

Это утверждение известно, как *принцип Коулмена*. Оно было высказано Коулменом и проиллюстрировано на нескольких примерах [137] (см. также [138]). Строгое доказательство вместе с ограничениями на его применимость было дано в статье [139]. Этот результат затем был обобщен в работах [140, 141]. Другая его формулировка, на языке теории стратов была дана еще до Коулмена [142]. В настоящее время доказано, что принцип Коулмана справедлив для всех компактных групп преобразований, полупростых групп, а также для унитарных представлений некомпактных групп.

Глава 21

Канонический формализм в механике точечных частиц

Трудно переоценить роль канонического (гамильтонова) формализма в классической и квантовой механике, а также в теории поля. Он предоставляет наиболее мощные методы интегрирования уравнений движения и является основой для канонического квантования различных моделей математической физики. К его недостаткам относится явное нарушение Лоренц-инвариантности моделей теории поля, поскольку время в гамильтоновом формализме играет выделенную роль. Это усложняет вычисления, проводимые в рамках теории возмущений. Однако принципиальные вопросы, связанные с физической интерпретацией математических моделей, невозможно решить без обращения к гамильтоновой формулировке. В настоящей главе рассматривается канонический формализм для системы точечных частиц, формализм Дирака для систем со связями и его обобщение на теорию поля.

В настоящей главе мы во многом следуем монографии [143]

21.1 Преобразование Лежандра

Рассмотрим *выпуклую функцию* $y = f(x)$ на интервале $-\infty \leq a < x < b \leq \infty$, т.е. функцию, у которой $f''(x) > 0$ при всех $x \in (a, b)$. Преобразованием Лежандра функции f на интервале (a, b) называется новая функция $g(p)$ нового переменного p , которая строится следующим образом (см. рис. 21.1). Нарисуем на плоскости x, y график функции f . Рассмотрим прямую $y = px$, где p – фиксированное чис-

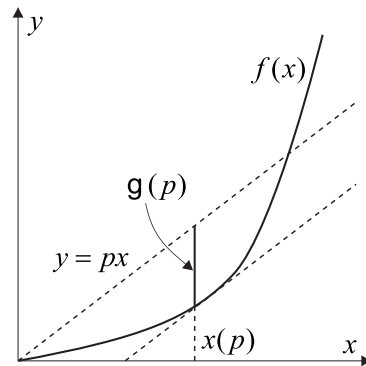


Рис. 21.1: Преобразование Лежандра $g(p)$ выпуклой функции $f(x)$.

ло. Найдем точку $x(p)$, в которой кривая дальше всего от прямой по вертикали, т.е.

функция $F(x, p) := px - f(x)$ имеет максимум по x при фиксированном p . Тогда, по определению, $g(p) := F(x(p), p)$. Функция $g(p)$ определена на некотором интервале $-\infty \leq c < p < d \leq \infty$.

Точка $x(p)$ определяется из условия экстремума $\partial F / \partial x = 0$ или

$$p = f'(x)$$

Ввиду выпуклости f , если это уравнение имеет решение, то оно единственно.

Пример 21.1.1. Нетрудно проверить, что функция $f(x) = \frac{mx^2}{2}$, $m = \text{const} > 0$ выпукла на всей вещественной оси $x \in \mathbb{R}$ и $p = mx$. Ее преобразование Лежандра определено для всех $p \in \mathbb{R}$ и имеет вид

$$g(p) = (xp - f)|_{x=p/m} = \frac{p^2}{2m} \quad \square$$

Пусть функция f достаточно гладкая. Тогда преобразование Лежандра переводит выпуклые функции в выпуклые. Это значит, что преобразование Лежандра можно применить дважды. Можно доказать [143], что преобразование Лежандра *инволютивно*, т.е. его квадрат равен тождественному преобразованию.

По определению, $F(x, p) := px - f(x) \leq g(p)$ для всех x и p . Отсюда вытекает *неравенство Юнга*

$$px \leq f(x) + g(p)$$

Преобразование Лежандра без труда обобщается на функции нескольких переменных. Пусть $f(\mathbf{x})$ – выпуклая функция нескольких переменных $\mathbf{x} = (x^1, \dots, x^n) \in \mathbb{R}^n$, т.е. квадратичная форма $dx^\alpha dx^\beta \partial_{\alpha\beta}^2 f$, $\alpha = 1, \dots, n$, положительно определена в некоторой окрестности $\mathbb{U} \subset \mathbb{R}^n$. Тогда преобразованием Лежандра называется функция $g(\mathbf{p})$ того же числа переменных $\mathbf{p} = (p_1, \dots, p_n)$, которая строится аналогично случаю одного переменного:

$$g(\mathbf{p}) = F(\mathbf{x}(\mathbf{p}), \mathbf{p}) = \max_{\mathbf{x}} F(\mathbf{x}, \mathbf{p}),$$

где

$$F(\mathbf{x}, \mathbf{p}) = x^\alpha p_\alpha - f(\mathbf{x}), \quad \text{и} \quad p_\alpha = \frac{\partial f}{\partial x^\alpha}.$$

Преобразование Лежандра $g(\mathbf{p})$ определено в некоторой окрестности $\mathbf{p} \in \mathbb{V} \subset \mathbb{R}^n$, которая определяется исходной функцией $f(\mathbf{x})$.

В этом определении мы различаем верхние и нижние индексы по следующей причине. Поскольку $\mathbf{x}$ – точка многообразия $\mathbb{R}^n$, то индексы ее координат мы пишем сверху. По определению, f – функция на $\mathbb{R}^n$. Поэтому набор переменных p_α определяет компоненты некоторого ковектора (1-формы). Отметим также, что в этом определении все координаты x^α равноправны.

Пример 21.1.2. Преобразованием Лежандра положительно определенной квадратичной формы

$$f(\mathbf{x}) = \frac{1}{2} g_{\alpha\beta} x^\alpha x^\beta,$$

где $g_{\alpha\beta} = g_{\beta\alpha}$ – постоянная матрица, снова является положительно определенной квадратичная форма

$$g = (x^\alpha p_\alpha - f(\mathbf{x}))|_{x^\alpha = g^{\alpha\beta} p_\beta} = \frac{1}{2} g^{\alpha\beta} p_\alpha p_\beta,$$

где $g^{\alpha\beta}$ – матрица, обратная к $g_{\alpha\beta}$. При этом значения обеих форм в соответствующих точках совпадают:

$$f(\mathbf{x}(\mathbf{p})) = g(\mathbf{p}), \quad g(\mathbf{p}(\mathbf{x})) = f(\mathbf{x}).$$

Обе функции $f(\mathbf{x})$ и $g(\mathbf{p})$ определены на всем евклидовом пространстве $\mathbb{R}^n$. $\square$

21.2 Гамильтонова динамика точечных частиц

Механика точечных частиц может быть описана на двух эквивалентных языках: лагранжевом и гамильтоновом. Каждый подход имеет свои преимущества и недостатки. В лагранжевом подходе совокупность N частиц описывается обобщенными координатами $q^i(t)$, $i = 1, \dots, N$, зависящих от времени t . Если каждая частица движется в трехмерном пространстве, то q^i представляет собой трехмерный вектор в евклидовом пространстве $\mathbb{R}^3$ для каждого значения индекса i . Для определенности будем считать, что каждая частица движется в одномерном пространстве. Тогда совокупность всех координат q^i можно также рассматривать, как координаты одной частицы в *конфигурационном* пространстве $\mathbb{R}^N$. Размерность конфигурационного пространства называется *числом степеней свободы* механической системы. Говорят, что механическая система имеет N степеней свободы.

Для простоты обозначений условимся считать, что символ с индексом q^i обозначает i -ю координату частицы, а по повторяющимся индексам производится суммирование. Если индекс отсутствует, то символ q обозначает весь набор координат $q := (q^1, \dots, q^N)$. В лагранжевом подходе уравнения движения имеют второй порядок, поэтому будем рассматривать дважды непрерывно дифференцируемые функции $q^i \in \mathcal{C}^2([t_1, t_2])$ для всех i на конечном отрезке $[t_1, t_2]$.

Предположим, что механическая система описывается некоторым действием

$$S[q] = \int_{t_1}^{t_2} dt L(q, \dot{q}, t), \quad (21.1)$$

где функция Лагранжа или лагранжиан $L(q, \dot{q}, t)$ зависит только от обобщенных координат q и их первых производных по времени $\dot{q}$, которые называются обобщенными *скоростями*. Рассмотрим вариационную задачу с фиксированными граничными условиями

$$q(t_1) = q_1, \quad q(t_2) = q_2, \quad (21.2)$$

т.е. траектория механической системы представляет собой кривую в конфигурационном пространстве, которая соединяет две фиксированные точки q_1 и q_2 . Тем самым вариации координат на концах интервала обращаются в нуль.

Обычно предполагают, что конфигурационное пространство представляет собой евклидово пространство $\mathbb{R}^N$ с заданной метрикой δ_{ij} , а обобщенные координаты – это декартовы координаты в $\mathbb{R}^N$. Кроме того, мы предполагаем, что лагранжиан в действии (21.1) представляет собой функцию (скалярное поле) от своих аргументов. Поэтому метрика необходима для построения инвариантов из координат q^i и скоростей $\dot{q}^i$, и она входит в действие. В этом действии евклидова метрика рассматривается как внешнее поле, и по ней варьирование не проводится. Конечно, после того, как задача поставлена, действие можно переписать в произвольной криволинейной системе координат в $\mathbb{R}^N$. Тогда обобщенные координаты q^i станут криволинейными координатами в $\mathbb{R}^N$, а компоненты евклидовой метрики $g_{ij}(x)$ будут нетривиальными функциями от x . Действие по этим компонентам не варьируется.

Замечание. При рассмотрении движения механической системы, естественная линейная структура в $\mathbb{R}^N$ не играет никакой роли и нигде не используется. Важно только наличие метрики. В общем случае можно считать, что конфигурационное пространство – это произвольное риманово многообразие (M, g) с заданной метрикой g , которая необходима для построения инвариантов. Действие в этом случае будет зависеть от метрики, которая рассматривается как внешнее поле, и по ней варьирование не проводится. Заметим также, что задание лагранжиана ничего не говорит о топологии конфигурационного пространства. Для того, чтобы задать топологию M необходимо сделать какие либо дополнительные предположения. В дальнейшем мы будем рассматривать в основном топологически тривиальные конфигурационные пространства, $M \approx \mathbb{R}^N$ с естественной топологией. $\square$

Уравнения Эйлера–Лагранжа (уравнения движения) для действия (21.1),

$$\frac{\delta S}{\delta q^i} = \frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = 0, \quad (21.3)$$

представляют собой систему уравнений не выше второго порядка. Функции $q^i(t)$, удовлетворяющие уравнениям Эйлера–Лагранжа, соответствуют стационарным точкам действия (21.1). Они определяют *траекторию* механической системы. Траектории частиц не зависят от выбора координат в конфигурационном пространстве $\mathbb{R}^N$, которые выбираются из соображений удобства.

Как правило, решение системы уравнений движения (21.3) содержит $2N$ произвольных постоянных, которые находятся из граничных условий (21.2). У краевой задачи решение может не существовать, а если оно существует, то может не быть единственным. Вместо краевой задачи можно также поставить задачу Коши, которая имеет единственное решение при корректной постановке. В этом случае произвольные постоянные, возникающие в решении уравнений движения, находятся из начальных данных для q и $\dot{q}$ в начальный момент времени $t = 0$. Именно эта задача наиболее часто ставится в физических приложениях для уравнений движения (21.3).

Перейдем к гамильтонову или каноническому формализму. В этом случае механическая система, состоящая из N частиц, описывается N обобщенными координатами q^i и N обобщенными импульсами

$$p_i := \frac{\partial L}{\partial \dot{q}^i}, \quad (21.4)$$

которые рассматриваются, как независимые переменные. Обобщенные координаты и импульсы являются координатами механической системы в $2N$ -мерном *фазовом* пространстве. Мы говорим, что координаты и импульсы являются канонически сопряженными переменными.

Предположим, что конфигурационное пространство с координатами q^i является N -мерным многообразием M . Тогда производные $\dot{q}^i$ являются компонентами касательного вектора вдоль траектории частицы. Это значит, что лагранжиан $L(q, \dot{q})$ является функцией (скалярным полем) на касательном расслоении $T(M)$. Тогда формула (21.4) определяет компоненты ковариантного вектора (1-формы). Это значит, что фазовое пространство есть ни что иное, как кокасательное расслоение $T^*(M)$ к конфигурационному пространству M .

Поскольку q^i и p_i являются координатами в базе и в кокасательном слое, то мы пишем координатные индексы, соответственно, вверх и вниз.

Предположим, что функция Лагранжа $L(q, \dot{q})$ выпукла по скоростям, т.е. квадратичная форма

$$\partial^2 L / \partial \dot{q}^i \partial \dot{q}^j, \quad (21.5)$$

которая называется *гесссианом*, положительно определена. Это условие не является ковариантным относительно произвольной замены координат на касательном расслоении $\mathbb{T}(\mathbb{M})$. Однако оно ковариантно относительно общих преобразований координат в конфигурационном пространстве $\mathbb{M}$. Действительно, при замене координат $q \mapsto q'(q)$ скорости преобразуются как векторы:

$$\dot{q}^i \mapsto \dot{q}'^i = \frac{\partial q'^i}{\partial q^j} \dot{q}^j.$$

Поскольку якобиан преобразования $\partial q'^i / \partial q^j$ не зависит от скоростей, то гесссиан преобразуется как ковариантный тензор второго ранга:

$$\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \mapsto \frac{\partial^2 L}{\partial \dot{q}'^i \partial \dot{q}'^j} = \frac{\partial q^k}{\partial q'^i} \frac{\partial}{\partial \dot{q}^k} \left(\frac{\partial q^l}{\partial q'^j} \frac{\partial L}{\partial \dot{q}^l} \right) = \frac{\partial q^k}{\partial q'^i} \frac{\partial q^l}{\partial q'^j} \frac{\partial^2 L}{\partial \dot{q}^k \partial \dot{q}^l}.$$

Таким образом, понятие выпуклости лагранжиана по скоростям не зависит от выбора координат в конфигурационном пространстве $\mathbb{M}$.

Замечание. Рассмотрим координаты q и время t в определении обобщенных импульсов (21.4) как параметры. Тогда формулы (21.4) задают скорости $\dot{q}$ как неявные функции обобщенных импульсов p . Из курса математического анализа известно, что уравнения (21.4) локально разрешимы относительно скоростей тогда и только тогда, когда гесссиан (21.5), который в данном случае совпадает с якобианом преобразования координат $\dot{q} \mapsto p$, является невырожденной матрицей. $\square$

Если функция Лагранжа выпукла по скоростям, то можно построить ее преобразование Лежандра (по скоростям):

$$H(q, p) := p_i \dot{q}^i - L(q, \dot{q}), \quad (21.6)$$

которое называется *функцией Гамильтона* или *гамильтонианом* системы. В силу определения преобразования Лежандра гамильтониан системы зависит только от обобщенных координат и импульсов. Это можно проверить и непосредственно

$$\frac{\partial H}{\partial \dot{q}^i} = p_i - \frac{\partial L}{\partial \dot{q}^i} = 0,$$

что следует из определения обобщенных импульсов (21.4).

Используя связь между гамильтонианом и лагранжианом (21.6), рассмотрим действие

$$S[q, p] = \int_{t_1}^{t_2} dt (p_i \dot{q}^i - H(q, p)) \quad (21.7)$$

как функционал от канонически сопряженных координат и импульсов. Соответствующая система уравнений Эйлера–Лагранжа имеет вид

$$\frac{\delta S}{\delta q^i} = -\dot{p}_i - \frac{\partial H}{\partial q^i} = 0, \quad (21.8)$$

$$\frac{\delta S}{\delta p_i} = \dot{q}^i - \frac{\partial H}{\partial p_i} = 0. \quad (21.9)$$

Нетрудно проверить, что N уравнений Эйлера–Лагранжа второго порядка (21.3) эквивалентны $2N$ уравнениям первого порядка (21.8). Понижение порядка уравнений движения произошло за счет введения новых независимых переменных – импульсов.

Поскольку конфигурационное и фазовое пространства имеют разную размерность, то уточним понятие эквивалентности. Для однозначного определения траектории частицы $q(t)$ в конфигурационном пространстве необходимо решить уравнения Эйлера–Лагранжа второго порядка (21.3), например, с начальными условиями $q(0) = q_0$, $\dot{q}(0) = q_1$. Траектория движения в фазовом пространстве $q(t), p(t)$ однозначно находится решением уравнений движения первого порядка (21.8), (21.9) с начальными условиями $q(0) = q_0$, $p(0) = p_0$. Каждая траектория в фазовом пространстве естественным образом проектируется на траекторию в конфигурационном пространстве $\{q(t), p(t)\} \mapsto \{q(t)\}$, при этом начальное условие для скорости q_1 определяется из уравнения (21.9). Обратно. Для любой траектории в конфигурационном пространстве $q(t)$ уравнение (21.9) определяет импульсы $p(t)$ и начальные условия p_0 так, что пара $q(t), p(t)$ является траекторией в фазовом пространстве. Здесь мы предполагаем, что уравнение (21.9) имеет единственное решение для p (в противном случае преобразование Лежандра не определено).

При постановке граничной задачи для действия (21.7) мы не можем рассматривать траектории, соединяющие две произвольные точки (q_1, p_1) и (q_2, p_2) фазового пространства, т.к. тогда возникнет $4N$ условий для $2N$ обыкновенных дифференциальных уравнений первого порядка (21.8), (21.9). В этом случае количество граничных условий превышает количество уравнений, и задача может не иметь решения. Поэтому можно рассматривать, например, те траектории, которые имеют начало и конец на N -мерных подмногообразиях фазового пространства, определяемых условиями (21.2), как и в лагранжевом подходе. При этом никаких граничных условий на импульсы не возникает, поскольку они входят в подинтегральное выражение без производных, и при вариации импульсов интегрирования по частям не происходит.

Замечание. Как правило для уравнений Эйлера–Лагранжа в фазовом пространстве ставится задача Коши, которая имеет единственное решение. При этом начальной точкой траектории может быть произвольная точка фазового пространства $T^*(\mathbb{R}^N) = \mathbb{R}^{2N}$. Тем самым мы предполагаем, что фазовое пространство так же как и конфигурационное пространство топологически тривиально. Обсуждение области определения различных функций в конфигурационном и фазовом пространствах требует знания явного вида лагранжиана, гамильтониана и анализа уравнений движения. Поскольку в общем случае учесть все возможности нельзя, то в дальнейшем при обсуждении общей схемы все рассматриваемые функции будут считаться определенными на всем фазовом пространстве и достаточно гладкими. В каждом конкретном случае области определения должны быть проанализированы, а постановка задачи уточнена. $\square$

Посмотрим на уравнения (21.8), (21.9) с другой точки зрения. Пусть заданы две функции от канонических переменных $f(q, p)$ и $g(q, p)$, т.е. два скалярных поля на фазовом пространстве. Определим для них *скобку Пуассона*

$$[f, g] := \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i}. \quad (21.10)$$

Легко проверить, что она обладает следующими свойствами:

- 1) $[af + bg, h] = a[f, g] + b[g, h]$, $a, b \in \mathbb{R}$ — линейность,
- 2) $[f, g] = -[g, f]$ — антисимметрия,
- 3) $[f, gh] = [f, g]h + g[f, h]$ — правило Лейбница,
- 4) $[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0$ — тождество Якоби,

и, следовательно, определяет пуассонову структуру на фазовом пространстве (см. раздел 19.3).

Скобку Пуассона (21.10) можно переписать в эквивалентном виде. Обозначим координаты фазового пространства через $\{x^\alpha\} := (q^1 \dots q^N, p_1 \dots p_N)$, $\alpha = 1, \dots, 2N$, или, короче, $x = (q, p)$. Тогда

$$[f, g] = \varpi^{-1\alpha\beta} \frac{\partial f}{\partial x^\alpha} \frac{\partial g}{\partial x^\beta}, \quad (21.11)$$

где

$$\varpi^{-1\alpha\beta} = [x^\alpha, x^\beta] = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$$

– обратная каноническая пуассонова структура. Каноническая форма ϖ невырождена и точна, $d\varpi = 0$, поскольку ее компоненты постоянны. Поэтому фазовое пространство представляет собой также симплектическое многообразие (см. раздел 19.2).

Рассмотрим простейшие свойства скобки Пуассона. Очевидно, скобки Пуассона произвольной функции f с самой собой и константой c равны нулю:

$$[f, f] = 0, \quad [f, c] = 0.$$

Из определения (21.10) следует, что скобка Пуассона обобщенной координаты с импульсом равна символу Кронекера:

$$[q^i, p_j] = \delta_j^i. \quad (21.12)$$

Рассмотрим траекторию в фазовом пространстве $q(t), p(t)$, где $t \in \mathbb{R}$. Тогда, используя скобку Пуассона, уравнения движения (21.8), (21.9) можно переписать в виде

$$\begin{aligned} \dot{q}^i &= [q^i, H] = \frac{\partial H}{\partial p_i}, \\ \dot{p}_i &= [p_i, H] = -\frac{\partial H}{\partial q^i}. \end{aligned} \quad (21.13)$$

При этом скобка Пуассона для координат фазового пространства (21.12) рассматривается как одновременная:

$$[q^i(t), p_j(t)] = \delta_j^i, \quad \forall t \in \mathbb{R}. \quad (21.14)$$

Скобка Пуассона для различных моментов времени $[q^i(t_1), p_j(t_2)]$ при $t_1 \neq t_2$ не определена. Вообще, эволюция во времени любой функции $f(t, q, p)$, зависящей от времени и канонических переменных определяется уравнением

$$\dot{f} = \frac{df}{dt} = \frac{\partial f}{\partial t} + [f, H]. \quad (21.15)$$

В частности, если функция Гамильтона не зависит от времени явно, то

$$\frac{dH}{dt} = [H, H] = 0,$$

ввиду антисимметрии скобки Пуассона. Это значит, что для заданной траектории механической системы в фазовом пространстве гамильтониан является интегралом движения и его численное значение сохраняется. Это значение называется *энергией* механической системы и определяется начальными данными. Системы, у которых гамильтониан не зависит от времени, называются *консервативными*.

Пример 21.2.1 (Гармонический осциллятор с затуханием). Рассмотрим функцию Лагранжа для одной точечной частицы (осциллятора), которая явно зависит от времени

$$L = \frac{1}{2} e^{2\mu t} (\dot{q}^2 - \omega^2 q^2), \quad \mu, \omega = \text{const.}$$

Постоянные ω и μ называются, соответственно, собственной частотой и коэффициентом затухания осциллятора. Обобщенный импульс частицы равен

$$p := \frac{\partial L}{\partial \dot{q}} = e^{2\mu t} \dot{q}.$$

Гамильтониан осциллятора также зависит от времени явно:

$$H = \frac{1}{2} e^{-2\mu t} p^2 + \frac{1}{2} e^{2\mu t} \omega^2 q^2.$$

Гамильтоновы и лагранжевы уравнения движения для осциллятора с затуханием имеют вид

$$\left. \begin{aligned} \dot{q} &= e^{-2\mu t} p, \\ \dot{p} &= -e^{2\mu t} \omega^2 q, \end{aligned} \right\} \Leftrightarrow e^{2\mu t} (\ddot{q} + 2\mu \dot{q} + \omega^2 q) = 0.$$

Общее решение этих уравнений параметризуется двумя постоянными: амплитудой A_0 и фазой φ ,

$$q = A_0 e^{-\mu t} \cos(\tilde{\omega} t + \varphi), \quad \tilde{\omega}^2 := \omega^2 - \mu^2.$$

Произвольная фаза φ соответствует произволу в выборе начала отсчета времени, и в дальнейшем мы положим $\varphi = 0$. При малых коэффициентах затухания $\mu^2 < \omega^2$ амплитуда колебаний экспоненциально затухает, если $\mu > 0$. Затухающие колебания происходят с частотой, меньшей собственной частоты осциллятора $\tilde{\omega} < \omega$. Для произвольной траектории численное значение гамильтониана зависит от времени явно

$$H = \frac{1}{2} A_0^2 \omega^2 \left[1 - \frac{1}{2} \cos 2\tilde{\omega} t - \frac{1}{2} \cos(2\tilde{\omega} t + 2\psi) \right],$$

где $\cos \psi := \mu/\omega$. □

Функция канонических переменных $f(q, p)$ называется *интегралом движения*, если $f(q, p) = \text{const}$ для любого решения канонических уравнений движения, т.е. является интегралом уравнений движения.

Теорема 21.2.1. Если известны два интеграла движения f и g , то их скобка Пуассона $[f, g]$ также является интегралом движения.

Доказательство. Прямая проверка равенства $d[f, g]/dt = 0$. □

Замечание. Теорема не гарантирует того, что вычисление скобки Пуассона двух интегралов движения дает новый интеграл движения. Часто она равна нулю, или интегралы движения f , g и $[f, g]$ — функционально зависимы. □

Если гамильтониан не зависит от какой-либо из координат, например, от q^1 , т.е. $\partial H / \partial q^1 = 0$, то эта координата называется *циклической*, а соответствующий обобщенный импульс сохраняется в силу второго уравнения (21.13)

$$p_1 = c = \text{const.}$$

При этом изменение остальных координат и импульсов во времени такое же, как в системе с координатами $q^2, \dots, q^N$, импульсами $p_2, \dots, p_N$ и функцией Гамильтона $H(q^2, \dots, q^N, c, p_2, \dots, p_N)$.

В большинстве физических приложений для системы уравнений Гамильтона (21.13) решается задача Коши, т.е. ищется решение системы уравнений (21.13) с заданными начальными условиями

$$q(0) = q_0, \quad p(0) = p_0. \quad (21.16)$$

Канонические уравнения движения (21.13) можно записать в виде

$$\dot{x} = X,$$

где $x = (q, p)$, а правая часть уравнений определяются векторным полем

$$X = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i}$$

на фазовом пространстве. С геометрической точки зрения решение уравнений Гамильтона с заданными начальными условиями задает интегральную кривую этого векторного поля, проходящую через точку (21.16). Как было показано в разделе 2.6.5, интегральные кривые векторного поля задают абелеву однопараметрическую группу преобразований многообразия, которым в данном случае является фазовое пространство. Эта группа преобразований в гамильтоновой динамике имеет специальное название.

Определение. *Фазовым потоком* называется однопараметрическая группа преобразований фазового пространства

$$s_t : \quad \mathbb{T}^*(\mathbb{M}) \ni \quad q(0), p(0) \mapsto q(t), p(t) \quad \in \mathbb{T}^*(\mathbb{M}),$$

где $q(t), p(t)$ – решение системы уравнений Гамильтона (21.13). $\square$

21.3 Потенциальное движение точечной частицы

Потенциальное движение точечной частицы массы m в евклидовом пространстве $\mathbb{R}^N$ с декартовыми координатами q^i , $i = 1, \dots, N$ описывается следующей функцией Лагранжа

$$L = m \frac{\dot{q}^i \dot{q}_i}{2} - U(q), \quad (21.17)$$

где $U(q) \geq 0$ – некоторая положительно определенная функция координат. В настоящем разделе подъем и опускание индексов производится с помощью евклидовой метрики $\delta_{ij} = \text{diag}(+ \dots +)$. Уравнения движения (*уравнения Ньютона*) при этом имеют вид

$$m \ddot{q}^i = - \frac{\partial U}{\partial q_i}. \quad (21.18)$$

Это – система обыкновенных дифференциальных уравнений второго порядка и их решение зависит от $2N$ произвольных постоянных. Для их определения решают, как правило, задачу Коши. То есть ищется решение уравнений (21.18) с заданными начальными условиями:

$$q(0) = q_0, \quad \dot{q}(0) = v_0.$$

В физических приложениях функции $U(q)$ обычно таковы, что решение этой задачи существует, единственно и определено при всех $t \in (-\infty, \infty)$. Заметим, что потенциальная энергия определена с точностью до постоянной, которая не влияет на уравнения движения.

Если потенциальная энергия имеет локальный экстремум в точке q_0 , то

$$\left. \frac{\partial U}{\partial q^i} \right|_{q=q_0} = 0.$$

Следовательно, постоянная траектория $q = q_0$ удовлетворяет системе уравнений (21.18). Таким образом, локальные экстремумы потенциальной энергии определяют положения равновесия частицы. Эти положения могут быть устойчивы или не устойчивы по отношению к малым возмущениям в зависимости от того является ли локальный экстремум минимумом или максимумом потенциальной энергии. Для положительно определенной потенциальной энергии существует по крайней мере одна точка равновесия.

Посмотрим на эту задачу с точки зрения принципа наименьшего действия. Во-первых, действие

$$S[q] = \int dt \left(m \frac{\dot{q}^i \dot{q}_i}{2} - U(q) \right)$$

не является положительно определенным. Поэтому нужно говорить не о минимуме действия, а только о стационарных точках. Во-вторых, для свободного движения, $U = 0$, траектории представляют собой прямые линии

$$q = v_0 t + q_0,$$

по которым частица движется с постоянной скоростью v_0 . Соответствующее действие при интегрировании по бесконечному интервалу расходится и говорить даже о стационарности действия в этом случае не имеет смысла. Подытожить сделанные замечания можно следующим образом. Потенциальное движение точечной частицы происходит таким образом, что для каждого конечного интервала времени (t_1, t_2) действие стационарно среди всех возможных траекторий, соединяющих точки q_1 и q_2 . При этом первую граничную точку можно выбрать произвольным образом, а вторая должна быть такой, чтобы задача Коши, поставленная в первой точке, имела решение для некоторой скорости $\dot{q}_1 = v_0$.

Переформулируем потенциальное движение точечной частицы на гамильтоновом языке. Компоненты импульса и гамильтониан точечной частицы равны

$$\begin{aligned} p_i &= m \dot{q}_i, \\ H &= \frac{p^i p_i}{2m} + U(q). \end{aligned} \quad (21.19)$$

Гамильтоновы уравнения движения принимают вид

$$\begin{aligned} \dot{q}^i &= \frac{p^i}{m}, \\ \dot{p}_i &= -\frac{\partial U}{\partial q^i}, \end{aligned} \quad (21.20)$$

которые эквивалентны уравнениям Ньютона (21.18). Заметим, что в данном случае первое уравнение совпадает с определением импульса частицы.

Гамильтониан точечной частицы (21.19) положительно определен и для заданной траектории его значение (энергия) постоянно. Первое и второе слагаемые в (21.19) называются, соответственно, *кинетической* и *потенциальной* энергией точечной частицы. Следующий пример показывает, что функция Гамильтона зависит от выбора системы координат в фазовом пространстве, а каноническая пуассонова структура – нет.

Пример 21.3.1. Рассмотрим частицу массы m , которая движется в трехмерном конфигурационном пространстве $q \in \mathbb{R}^3$ с заданным потенциалом U . Гамильтониан является функцией на фазовом пространстве и в произвольной криволинейной системе координат имеет вид

$$H = \frac{1}{2m} g^{ij} p_i p_j + U(q),$$

где $g^{ij} = g^{ij}(q)$ – обратная метрика в выбранной системе отсчета. Например, в декартовых координатах x, y, z

$$H = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + U(x, y, z).$$

В цилиндрических координатах r, φ, z

$$H = \frac{1}{2m} \left(p_r^2 + \frac{p_\varphi^2}{r^2} + p_z^2 \right) + U(r, \varphi, z).$$

В сферических координатах r, θ, φ

$$H = \frac{1}{2m} \left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2 \theta} \right) + U(r, \theta, \varphi). \quad \square$$

Каноническая пуассонова структура была определена в произвольной, в общем случае криволинейной, системе координат. Проверим корректность этого определения относительно преобразования координат, поскольку при переходе от одной системы координат к другой структурные функции могли бы измениться. Допустим, что мы определили каноническую пуассонову структуру в декартовой системе координат. В рассматриваемом случае преобразование координат имеет вид $q, p \mapsto Q(q), P(q, p)$, где большие буквы обозначают криволинейные (сферические или цилиндрические) координаты в конфигурационном пространстве. Скобки Пуассона новых координат и импульсов равны:

$$\begin{aligned} [Q^i, Q^j] &= 0, \\ [Q^i, P_j] &= \frac{\partial Q^i}{\partial q^k} \frac{\partial P_j}{\partial p_k}, \\ [P_i, P_j] &= \frac{\partial P_i}{\partial q^k} \frac{\partial P_j}{\partial p_k} - \frac{\partial P_i}{\partial p_k} \frac{\partial P_j}{\partial q^k}. \end{aligned}$$

Поскольку компоненты обобщенных импульсов являются компонентами ковектора, то они преобразуются по правилу

$$P_i = \frac{\partial q^k}{\partial Q^i} p_k.$$

Теперь нетрудно проверить выполнение скобок Пуассона:

$$[Q^i, P_j] = \delta_j^i, \quad [Q^i, Q^j] = 0, \quad [P_i, P_j] = 0.$$

Таким образом, структурные функции имеют канонический вид и не зависят от выбора системы координат в конфигурационном пространстве $\mathbb{R}^3$. Рассматриваемые преобразования координат $q, p \mapsto Q, P$ в фазовом пространстве относятся к классу канонических преобразований, которые будут рассмотрены позже в разделе 21.11.

Определение. Преобразование координат $q = q(Q)$ конфигурационного пространства $\mathbb{R}^N$ называется *точечным*. $\square$

21.4 Комплексный язык в гамильтоновом формализме

В фазовом пространстве координаты q^i и импульсы p_i имеют разные трансформационные свойства при преобразовании координат конфигурационного пространства $q^i \mapsto q'^i(q)$. По этой причине мы различаем верхние и нижние индексы. Следовательно, запись $q^i + p_i$ некорректна. Чтобы записать гамильтонову динамику в комплексных переменных мы предположим, что на конфигурационном пространстве задана евклидова метрика, и латинские индексы $i, j = 1, \dots, N$ можно поднимать и опускать, например, $p^i := \delta^{ij} p_j$. Тем самым мы сужаем группу преобразований координат конфигурационного пространства до группы вращений $\mathcal{O}(n, \mathbb{R})$. Это удобно. В конечном итоге, гамильтонова система определяется заданием гамильтониана $H(q, p)$, который является достаточно гладкой вещественнозначной функцией на фазовом пространстве. При задании гамильтониана можно отказаться от существования евклидовой метрики на конфигурационном пространстве и рассматривать функции от q^i и p_i .

Введем комплексные координаты на фазовом пространстве:

$$\begin{aligned} z^i &:= q^i + ip^i, \\ \bar{z}^i &:= q^i - ip^i, \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} q^i &= \frac{1}{2}(z^i + \bar{z}^i), \\ p^i &= \frac{1}{2i}(z^i - \bar{z}^i). \end{aligned} \quad (21.21)$$

Тогда $H(q, p) = H(z, \bar{z})$, и гамильтоновы уравнения движения,

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}, \quad (21.22)$$

можно переписать в виде

$$i\dot{z}^i = 2\frac{\partial H}{\partial \bar{z}_i}. \quad (21.23)$$

Комплексно сопряженные уравнения выглядят следующим образом

$$-i\dot{\bar{z}}^i = 2\frac{\partial \bar{H}}{\partial z_i}. \quad (21.24)$$

Таким образом, система из $2N$ гамильтоновых уравнений (21.22) эквивалентна системе N комплексных уравнений (21.23) или (21.24). Гамильтониан, по определению, должен быть вещественнозначным:

$$\bar{H}(\bar{z}, z) = H(z, \bar{z}).$$

Тем самым мы превратили $2N$ -мерное фазовое пространство в N -мерное комплексное многообразие, на котором гамильтониан задается вещественнозначной функцией $H(z, \bar{z})$, и динамика системы определяется дифференциальным уравнением первого порядка (21.23).

Пример 21.4.1. Рассмотрим линейную гамильтонову систему с квадратичным гамильтонианом

$$H = \frac{1}{2} z^i D_{ij} z^j + z^i \tilde{H}_{ij} \bar{z}^j + \frac{1}{2} \bar{z}^i \bar{D}_{ij} \bar{z}^j, \quad (21.25)$$

где D и $\tilde{H}$ – некоторые комплексные $n \times n$ матрицы. Для вещественности гамильтониана необходимо и достаточно, чтобы матрица $\tilde{H}$ была эрмитовой, $\tilde{H} = \tilde{H}^\dagger$, а симметричная матрица D может быть произвольна. Комплексные уравнения движения принимают вид

$$i\dot{z}^i = 2z^j \tilde{H}_j^i + 2\bar{z}^j \bar{D}_j^i. \quad (21.26)$$

Если $D = 0$, то гамильтоново уравнение (21.26) является линейным уравнением Шредингера

$$i\dot{z} = 2z\tilde{H},$$

где опущены матричные индексы и вектор z записан в виде строки. Решение задачи Коши $z(0) = z_0$ для этого уравнения имеет вид

$$z = z_0 e^{-2i\tilde{H}t}, \quad e^{-2i\tilde{H}t} \in \mathbb{U}(N, \mathbb{C}).$$

В этом случае компоненты $\{z^i\} \in \mathbb{C}^N$ образуют вектор гильбертова пространства, и его норма $\|z\| := z^i \bar{z}_i = \|z_0\|$ сохраняется во времени. $\square$

21.5 Лемма Стокса

В этом и следующих разделах механика частиц будет рассмотрена с более общей точки зрения. Начнем с геометрического рассмотрения, которое затем применим к гамильтоновой динамике точечных частиц. Пусть на многообразии $\mathbb{M}$ нечетной размерности $\dim \mathbb{M} = 2N + 1$ задана 2-форма

$$B = \frac{1}{2} dx^\alpha \wedge dx^\beta B_{\alpha\beta}.$$

Поскольку матрица, задающая 2-форму в локальной системе координат антисимметрична, $B_{\alpha\beta} = -B_{\beta\alpha}$, а многообразие нечетномерно, то ее определитель равен нулю. Это значит, что в каждой точке $x \in \mathbb{M}$ у нее существует по крайней мере один нетривиальный собственный вектор $X = X^\alpha \partial_\alpha$ с нулевым собственным значением. Отсюда вытекает, что для 2-формы B существует *нулевое* векторное поле X со свойством

$$B(X, Y) = X^\alpha Y^\beta B_{\alpha\beta} = 0, \quad \forall Y = Y^\alpha \partial_\alpha.$$

Пространство нулевых векторных полей линейно.

Определение. 2-форма называется *неособой*, если размерность пространства нулевых векторов минимальна, т.е. равна нулю или единице на многообразиях четной и нечетной размерности, соответственно. $\square$

Пример 21.5.1. Рассмотрим 2-форму ϖ в фазовом пространстве $\mathbb{R}^{2N}$ с координатами $x = (x^1 \dots x^{2N}) = (q^1 \dots q^N, p_1 \dots p_N)$,

$$\varpi = dp_i \wedge dq^i = dx^{N+1} \wedge dx^1 + dx^{N+2} \wedge dx^2 + \dots + dx^{2N} \wedge dx^N = \frac{1}{2} dx^\alpha \wedge dx^\beta \varpi_{\alpha\beta}, \quad (21.27)$$

где ϖ – каноническая симплектическая форма (19.1). Эта форма неособа, т.к. ее определитель равен единице $\det \varpi_{\alpha\beta} = 1$ и размерность пространства нулевых векторов равна нулю. $\square$

Пусть на многообразии $\mathbb{M}$, $\dim \mathbb{M} = 2N + 1$, задана 1-форма $A = dx^\alpha A_\alpha$. Предположим, что внешний дифференциал этой формы является неособым. Это значит, что в каждой точке многообразия $x \in \mathbb{M}$ существует единственный, с точностью до умножения на постоянную, нулевой вектор X такой, что

$$dA(X, Y) = 0, \quad \forall Y. \quad (21.28)$$

Тем самым 1-форма A определяет нулевое векторное поле X на $\mathbb{M}$ с точностью до умножения на отличную от нуля функцию. Назовем интегральные кривые $x^\alpha(t)$,

$$\frac{dx^\alpha}{dt} = X^\alpha,$$

нулевого векторного поля *характеристиками* формы A . Далее, пусть γ_1 – замкнутая кривая на $\mathbb{M}$, которая ни в одной своей точке не касается характеристик. Тогда множество характеристик, выходящих из точек кривой γ_1 , образует *трубку характеристик*. Трубка характеристик определена по крайней мере в некоторой окрестности кривой γ_1 .

Лемма 21.5.1 (Стокс). *Интеграл от 1-формы A с неособым внешним дифференциалом dA по любой из двух замкнутых кривых γ_1 и γ_2 , охватывающих одну и ту же трубку характеристик, одинаков:*

$$\oint_{\gamma_1} A = \oint_{\gamma_2} A,$$

если $\gamma_1 - \gamma_2 = \partial S$, где S – часть трубки характеристик.

Доказательство. По формуле Стокса справедливы равенства

$$\oint_{\gamma_1} A - \oint_{\gamma_2} A = \oint_{\partial S} A = \int_S dA.$$

Этот интеграл обращается в нуль, т.к. значение формы dA на любой паре векторов, касательных к трубке характеристик, равен нулю в силу (21.28). $\square$

21.6 Канонические уравнения Гамильтона

Из безобидной, на первый взгляд, леммы Стокса непосредственно вытекают все основные положения гамильтоновой динамики. Предположим, что механической системе соответствует топологически тривиальное фазовое пространство $\mathbb{T}^*(\mathbb{R}^N) \approx \mathbb{R}^{2N}$. Рассмотрим *расширенное фазовое пространство* $\mathbb{R}^{2N+1}$ с координатами $q^1, \dots, q^N, p_1, \dots, p_N, t$, в котором время t рассматривается, как независимая дополнительная координата.

Замечание. Предположение о том, что частица движется в тривиальном конфигурационном пространстве $\mathbb{R}^N$ и, следовательно, в тривиальном фазовом пространстве $T^*(\mathbb{R}^N) \approx \mathbb{R}^{2N}$ не означает, что приведенные ниже формулы справедливы только для движения в топологически тривиальных пространствах. Данное предположение просто означает, что мы рассматриваем механическую систему в определенной карте. При движении частицы по нетривиальному многообразию необходимо дополнительно проследить за склейкой карт. В дальнейшем мы не будем обсуждать этот вопрос. $\square$

Пусть на расширенном фазовом пространстве задана некоторая функция Гамильтона $H(q, p, t)$. Определим 1-форму на расширенном фазовом пространстве

$$A := dq^i p_i - dt H. \quad (21.29)$$

Эта форма называется *интегральным инвариантом Пуанкаре–Картана*. Точнее, интегральный инвариант получается после интегрирования формы A по замкнутому контуру в расширенном фазовом пространстве. Смысл названия будет ясен из дальнейшего рассмотрения. Внешний дифференциал 1-формы (21.29) является неособым, т.к. каноническая форма $\varpi = dp_i \wedge dq^i$ неособа.

Теорема 21.6.1. *Характеристики формы A в расширенном фазовом пространстве $\mathbb{R}^{2N+1}$ однозначно проектируются на фазовое пространство q, p , т.е. задаются функциями $q(t)$, $p(t)$. Эти функции удовлетворяют уравнениям Гамильтона:*

$$\begin{aligned} \frac{dq^i}{dt} &= \frac{\partial H}{\partial p_i}, \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial q^i}, \end{aligned} \quad (21.30)$$

Другими словами, характеристики формы A представляют собой траектории фазового потока в расширенном фазовом пространстве, т.е. интегральные кривые канонических уравнений (21.30).

Доказательство. Дифференциал формы (21.29) равен

$$dA = dp_i \wedge dq^i - \frac{\partial H}{\partial q^i} dq^i \wedge dt - \frac{\partial H}{\partial p_i} dp_i \wedge dt$$

Эта 2-форма неособа, т.к. ранг матрицы, составленной из ее координат, равен $2N$. Прямая подстановка показывает, что вектор

$$X = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i} + \frac{\partial}{\partial t} \quad (21.31)$$

является нулевым вектором формы A

$$dA(X, Y) = -\frac{\partial H}{\partial q^i} Y^i - Y_i \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i} Y^0 \frac{\partial H}{\partial q^i} + Y^i \frac{\partial H}{\partial q^i} + \frac{\partial H}{\partial q^i} Y^0 \frac{\partial H}{\partial p_i} + Y_i \frac{\partial H}{\partial p_i} = 0.$$

Это значит, что векторное поле (21.31) задает направление характеристик 1-формы (21.29). С другой стороны, вектор (21.31) является вектором скорости фазового потока (21.30). Действительно, интегральные кривые $\{q(\tau), p(\tau), t(\tau)\}$ в расширенном фазовом пространстве задаются уравнениями

$$\frac{dq^i}{d\tau} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{d\tau} = -\frac{\partial H}{\partial q^i}, \quad \frac{dt}{d\tau} = 1. \quad (21.32)$$

В силу последнего уравнения они однозначно проектируются на фазовое пространство. Таким образом, интегральные кривые (21.30) представляют собой проекции характеристик формы (21.29) на фазовое пространство. $\square$

Применим теперь лемму Стокса и интегральному инварианту Пуанкаре–Картана.

Теорема 21.6.2. Пусть две замкнутые кривые γ_1 и γ_2 охватывают одну и ту же трубку характеристик формы (21.29). Тогда интегралы по ним от интегрального инварианта Пуанкаре–Картана одинаковы:

$$\oint_{\gamma_1} (dq p - dt H) = \oint_{\gamma_2} (dq p - dt H).$$

Именно поэтому 1-форма $dq p - dt H$ называется интегральным инвариантом.

В частном случае, когда замкнутые кривые лежат в фазовом подпространстве, соответствующем постоянному значению времени, $t = \text{const} \Leftrightarrow dt = 0$, получаем

Следствие. Фазовый поток сохраняет интеграл

$$\oint_{\gamma} dq p$$

вдоль произвольной замкнутой кривой в фазовом пространстве.

Форма $dq p := dq^i p_i$ в фазовом пространстве называется *относительным интегральным инвариантом Пуанкаре* или *формой Лиувилля*. Точнее, интегральный инвариант получается после интегрирования формы Лиувилля по замкнутому контуру γ . Он имеет простой геометрический смысл. Пусть S – двумерная ограниченная ориентированная поверхность с кусочно гладкой границей такая, что $\partial S = \gamma$, тогда по формуле Стокса

$$\oint_{\gamma} dq p = \int_S dp \wedge dq,$$

где $dp \wedge dq := dp_i \wedge dq^i$. Отсюда вытекает

Следствие. Фазовый поток $s_t := \exp(tX_H)$ сохраняет сумму ориентированных площадей проекций поверхности на N координатных плоскостей q^i, p_i :

$$\int_S dp \wedge dq = \int_{s_t S} dp \wedge dq.$$

В этом смысле каноническая 2-форма $\varpi = dp \wedge dq$ является абсолютным интегральным инвариантом фазового потока s_t .

Замечание. Эпитеты “относительный” и “абсолютный” интегральный инвариант не несут глубокого смысла. Исторически сложилась так, что относительным называют инвариант, полученный после интегрирования по замкнутому (компактному и без края) многообразию. Абсолютным называют инвариант, возникающий после интегрирования некоторой формы по компактному многообразию с краем. $\square$

21.7 Принцип Мопертюи

В приложениях часто встречаются гамильтонианы, не зависящие от времени явно $H = H(q, p)$. В этом случае численное значение гамильтониана (энергия) на фазовой траектории сохраняется

$$\frac{dH}{dt} = [H, H] = 0.$$

Покажем, что наличие интеграла энергии позволяет понизить размерность расширенного фазового пространства $\mathbb{R}^{2N+1}$ на две единицы и свести задачу к интегрированию некоторой системы канонических уравнений в $(2N - 1)$ -мерном пространстве.

Предположим, что в некоторой области фазового пространства уравнение

$$H(q, p) = E$$

можно решить относительно p_1 :

$$p_1 = K(Q, P, T; E),$$

где $Q := (q^2, \dots, q^N)$, $P := (p_2, \dots, p_N)$ и $T = -q^1$. Тогда

$$pdq - Hdt = PdQ - KdT - d(Ht) + tdH. \quad (21.33)$$

Пусть γ – интегральная кривая канонических уравнений (21.30). Она лежит на $2N$ -мерном подмногообразии $H(q, p) = E$ в расширенном фазовом пространстве $\mathbb{R}^{2N+1}$. Спроектируем расширенное фазовое пространство на фазовое пространство $\mathbb{R}^{2N}$. При этом поверхность $H = E$ проектируется на $(2N - 1)$ -мерное подмногообразие $\mathbb{M}^{2N-1}$ фазового пространства, которое определяется тем же уравнением $H = E$, а кривая γ – на кривую $\bar{\gamma}$, лежащую на этом подмногообразии. Переменные Q, P, T образуют локальную систему координат на $\mathbb{M}^{2N-1}$. Но характеристики формы $PdQ - KdT$ удовлетворяют уравнениям Гамильтона, поскольку $dH = 0$ на $\mathbb{M}^{2N-1}$, а полный дифференциал $d(Ht)$ в (21.33) не влияет на уравнения движения. Это доказывает следующее утверждение.

Теорема 21.7.1. *Если гамильтониан не зависит от времени явно, то фазовые траектории канонических уравнений (21.30) на подмногообразии $\mathbb{M}^{2N-1}$ фазового пространства, определяемом уравнением $H(q, p) = E$, удовлетворяют каноническим уравнениям*

$$\frac{dq^i}{dq^1} = -\frac{\partial K}{\partial p_i}, \quad \frac{dp_i}{dq^1} = \frac{\partial K}{\partial q^i}, \quad i = 2, \dots, N, \quad (21.34)$$

где функция $K(q^1, \dots, q^N, p_2, \dots, p_N, E)$ определяется уравнением

$$H(q^1, \dots, q^N, K, p_2, \dots, p_N) = E.$$

Замечание. В этой теореме роль времени играет первая координата q^1 , и канонические уравнения определяют не динамику системы, а форму траектории. $\square$

Покажем, каким образом канонические уравнения Гамильтона связаны с принципом наименьшего действия. Рассмотрим интегральную кривую уравнений (21.32) в расширенном фазовом пространстве (q, p, t) , соединяющую две точки (q_0, p_0, t_0) и (q_1, p_1, t_1) .

Теорема 21.7.2. *Кривая γ является стационарной точкой интеграла*

$$\int_{\gamma} (dq p - dt H) \quad (21.35)$$

при таких вариациях γ , когда концы кривой остаются на N -мерных подмногообразиях ($q = q_0, t = t_0$) и ($q = q_1, t = t_1$).

Доказательство. В силу третьего уравнения (21.32) вариацию интеграла (21.35) можно записать как вариацию функционала действия

$$\delta \int_{\gamma} dt (p \dot{q} - H) = p \delta q \Big|_{q_0}^{q_1} + \int_{\gamma} dt \left[\left(\dot{q}^i - \frac{\partial H}{\partial p_i} \right) \delta p_i - \left(\dot{p}_i + \frac{\partial H}{\partial q^i} \right) \delta q^i \right].$$

Отсюда вытекает, что для исчезновения граничных вкладов достаточно зафиксировать только значения обобщенных координат при $t = t_{0,1}$. $\square$

Понижение размерности расширенного фазового пространства при гамильтониане, не зависящем от времени, позволяет по новому взглянуть и в определенной степени оправдать употребление термина “принцип наименьшего действия” в механике. Фазовые траектории исходных канонических уравнений (21.30) целиком лежат на подмногообразии $\mathbb{M} \subset \mathbb{R}^{2n+1}$, $\dim \mathbb{M} = 2N - 1$, соответствующем фиксированному значению энергии, и являются характеристиками формы $dq p = dQP - dTK$. Отсюда следует

Теорема 21.7.3. *Если функция Гамильтона не зависит от времени явно, то фазовые траектории канонических уравнений (21.30), лежащие на подмногообразии $\mathbb{M}^{2n-1}$, соответствующем фиксированному значению энергии $H(q, p) = E$, являются стационарными точками интеграла*

$$\int dq p$$

в классе кривых, лежащих на $\mathbb{M}^{2n-1}$ и соединяющих подпространства $q = q_0$ и $q = q_1$.

Рассмотрим теперь проекцию экстремали, лежащую на подмногообразии постоянной энергии $\mathbb{M}$ на конфигурационное пространство. Эта кривая соединяет точки с координатами q_0 и q_1 . Пусть $q = q(\tau)$, $\tau \in [a, b]$ – некоторая кривая, соединяющая те же точки $q(a) = q_0$ и $q(b) = q_1$ в конфигурационном пространстве. Она является проекцией некоторой кривой γ на подмногообразии $\mathbb{M}$. Эту кривую нетрудно построить. Для этого достаточно найти обобщенный импульс $p = \partial L / \partial \dot{q}$, где $\dot{q} := dq/d\tau$ – вектор скорости кривой в конфигурационном пространстве. Если параметр τ подобран так, что $H(q, p) = E$, то мы получаем кривую $q = q(\tau)$, $p = \partial L / \partial \dot{q}$ на поверхности $\mathbb{M}$. Применяя предыдущую теорему, получаем

Следствие. Среди всех кривых $q = q(\tau)$, $\tau \in [a, b]$ соединяющих точки q_0 и q_1 в конфигурационном пространстве и параметризованных так, что функция Гамильтона имеет фиксированное значение

$$H(q, \partial L / \partial \dot{q}) = E$$

траекторией движения механической системы является экстремаль *укороченного действия*

$$\int_a^b d\tau p_i \dot{q}^i = \int_a^b d\tau \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i. \quad (21.36)$$

Это следствие называется *принципом наименьшего действия Мопертюи*. Важно отметить, что отрезок $\tau \in [a, b]$ не фиксирован и может быть разным у сравниваемых кривых. Зато одинаковой должна быть энергия. Принцип Мопертюи определяет только форму траектории, а для определения зависимости координат от времени необходимо воспользоваться условием постоянства энергии. Заметим также, что во многих задачах подынтегральное выражение в (21.36) положительно определено. В таких случаях движение происходит действительно по экстремалиям укороченного действия, и можно говорить о принципе наименьшего действия.

Пример 21.7.1. Покажем, что если материальная точка движется по риманову многообразию $(\mathbb{M}, g)$ только под действием сил инерции, то движение происходит по экстремалиям. Пусть $q(t) = \{q^i(\tau)\}$ – траектория частицы и $ds^2 = dq^i dq^j g_{ij}$ – интервал риманова многообразия. Тогда функция Лагранжа для инерциального движения частицы определяется только кинетической энергией:

$$L = T = H = \frac{1}{2} \left(\frac{ds}{d\tau} \right)^2 = \frac{1}{2} g_{ij} \frac{dq^i}{d\tau} \frac{dq^j}{d\tau}, \quad \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i = 2T = \left(\frac{ds}{d\tau} \right)^2.$$

Чтобы обеспечить фиксированное значение энергии вдоль траектории, параметр τ необходимо выбрать пропорциональным длине траектории $d\tau = ds/\sqrt{2E}$ (канонический параметр). Тогда укороченное действие принимает вид

$$\int_a^b d\tau \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i = \sqrt{2E} \int_a^b ds.$$

Это означает, что инерциальное движение происходит по экстремалиям римановой метрики g_{ij} . $\square$

Пример 21.7.2. Пусть движение материальной точки происходит по риманову многообразию $(\mathbb{M}, g)$ с интервалом $ds^2 = dq^i dq^j g_{ij}$ в потенциальном поле $U(q) \in \mathcal{C}^k(\mathbb{M})$. Функция Лагранжа и гамильтониан имеют вид

$$L = T - U, \quad H = T + U, \quad T = \frac{1}{2} \left(\frac{ds}{d\tau} \right)^2.$$

Чтобы обеспечить фиксированное значение энергии $H = E$, параметр τ вдоль траектории необходимо выбрать пропорциональным длине:

$$d\tau = \frac{ds}{\sqrt{2(E - U)}}.$$

Тогда укороченное действие примет вид

$$\int_{\gamma} d\tau \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i = \int_{\gamma} ds \sqrt{2(E - U)}.$$

Это значит, что движение материальной частицы по риманову многообразию $(\mathbb{M}, g)$ в потенциальном поле U происходит вдоль экстремалей метрики $\rho = g_{ij} 2(E - U)$, которая связана с исходной метрикой множителем $2(E - U)$. Эта метрика имеет особенность на границе $U(q) = E$. $\square$

Если начальная и конечная точки экстремали в рассмотренных примерах достаточно близки, то экстремум длины является минимумом. Это оправдывает название “принцип наименьшего действия”.

Пример 21.7.3 (Метод факторизации). Для нахождения траектории частицы иногда удобно выделить из гамильтониана $H(q, p)$ отличный от нуля множитель. Пусть $f(q, p) > 0$ – некоторая положительная функция на фазовом пространстве. Рассмотрим новый гамильтониан

$$\tilde{H} := f(q, p)(H(q, p) - E), \quad E = \text{const.}$$

Соответствующие уравнения движения имеют вид

$$\begin{aligned} \frac{dq}{ds} &= f \frac{\partial H}{\partial p} + \frac{\partial f}{\partial p}(H - E), \\ \frac{dp}{ds} &= -f \frac{\partial H}{\partial q} - \frac{\partial f}{\partial q}(H - E). \end{aligned} \quad (21.37)$$

Эти уравнения на инвариантной поверхности фазового пространства, определяемой уравнением $\tilde{H} = 0$ или $H = E$, эквивалентны гамильтоновым уравнениям (21.13) для исходного гамильтониана H . Уравнения (21.37) позволяют определить только форму траектории. Для определения эволюции частицы во времени достаточно проинтегрировать уравнение

$$\frac{dt}{ds} = f(q(s), p(s)),$$

где $q(s), p(s)$ – соответствующая траектория. □

21.8 Уравнение Гамильтона–Якоби

Рассмотрим *расширенное конфигурационное* пространство $\mathbb{R}^{n+1}$ с координатами $\{q, t\} = (q^1, \dots, q^n, t)$, которое получается из конфигурационного пространства $\mathbb{R}^n$ добавлением еще одного измерения – времени.

Определение. *Функцией действия* $S(q, t)$ называется интеграл

$$S(q, t) = \int_{\gamma} d\tau L, \quad (21.38)$$

где $L(q, \dot{q}, \tau)$ – функция Лагранжа, и интегрирование ведется вдоль экстремали $\gamma := \{q(\tau), t(\tau)\}$, с фиксированным началом q_0, t_0 и переменным концом q, t расширенного конфигурационного пространства. □

Это определение корректно по крайней мере в малой окрестности начальной точки q_0, t_0 . Точнее, функция действия (21.38) определена в некоторой окрестности $\mathbb{U}$ точки q_0, t_0 , если любую точку этой окрестности можно соединить с точкой q_0, t_0 экстремалью, целиком лежащей в $\mathbb{U}$, и эта экстремаль единственна. Если точка q, t лежит далеко от q_0, t_0 , то интеграл (21.38) может не определить функцию действия $S(q, t)$ по двум причинам. Во-первых, может существовать несколько экстремалей, соединяющих точки q_0, t_0 и q, t . Во-вторых, возможно, что точки q_0, t_0 и q, t вообще нельзя соединить экстремалью. С другой стороны, для любой точки q, t расширенного конфигурационного пространства можно так подобрать начальную точку q_0, t_0 , что функция действия будет определена в некоторой окрестности точки q, t . В дальнейшем мы будем считать, что точка q_0, t_0 фиксирована, и рассматривать те области расширенного конфигурационного пространства, где функция действия определена.

Аргументом функции действия является верхний предел интеграла (21.38), поэтому ее дифференциал равен

$$dS(q, t) = dq^i p_i - dt H, \quad (21.39)$$

где $p_i := \partial L / \partial \dot{q}^i$ и $H := \dot{q}^i p_i - L$ определяются в конечной точке кривой γ . Таким образом, дифференциал функции действия dS является интегральным инвариантом Пуанкаре–Картана (21.29).

Теорема 21.8.1. *Функция действия удовлетворяет уравнению Гамильтона–Якоби*

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}, t\right) = 0. \quad (21.40)$$

Доказательство. Достаточно заметить, что из выражения (21.39) следуют равенства

$$\frac{\partial S}{\partial t} = -H(q, p, t), \quad p_i = \frac{\partial S}{\partial q^i}. \quad \square$$

Очевидно, что любое решение уравнения Гамильтона–Якоби определено с точностью до аддитивной постоянной, которая соответствует произволу в выборе начальной точки q_0, t_0 экстремали γ в определении функции действия (21.38).

Уравнение Гамильтона–Якоби представляет собой уравнение в частных производных первого порядка. Покажем, что его интегрирование сводится к интегрированию обыкновенных дифференциальных уравнений Гамильтона. Поставим для уравнения Гамильтона–Якоби задачу Коши, т.е. будем искать решение уравнения (21.40) с начальным условием

$$S(q, t_0) = S_0(q). \quad (21.41)$$

Чтобы построить решение этой задачи, рассмотрим задачу Коши для уравнений Гамильтона (21.30) с начальными условиями:

$$q^i(t_0) = q_0^i, \quad p_i(t_0) = \left. \frac{\partial S_0}{\partial q^i} \right|_{q_0}.$$

Соответствующее этим начальным условиям решение изображается в расширенном конфигурационном пространстве кривой $q(t)$, которая является стационарной точкой действия $\int dt L$, где $L(q, \dot{q}, t)$ есть преобразование Лежандра по импульсу p от гамильтониана $H(q, p, t)$. Эта траектория называется *характеристикой* задачи Коши для уравнения Гамильтона–Якоби, выходящей из точки q_0, t_0 . При временах, достаточно близких к t_0 , значения $q(t), t$ можно принять за координаты в окрестности точки q_0, t_0 расширенного конфигурационного пространства. Построим теперь функцию действия

$$S(q, t) = S_0(q_0) + \int_{q_0, t_0}^{q, t} d\tau L(q, \dot{q}, \tau),$$

где интегрирование ведется вдоль экстремали, соединяющей точки q_0, t_0 и q, t . После этого проверяется, что эта функция действия удовлетворяет уравнению Гамильтона–Якоби (21.40) и начальному условию (21.41). Можно также доказать, что это решение задачи Коши для уравнения Гамильтона–Якоби единственно.

Если гамильтониан не зависит от времени явно, то зависимость функции действия от времени легко находится:

$$S(q, t) = -Et + W(q), \quad (21.42)$$

где $W(q)$ – *укороченная функция действия*, зависящая только от координат, и $E = \text{const}$. Из уравнения Гамильтона–Якоби (21.40) следует, что укороченная функция действия должна удовлетворять уравнению в частных производных

$$H\left(q, \frac{\partial W}{\partial q}\right) = E. \quad (21.43)$$

Отсюда следует, что постоянная E равна энергии системы. Это уравнение называется *укороченным уравнением Гамильтона–Якоби*.

Пример 21.8.1. Поясним определение функции действия на простом примере свободной точечной частицы единичной массы, движущейся в евклидовом пространстве $\mathbb{R}^N$. В этом случае лагранжиан и гамильтониан частицы имеют вид

$$L = \frac{1}{2} \dot{q}^i \dot{q}_i, \quad H = \frac{1}{2} p^i p_i,$$

где подъем и опускание индексов осуществляется с помощью символов Кронекера. Траекториями свободной частицы являются прямые линии и только они. Экстремаль $q^i(\tau)$, соединяющая точки q_0^i, t_0 и q^i, t , задается линейной функцией

$$q^i(\tau) = q_0^i + \frac{q^i - q_0^i}{t - t_0}(\tau - t_0),$$

описывающей равномерное прямолинейное движение частицы со скоростью $\dot{q}^i = (q^i - q_0^i)/(t - t_0)$. Нетрудно вычислить соответствующую функцию действия

$$S(q, t) = \int_{t_0}^t d\tau \frac{1}{2} \left(\frac{q - q_0}{t - t_0} \right)^2 = \frac{1}{2} \frac{(q - q_0)^2}{t - t_0},$$

где

$$(q - q_0)^2 := (q^i - q_0^i)(q_i - q_{0i}).$$

Функция действия определена для всех $q \in \mathbb{R}^N$, $t > t_0$ и удовлетворяет уравнению Гамильтона–Якоби:

$$\frac{\partial S}{\partial t} + \frac{1}{2} \frac{\partial S}{\partial q^i} \frac{\partial S}{\partial q_i} = 0.$$

Поскольку гамильтониан частицы не зависит от времени явно, то для любой траектории частицы энергия сохраняется:

$$E = \frac{1}{2} \left(\frac{q - q_0}{t - t_0} \right)^2 = \text{const}.$$

Отсюда можно выразить время через координаты

$$t - t_0 = \sqrt{\frac{(q - q_0)^2}{2E}}.$$

Поскольку $S = E(t - t_0)$, то укороченное действие равно

$$W = Et + S = 2E(t - t_0) + Et_0 = \sqrt{2E(q - q_0)^2} + Et_0.$$

Нетрудно проверить, что укороченное действие удовлетворяет укороченному уравнению Гамильтона–Якоби

$$E = \frac{1}{2} \frac{\partial W}{\partial q^i} \frac{\partial W}{\partial q_i}. \quad \square$$

Уравнение Гамильтона–Якоби предоставляет мощный метод решения задач механики точечных частиц. При обобщении этого метода на теорию поля возникают существенные осложнения. В настоящее время, насколько известно автору, это интересное обобщение не развито.

21.9 Принцип Гюйгенса

Многие понятия гамильтоновой механики возникли при перенесении на общие вариационные принципы весьма простых и наглядных понятий геометрической оптики. В настоящем разделе мы обсудим некоторые аспекты геометрической оптики.

Рассмотрим свет, распространяющийся в среде (конфигурационном пространстве), которую мы отождествим с евклидовым пространством $\mathbb{R}^3$ с декартовыми координатами q^i , $i = 1, 2, 3$. В общем случае скорость света зависит от точки $q \in \mathbb{R}^3$ (неоднородная среда) и направления луча света (неизотропная среда). Неизотропность среды можно описать, задав в каждой точке q поверхность в касательном пространстве $T_q(\mathbb{R}^3)$. Для этого отложим в начале координат касательного пространства $T_q(\mathbb{R}^3)$ вектор скорости распространения света в точке q . Эта поверхность называется *индикатрисой*. Согласно *принципу Ферма*, свет распространяется в среде из точки q_0 в точку q за кратчайшее время.

Если среда изотропна, то принцип Ферма приобретает простую математическую форму. Пусть $q(t)$ – траектория луча. Тогда для скорости света $v^i := dq^i/dt$ справедливо равенство

$$v^2 dt^2 = ds^2,$$

где $v^2 := v^i v^j \delta_{ij}$ и $ds^2 := dq^i dq^j \delta_{ij}$. Отсюда следует, что время прохождения луча между точками q_0 и q вдоль траектории $q(t)$ равно

$$t = \int_{q_0}^q \frac{ds}{v},$$

где $v := \sqrt{v^2}$ и $ds := \sqrt{ds^2}$. Для изотропной среды показатель преломления n равен отношению скорости света c в вакууме к скорости света v в среде, $n = c/v$. Поэтому время распространения луча света дается интегралом

$$t = \frac{1}{c} \int_{q_0}^q ds n. \quad (21.44)$$

Таким образом, принцип Ферма для распространения лучей света в изотропной среде сводится к вариационной задаче для действия (21.44), в котором показатель преломления $n(q)$ является заданной функцией на конфигурационном пространстве.

Распространение света в среде можно также описать на языке волновых фронтов. Допустим, что в момент времени $t = 0$ в точке q_0 произошла вспышка света. Рассмотрим множество точек, до которых свет дойдет за время, меньшее или равное $t > 0$. Граница этого множества $\Phi_0(t)$, которая представляет собой некоторую поверхность в $\mathbb{R}^3$, называется *волновым фронтом* точки q_0 через время t и состоит из точек, до которых свет дойдет за время t и не может прийти быстрее. При этом мы предполагаем, что среда такова, что волновые фронты представляют собой гладкие поверхности в $\mathbb{R}^3$. Тогда между волновыми фронтами для разных моментов времени t имеется замечательное соотношение.

Теорема 21.9.1 (Принцип Гюйгенса). *Рассмотрим волновой фронт $\Phi_0(t)$ точки q_0 в момент времени t . Для каждой точки этого фронта $q \in \Phi_0(t)$ построим волновой фронт $\Phi_q(s)$ через время $s > 0$. Тогда волновой фронт $\Phi_0(t + s)$ точки q_0 через время $t + s$ будет огибающей поверхностью всех фронтов $\Phi_q(s)$ для $q \in \Phi_0(t)$.*

Доказательство. Пусть $q_{t+s} \in \Phi_0(t+s)$. Тогда существует путь из начальной точки q_0 в точку q_{t+s} , по которому свет распространяется за время $t+s$ и нет более короткого. Рассмотрим точку q_t на этом пути, до которой свет идет время t . Никакого более короткого пути из q_0 в q_t не существует, иначе путь не был бы кратчайшим. Поэтому точка q_t лежит на фронте $\Phi_0(t)$. Точно так же путь из q_t в q_{t+s} свет про-

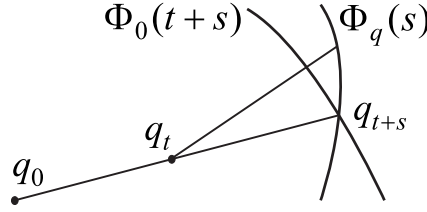


Рис. 21.2: Невозможность пересечения фронтов

ходит за время s и нет более короткого пути между этими точками. Поэтому точка q_{t+s} лежит на фронте $\Phi_q(s)$ точки $q(t)$ через время s . Осталось показать, что фронты $\Phi_q(s)$ и $\Phi_0(t+s)$ в точке $q(t+s)$ касаются. Действительно, предположим, что фронты пересекаются, как показано на рис.21.2. Тогда в некоторые точки фронта $\Phi_0(t+s)$ из точки q_t можно было бы добраться за время, меньшее s . В свою очередь это значит, что в эти точки из q_0 можно было бы добраться за время, меньшее $t+s$, что противоречит определению фронта $\Phi_0(t+s)$. $\square$

Разумеется, точку q_0 в принципе Гюйгенса можно заменить на кривую, поверхность или вообще на произвольное замкнутое множество, а трехмерное евклидово пространство на произвольное дифференцируемое многообразие $\mathbb{M}$. При этом свет можно заменить на распространение любого возмущения, передающегося локально, т.е. вдоль линии $\gamma \in \mathbb{M}$, на которой некоторый функционал принимает наименьшее значение.

Принцип Гюйгенса приводит к двум способам описания распространения возмущений. Во-первых, можно следить за *лучами*, т.е. кратчайшими путями, вдоль которых распространяется свет (корпускулярная точка зрения). В этом случае распространение света задается вектором скорости $\dot{q}$ в каждой точке конфигурационного пространства, т.е. некоторой точкой на индикатрисе. Во-вторых, мы можем следить за распространением волнового фронта (волновая точка зрения).

Определим скорость движения p волнового фронта следующим образом. Для каждой точки q_0 определим функцию $S_0(q)$, как наименьшее время распространения света из точки q_0 в точку q . По построению, поверхность уровня функции $S_0(q)$ является волновым фронтом:

$$\Phi_0(t) := \{q \in \mathbb{R}^3 : S_0(q) = t\}.$$

Определение. Ковектор с компонентами

$$p_i := \frac{\partial S_0(q)}{\partial q^i}$$

называется *ковектором нормальной медлительности фронта*. $\square$

Такое название связано с тем, что чем больше градиент функции $S_0(q)$, тем медленнее движется фронт. Действительно, в линейном приближении справедливо равенство

$$S_0(q + dq) = S_0(q) + p_i dq^i = t + dt.$$

Отсюда следует, что чем больше расстояние dq , пройденное светом за время dt , тем меньше ковектор нормальной медлительности фронта p . Векторы X , касательные к волновому фронту, определяются, как ортогональное дополнение к ковектору нормальной медлительности фронта:

$$X^i p_i = 0.$$

Если на конфигурационном пространстве определена риманова метрика g_{ij} , то можно определить вектор нормальной медлительности фронта с компонентами $p^i := g^{ij} p_j$. По построению, этот вектор ортогонален волновому фронту.

Если среда анизотропна, то направление лучей света $\dot{q}$ и направление движения фронта p не совпадают. Однако они связаны между собой простым соотношением, которое легко выводится из принципа Гюйгенса. Напомним, что оптические свойства среды в точке q характеризуются поверхностью в касательном пространстве $\mathbb{T}_q(\mathbb{R}^3)$ – индикатрисой.

Определение. Направление гиперплоскости (т.е. вектор, перпендикулярный к данной гиперплоскости), касающейся индикатрисы в точке $\dot{q}$, называется *сопряженным* к направлению $\dot{q}$. $\square$

Предложение 21.9.1. *Направление p волнового фронта $\Phi_0(t)$ в точке q_t сопряжено направлению луча $\dot{q}$ в данной точке.*

Доказательство. Рассмотрим точки $q_{t-\epsilon}$, лежащие на луче с началом в точке q_0 и концом в q_t , при малых $\epsilon > 0$. Фронт волны $\Phi_{q_{t-\epsilon}}(\epsilon)$ в точке $q_{t-\epsilon}$ в момент времени ϵ можно представить в виде

$$\Phi_{q_{t-\epsilon}}(\epsilon) = \{q_{t-\epsilon}^i + \dot{q}_t^i \epsilon\},$$

где вектор скорости $\dot{q}$ пробегает всю индикатрису, с точностью порядка $o(\epsilon)$. Из принципа Гюйгенса следует, что фронт $\Phi_{q_{t-\epsilon}}(\epsilon)$ касается фронта $\Phi_0(t)$ в точке q_t . В пределе $\epsilon \rightarrow 0$ получаем утверждение предложения. $\square$

Теперь сравним геометрическую оптику с построениями предыдущего раздела. Лучом является траектория частицы $q(t)$ в конфигурационном пространстве $\mathbb{R}^3$. Ковектор нормальной медлительности фронта – это импульс частицы p . Принцип Ферма соответствует принципу наименьшего действия. Функция $S_0(q)$ есть ни что иное, как функция действия (21.38), в которой время t соответствует концу траектории $q = q(t)$. Поверхности уровня функции действия $S(q, t)$ соответствуют волновым фронтам.

21.10 Переменные действие-угол

Знание интегралов движения позволяет упростить задачу. При этом интерес представляют функционально независимые интегралы движения. Из определения функциональной независимости следует, что на фазовом пространстве размерности $2N$ может существовать не более $2N$ функционально независимых интегралов движения.

Чтобы проинтегрировать систему из $2N$ обыкновенных дифференциальных уравнений первого порядка, достаточно знать $2N$ независимых первых интегралов. Оказывается, что если задана каноническая система уравнений движения, то ситуация существенно проще: достаточно знать только N независимых первых интегралов. Это происходит потому что каждый интеграл движения позволяет понизить порядок системы уравнений не на одну, а на две единицы.

Определение. Функция $F \in C^1(\mathbb{M})$ на пуассоновом многообразии $\mathbb{M}$ называется *первым интегралом* механической системы с гамильтонианом H , если ее скобка Пуассона с гамильтонианом равна нулю, $[F, H] = 0$. Две функции $F_1, F_2 \in C^1(\mathbb{M})$ находятся в *инволюции*, если их скобка Пуассона равна нулю, $[F_1, F_2] = 0$. $\square$

Заметим, что условие $[F, H] = 0$ эквивалентно условию $\dot{f} = 0$, т.к. $\dot{f} = [F, H]$

Теорема 21.10.1 (Лиувилль). *Предположим, что на симплектическом многообразии $\mathbb{M}$ размерности $2N$ заданы N функционально независимых дифференцируемых функций $\{F_i\}$, $i = 1, \dots, N$, которые находятся в инволюции, $[F_i, F_j] = 0$. Рассмотрим N -мерное подмногообразие $\mathbb{M}_f \hookrightarrow \mathbb{M}$, которое является множеством уровня функций F_i :*

$$\mathbb{M}_f := \{x \in \mathbb{M} : F_i = f_i = \text{const}, i = 1, \dots, N\}.$$

Тогда справедливы следующие утверждения:

- 1) Подмногообразие $\mathbb{M}_f$ инвариантно относительно фазового потока с функцией Гамильтона $H = F_i$ при любом фиксированном i .
- 2) Если подмногообразие $\mathbb{M}_f$ компактно и связно, то оно диффеоморфно N -мерному тору $\mathbb{T}^N$.
- 3) Фазовый поток с функцией Гамильтона H определяет на $\mathbb{M}_f$ условно-периодическое движение, т.е. в угловых координатах $\{\varphi^i, \text{mod } 2\pi\}$ на торе уравнения движения имеют вид

$$\frac{d\varphi^i}{dt} = \omega^i, \quad \omega^i = \omega^i(f) = \text{const}, \quad \forall i. \quad (21.45)$$

- 4) Канонические уравнения движения с функцией Гамильтона H интегрируются в квадратурах.

Доказательство. См., например, [143], глава 10, §49, раздел А.

Прокомментируем первые два утверждения теоремы.

Поскольку фазовое пространство $\mathbb{M}$ является симплектическим и, следовательно, пуассоновым многообразием, то каждой функции F_i ставится в соответствие векторное поле X_i по правилу (19.29). При этом инволютивность первых интегралов движения влечет за собой инволютивность распределения векторных полей $\{X_i\}$. Согласно теореме Фробениуса у этого распределения существует интегральное подмногообразие, которым является $\mathbb{M}_f \hookrightarrow \mathbb{M}$. Отсюда следует инвариантность $\mathbb{M}_f$ относительно фазового потока для любой функции F_i .

Тор возникает из-за того, что векторные поля X_i , которые являются генераторами группы однопараметрических преобразований $\mathbb{M}_f$, не просто находятся в инволюции, а коммутируют между собой. $\square$

Замечание. Если гамильтониан механической системы H не зависит от времени, то его можно выбрать в качестве одного из первых интегралов F_i . $\square$

Определение. Гамильтонова механическая система, называется *интегрируемой*, если она имеет N или более функционально независимых интегралов движения. $\square$

Если фазовое пространство механической системы таково, что выполнены условия теоремы Лиувилля, то на фазовом пространстве $\mathbb{M}$ существует выделенная система координат (действие-угол), в которой уравнения движения выглядят особенно просто. Для определенности будем считать, что гамильтониан системы совпадает с первым интегралом движения: $H = F_1$. В теореме Лиувилля утверждается, что

подмногообразие $\mathbb{M}_f \hookrightarrow \mathbb{M}$ является N -мерным тором, инвариантно относительно фазового потока и на нем существуют угловые координаты φ^i , для которых уравнения движения имеют вид (21.45). Общее решение этих уравнений имеет простой вид

$$\varphi = \varphi_0 + \omega t, \quad \varphi_0 = \text{const.}$$

Поскольку интегралы движения функционально независимы, то в некоторой окрестности подмногообразия $\mathbb{M}_f$ в качестве координат можно выбрать совокупность функций $\{F, \varphi\}$. В этой системе координат уравнения движения принимают простой вид

$$\frac{dF}{dt} = 0, \quad \frac{d\varphi}{dt} = \omega(F),$$

и легко интегрируются:

$$F(t) = F(0), \quad \varphi(t) = \varphi_0 + \omega(F(0))t.$$

В общем случае в координатах $\{F, \varphi\}$ симплектическая форма не будет иметь канонического вида. Однако существует такой набор функционально независимых функций $\{I_i = I_i(F)\}$, $i = 1, \dots, N$, что переменные $\{I, \varphi\}$ образуют такую систему координат в окрестности $\mathbb{M}_f$, в которой симплектическая форма имеет канонический вид

$$\varpi = dI_i \wedge d\varphi^i. \quad (21.46)$$

Определение. Координаты $\{\varphi^1, \dots, \varphi^N, I_1, \dots, I_N\}$ в некоторой окрестности подмногообразия фазового пространства $\mathbb{M}_f \subset \mathbb{R}^{2N}$, в которых канонические уравнения движения имеют вид

$$\frac{dI}{dt} = 0, \quad \frac{d\varphi}{dt} = \omega(I),$$

и симплектическая форма является канонической (21.46) называются *переменными действие-угол*. $\square$

Переменные действие-угол являются координатами на фазовом пространстве. При этом угловые переменные удобно рассматривать как обобщенные координаты, а переменные действия – как сопряженные импульсы. Из определения сразу следует, что преобразование координат $q, p \mapsto \varphi, I$, если оно существует, является каноническим.

Переменные I так же как и функции F являются первыми интегралами движения. В переменных действие-угол гамильтониан механической системы имеет вид $H = F_1 = H(I)$ и не зависит от угловых переменных φ . То есть каждая из угловых координат φ является циклической.

Теорема 21.10.2. Координаты действие-угол $\{\varphi, I\}$ существуют в некоторой окрестности подмногообразия $\mathbb{M}_f \hookrightarrow \mathbb{R}^{2N}$.

Доказательство. Рассмотрим множество торов $\mathbb{M}_f$, соответствующих различным значениям интегралов движения F_i . Пусть γ_i , $i = 1, \dots, N$, – базисные одномерные циклы (окружности) на торах $\mathbb{M}_f$, т.е. приращение координаты φ^i на цикле γ_j равно 2π , если $i = j$ и 0, если $i \neq j$. Каждое подмногообразие $\mathbb{M}_f \hookrightarrow \mathbb{R}^{2N}$ является нулевым. Действительно, векторные поля X_i , соответствующие интегралам движения F_i , образуют базис касательных пространств $\mathbb{T}_x(\mathbb{M}_f)$ для всех $x \in \mathbb{M}_f$ и попарно коммутируют. Поэтому значение канонической симплектической формы $\varpi = dp_i \wedge dq^i$ на двух произвольных базисных полях равно нулю:

$$\varpi(X_i, X_j) = \varpi_{kl} X_i^k X_j^l = \varpi^{-1kl} \partial_k F_i \partial_l F_j = [F_i, F_j] = 0.$$

Следовательно, подмногообразие $\mathbb{M}_f$ является нулевым. Отсюда вытекает, что 1-форма $dq p$ замкнута на $\mathbb{M}_f$, т.е. ее внешняя производная $dp \wedge dq$ обращается в нуль на подмногообразии $\mathbb{M}_f$. Положим

$$I_i(F) := \oint_{\gamma_i} dq p.$$

Эти функции зависят от значений интегралов движения F_i , определяющих тор $\mathbb{M}_f$, но не зависят от выбора базисных циклов γ_i , т.к. форма $dq p$ замкнута. Это означает, что функции $I_i(F)$ не зависят от координат φ на торе.

Теперь совершим два канонических преобразования. Поскольку скобка Пуассона на интегралах движения F_i между собой равна нулю, то их можно выбрать в качестве новых импульсов: $P_i = F_i$. Совершим первое каноническое преобразование $q, p \mapsto Q, F$ с производящей функцией $S_4(p, F)$ (см. следующий раздел), зависящей от новых и старых импульсов F и p . Тогда старые и новые координаты определены соотношениями (21.75):

$$q = -\frac{\partial S_4}{\partial p}, \quad Q = \frac{\partial S_4}{\partial F}.$$

Подмногообразие $\mathbb{M}_f$ в новой системе координат задано соотношениями $P = 0$. Следовательно, новые координаты Q образуют некоторую систему координат на $\mathbb{M}_f$.

Зафиксируем точку $Q_0 \in \mathbb{M}_f$. В некоторой ее окрестности можно выбрать систему координат Q, I , т.к. функции $I = I(P)$ зависят только от импульсов $P = F$. При этом импульсы P можно выразить, как функции от Q и I , т.е. $P = P(Q, I)$. Поскольку преобразование $q, p \mapsto Q, P$ каноническое, то 1-форма dQP замкнута. Поэтому в односвязной окрестности $\mathbb{U}$ точки $Q_0 \in \mathbb{U} \subset \mathbb{M}_f$ определена функция

$$S_2(Q, I) = \int_{Q_0}^Q dQ' P(Q', I).$$

Этот интеграл не зависит от кривой, целиком лежащей в $\mathbb{U}$ и соединяющей точки Q_0 и Q . Следовательно, функцию $S_2(Q, I)$ можно выбрать в качестве производящей функции второго канонического преобразования $Q, P \mapsto \varphi, I$, зависящей от старых координат Q и новых импульсов I . Формулы преобразования имеют вид (21.73):

$$P = \frac{\partial S_2}{\partial Q}, \quad \varphi = \frac{\partial S_2}{\partial I}.$$

Таким образом, преобразование координат $q, p \mapsto \varphi, I$ является произведением двух канонических преобразований $q, p \mapsto Q, P$ и $Q, P \mapsto \varphi, I$ и, следовательно, само является каноническим. По построению, оно определено в окрестности тора $\mathbb{M}_f$. $\square$

Замечание. В доказательстве теоремы переход от координат q, p к переменным действие-угол φ, I содержит только алгебраические операции и интегрирование. Это доказывает утверждение 4 теоремы Лиувилля 21.10.1. $\square$

Переменные действие-угол определены неоднозначно. Очевидно, что координаты можно сдвигать:

$$I' = I + \text{const}, \quad \varphi' = \varphi + \text{const}(I).$$

21.11 Канонические преобразования

Наиболее мощный и гибкий метод нахождения точных решений уравнений движения дают канонические преобразования. По-прежнему, обозначим конфигурационное и фазовое пространства, соответственно, через $\mathbb{R}^N$, и $\mathbb{T}^*(\mathbb{R}^N) \approx \mathbb{R}^{2N}$. Будем считать, что на фазовом пространстве заданы координаты $\{x^\alpha\} = \{q^i, p_i\} := (q^1, \dots, q^N, p_1, \dots, p_N)$ и каноническая пуассонова структура (19.33).

Определение. Достаточно гладкое биективное отображение g фазового пространства (диффеоморфизм) $\mathbb{T}^*(\mathbb{R}^N)$ на себя называется каноническим, если оно сохраняет каноническую симплектическую форму:

$$g^* \varpi = \varpi,$$

где $\varpi := dp \wedge dq := dp_i \wedge dq^i$. □

В координатах это определение записывается следующим образом. Отображение g фазового пространства задается $2N$ функциями, которые в общем случае могут зависеть от времени,

$$g : \mathbb{T}^*(\mathbb{R}^N) \ni (q^i, p_i) \mapsto (Q^i(q, p, t), P_i(q, p, t)) \in \mathbb{T}^*(\mathbb{R}^N), \quad (21.47)$$

где время t входит в качестве параметра. Поскольку ранг 2-формы ϖ равен $2N$, то канонические преобразования задают преобразование координат (диффеоморфизм) фазового пространства с отличным от нуля якобианом:

$$\frac{\partial(Q, P)}{\partial(q, p)} \neq 0.$$

Условие сохранения канонической симплектической формы ϖ принимает вид

$$dp_i \wedge dq^i = dP_i \wedge dQ^i. \quad (21.48)$$

Это равенство приводит к дифференциальным уравнениям на функции $Q^i(q, p, t)$, $P_i(q, p, t)$, которые мы рассмотрим несколько позже.

Определение канонических преобразований можно записать в эквивалентном интегральном виде

$$\int_S dp \wedge dq = \int_{gS} dp \wedge dq, \quad (21.49)$$

где интегрирование ведется по произвольной двумерной поверхности $S \subset \mathbb{T}^*(\mathbb{R}^N)$ с кусочно гладкой границей. Поскольку кокасательное расслоение $\mathbb{T}^*(\mathbb{R}^N)$ односвязно, то к интегралу (21.49) можно применить формулу Стокса:

$$\int_S dp \wedge dq = \oint_{\partial S} dq p,$$

где мы предположили, что граница ∂S такова, что dq на ней может обратиться в нуль только в изолированных точках. Поэтому для односвязных фазовых пространств интегральное определение канонического преобразования (21.49) можно переписать в виде

$$\oint_{\gamma} dq p = \oint_{g\gamma} dq p, \quad (21.50)$$

где γ – произвольная кусочно гладкая замкнутая кривая в фазовом пространстве.

В общем случае не односвязных многообразий из равенства (21.50) следует (21.49), но не наоборот. То есть сохранение 1-формы $dq p$ является достаточным условием того, что преобразование будет каноническим.

Пример 21.11.1. Каноническая пуассонова структура определена на произвольном кокасательном расслоении $T^*(M)$. При этом вид структурных функций не зависит от выбора координат на конфигурационном пространстве M . Поэтому преобразование $q, p \mapsto Q(q), P(q, p)$, где Q зависит только от q , является каноническим (см., пример 21.3.1). $\square$

Пример 21.11.2. Пусть фазовое пространство топологически тривиально, $T^*(M) \approx \mathbb{R}^{2N}$. Поскольку линейное преобразование координат фазового пространства, порождаемое симплектической группой $SP(N, \mathbb{R})$ (см. раздел 19.1), сохраняет каноническую симплектическую форму, то оно является каноническим. $\square$

На практике наиболее полезными и содержательными являются нелинейные канонические преобразования, связанные со спецификой задачи: симметриями и наличием интегралов движения.

Пример 21.11.3. Фазовый поток является каноническим отображением в силу следствия из теоремы 21.6.2. $\square$

Пример 21.11.4 (Теория возмущений). Допустим, что гамильтониан механической системы состоит из двух слагаемых,

$$H(q, p) = H_0(q, p) + H_1(q, p). \quad (21.51)$$

При этом гамильтониан H_0 настолько прост, что для него известно общее решение

$$q = q(t; q_0, p_0), \quad p = p(t; q_0, p_0), \quad (21.52)$$

задачи Коши с начальными условиями:

$$q(0; q_0, p_0) = q_0, \quad p(0; q_0, p_0) = p_0.$$

Как было отмечено, фазовый поток определяет каноническое преобразование, которое задается функциями (21.52). Обратное преобразование $q, p \mapsto q_0, p_0$ также является каноническим. При этом преобразовании гамильтониан H_0 обращается в константу и уравнения движения исходной гамильтоновой системы упрощаются:

$$\dot{q}_0 = \frac{\partial \tilde{H}_1}{\partial p_0}, \quad \dot{p}_0 = -\frac{\partial \tilde{H}_1}{\partial q_0},$$

где

$$\tilde{H}_1(q_0, p_0, t) = H_1(q(t; q_0, p_0), p(t; q_0, p_0)). \quad (21.53)$$

Обозначим через $\tilde{q}_0(t; q_0, p_0)$ и $\tilde{p}_0(t; q_0, p_0)$ решение задачи Коши для полученной системы уравнений движения с начальными данными:

$$\tilde{q}_0(0; q_0, p_0) = q_0, \quad \tilde{p}_0(0; q_0, p_0) = p_0.$$

Это решение обладает замечательным свойством: для невозмущенной системы оно постоянно и совпадает с начальными данными. Таким образом, задача Коши для гамильтоновой системы (21.51) сведена к задаче Коши для возмущения, которое определяется гамильтонианом (21.53). Формулы для решения исходной задачи в старых координатах q, p получаются после подстановки решения $\tilde{q}_0, \tilde{p}_0$ в общее решение (21.52) невозмущенной задачи. $\square$

Из определения канонических преобразований фазового пространства следует, что фазовое пространство вместе с каноническими отображениями образуют группу преобразований. Эта группа является подгруппой группы общих преобразований координат фазового пространства $\text{diff } T^*(M)$. В общем случае канонические преобразования $q^i, p_i \mapsto Q^i, P_i$ не сохраняют структуру кокасательного расслоения на фазовом пространстве, соответствующую координатам q^i, p_i , т.к. могут перемешивать канонические координаты и импульсы.

По определению, каноническое преобразование сохраняет 2-форму ϖ (21.27) и, следовательно, все ее внешние степени. Отсюда вытекает

Теорема 21.11.1. *Канонические преобразования сохраняют интегральные инварианты $\varpi^2, \dots, \varpi^N$.*

Поскольку $2N$ -форма ϖ^N существует и невырождена, то на фазовом пространстве существует также форма объема v , которая ей пропорциональна (19.12). Отсюда вытекает

Следствие. Канонические преобразования сохраняют форму объема v фазового пространства. Другими словами, якобиан канонического преобразования координат равен единице,

$$\det \frac{\partial(Q, P)}{\partial(q, p)} = 1.$$

Это значит, что объем

$$V = \int_{\mathbb{U}} v$$

произвольной области $\mathbb{U}$ фазового пространства сохраняется при каноническом преобразовании: $gV = V$.

Поскольку фазовый поток для произвольной гамильтоновой системы является каноническим преобразованием, то он также имеет интегральные инварианты $\varpi, \varpi^2, \dots, \varpi^N$. Последний из этих инвариантов есть фазовый объем. Отсюда следует

Теорема 21.11.2 (Лиувилль). *Фазовый поток сохраняет объем произвольной области фазового пространства.*

Верно также обратное утверждение. А именно, пусть задан функционал

$$I_k = \int_{\mathbb{U}} d^{2k}x F(t, x), \quad k = 1, \dots, N,$$

где $\mathbb{U} \subset \mathbb{R}^{2N}$ – произвольное четномерное подмногообразие фазового пространства размерности $\dim \mathbb{U} = 2k$ с кусочно гладкой границей и $F(t, x)$ – некоторая функция времени и канонических переменных. Если этот функционал инвариантен относительно фазового потока с произвольным гамильтонианом, то он называется *интегральным инвариантом*.

Теорема 21.11.3. *Любой интегральный инвариант I_k фазового потока с произвольным гамильтонианом имеет вид*

$$I_k = c_k \int_{\mathbb{U}} \varpi^k,$$

где c_k – некоторые постоянные, для всех $k = 1, \dots, N$.

Доказательство. Дано сравнительно недавно [144]. $\square$

Важность канонических преобразований заключается в том, что они позволяют менять координаты в фазовом пространстве не меняя вида канонических уравнений движения. Чтобы показать это, рассмотрим преобразование координат в расширенном фазовом пространстве $\mathbb{R}^{2N+1}$, с координатами q, p, t . Как было показано в разделе 21.6, фазовый поток взаимно однозначно связан с характеристиками интегрального инварианта Пуанкаре–Картана в расширенном фазовом пространстве. Поскольку характеристики 1-формы являются инвариантным объектом, то замену координат $q, p, t \mapsto Q, P, T$ удобно провести в расширенном фазовом пространстве.

Теорема 21.11.4. Пусть Q, P, T – новая система координат в расширенном фазовом пространстве $\mathbb{R}^{2N+1}$ и $K(Q, P, T)$, $S(Q, P, T)$ – такие функции, что выполнено условие

$$dq p - dt H = dQ P - dT K + dS. \quad (21.54)$$

Тогда траектории фазового потока (21.30) изображаются в координатах Q, P, T интегральными кривыми канонических уравнений

$$\frac{dQ^i}{dT} = \frac{\partial K}{\partial P_i}, \quad \frac{dP_i}{dT} = -\frac{\partial K}{\partial Q^i}. \quad (21.55)$$

Доказательство. Поскольку $ddS = 0$, то 1-форма dS не влияет на характеристики формы (21.54). Это значит, что фазовый поток определяется также характеристиками формы $dQ P - dT K$. Уравнения (21.55) следуют из теоремы 21.6.1. $\square$

Каноническим преобразованиям координат в фазовом пространстве $q, p \mapsto Q, P$ соответствует преобразование координат $q, p, t \mapsto Q, P, t$ в расширенном фазовом пространстве, которое не меняет времени.

Теорема 21.11.5. В новых координатах Q, P канонические уравнения движения (21.30) имеют канонический вид (21.55) со старой функцией Гамильтона, выраженной через новые координаты $K(Q, P, t) = H(q(Q, P), p(Q, P), t)$.

Доказательство. Рассмотрим 1-форму $dq p - dQ P$ в фазовом пространстве. Тогда для любой замкнутой кривой γ справедливо равенство

$$\oint_{\gamma} (dq p - dQ P) = 0,$$

поскольку преобразование является каноническим. Это равенство можно записать в дифференциальной форме $dq p - dQ P = dS$, где $S(Q, P)$ – некоторая функция на фазовом пространстве. Следовательно, в расширенном фазовом пространстве имеем равенство

$$dq p - dt H = dQ P - dt H + dS, \quad (21.56)$$

и справедлива предыдущая теорема. $\square$

Таким образом, мы показали, что при канонических преобразованиях вид гамильтоновых уравнений движения не меняется. Часто это свойство принимается за определение канонических преобразований. Следующий пример показывает, что такое определение является более общим, чем принятое в настоящей монографии.

Пример 21.11.5. Преобразование координат, которое заключается в растяжке импульсов:

$$\varphi_c : \mathbb{R}^{2N} \ni (q, p) \mapsto (Q = q, P = cp) \in \mathbb{R}^{2N}, \quad (21.57)$$

где $c \neq 0$ – некоторая постоянная, сохраняет гамильтонову форму уравнений движения с гамильтонианом $K = cH$. Тем не менее оно не является каноническим, т.к. каноническая форма

$$\varpi = dp \wedge dq = \frac{1}{c} dP \wedge dQ$$

меняет свой вид.

В более общем случае преобразование $Q = bq, P = cp$, где b и c – произвольные отличные от нуля постоянные, также сохраняет вид гамильтоновых уравнений движения. При этом новый гамильтониан равен $K = bcH$. $\square$

По сути дела этот пример описывает все отличия в разных определениях канонического преобразования.

Предложение 21.11.1. Пусть при преобразовании координат $\varphi : q, p \mapsto Q, P$ фазового пространства $\mathbb{R}^{2N}$ гамильтоновы уравнения движения (21.13) сохраняют свой вид:

$$\dot{Q}^i = \frac{\partial K}{\partial P_i}, \quad \dot{P}_i = -\frac{\partial K}{\partial Q^i},$$

где $K(Q, P)$ – некоторый новый гамильтониан. Тогда

$$\oint_{\varphi(\gamma)} dQP = c \oint_{\gamma} dq p, \quad (21.58)$$

где $c \neq 0$ – некоторая постоянная.

Доказательство. Рассмотрим расширенное фазовое пространство $\mathbb{R}^{2N+1}$. Пусть γ – замкнутый контур в фазовом пространстве, соответствующий фиксированному моменту времени $t = \text{const}$. Тогда в силу следствия из теоремы 21.6.2 интегралы

$$\oint_{\gamma} dq p \quad \text{и} \quad \oint_{\varphi(\gamma)} dQP$$

являются интегральными инвариантами фазовых потоков. По теореме Стокса интеграл по контуру γ преобразуется в интеграл по поверхности, натянутой на данный контур. Используя теорему 21.11.3, заключаем, что

$$\oint_{\varphi(\gamma)} dQP = c \oint_{\gamma} dq p,$$

где $c \neq 0$ – некоторая постоянная. $\square$

Ясно, что множество преобразований, сохраняющих вид гамильтоновых уравнений, образует группу преобразований фазового пространства.

Следствие. Любое преобразование координат $q, p \mapsto Q, P$ фазового пространства, которое сохраняет форму гамильтоновых уравнений движения, является композицией преобразования φ_c (21.57) и некоторого канонического преобразования. $\square$

Доказательство. При преобразовании координат $q, p \mapsto Q, P$ относительный интегральный инвариант Пуанкаре–Картана умножается на некоторую постоянную c (21.58). Представим это преобразование в виде двух последовательных преобразования координат:

$$\varphi_c : q, p \mapsto \tilde{q}, \tilde{p} \quad \text{и} \quad g : \tilde{q}, \tilde{p} \mapsto Q, P.$$

Поскольку при первом преобразовании инвариант умножается на ту же постоянную c , то при втором преобразовании координат он сохраняется,

$$g : \oint_{\varphi_c(\gamma)} d\tilde{q} \tilde{p} = \oint_{g \circ \varphi_c(\gamma)} dQ P.$$

Ввиду произвольности контура γ это значит, что преобразование g является каноническим. $\square$

Поскольку преобразование φ_c (21.57) является простой растяжкой координат, то канонические преобразования составляют нетривиальную и наиболее содержательную часть преобразований координат, сохраняющих вид гамильтоновых уравнений движения.

Сформулируем три критерия для канонических преобразований. С этой целью введем новое понятие. Обозначим координаты фазового пространства через $\{x^\alpha\} = \{q^i, p_i\}$, $\alpha = 1, \dots, 2N$.

Определение. Пусть на фазовом пространстве $\mathbb{R}^{2N}$ с координатами $\{x^\alpha\} = \{q^i, p_i\}$ и канонической симплектической формой ϖ задано преобразование координат $\{x^\alpha\} \mapsto \{y^\alpha\} = \{Q^i(x), P_i(x)\}$. Назовем *скобкой Лагранжа* двух новых координат y^α и y^β для заданного набора функций $q^i(y^\alpha, y^\beta)$, $p_i(y^\alpha, y^\beta)$, определяющих преобразование координат, следующее выражение

$$[y^\alpha, y^\beta]_L := \varpi_{\delta\gamma} \frac{\partial x^\gamma}{\partial y^\alpha} \frac{\partial x^\delta}{\partial y^\beta} = \frac{\partial q^i}{\partial y^\alpha} \frac{\partial p_i}{\partial y^\beta} - \frac{\partial q^i}{\partial y^\beta} \frac{\partial p_i}{\partial y^\alpha}. \quad \square \quad (21.59)$$

Скобка Лагранжа антисимметрична и определена для любой пары новых координат y^α и y^β . При этом мы не требуем, чтобы преобразование координат $x \mapsto y(x)$ было каноническим.

Замечание. Если на фазовом пространстве просто заданы две дифференцируемые функции $F, G \in C^1(\mathbb{R}^{2N})$, то скобка Лагранжа для них неопределена, т.к. не определены частные производные $\partial x^\alpha / \partial F$ и $\partial x^\alpha / \partial G$. $\square$

Теорема 21.11.6. Для того, чтобы преобразование координат фазового пространства $q, p \mapsto Q, P$ было каноническим, необходимо и достаточно выполнения следующих условий для скобок Лагранжа новых координат:

$$[Q^i, Q^j]_L = 0, \quad [P_i, P_j]_L = 0, \quad [Q^i, P_j]_L = \delta_j^i. \quad (21.60)$$

Доказательство. Подставляя в определение канонических преобразований (21.48) функции $q^i(Q, P)$, $p_i(Q, P)$, определяющие канонические преобразования, получим равенство

$$\begin{aligned} dP_i \wedge dQ^i &= \frac{1}{2} dQ^l \wedge dQ^k \left(\frac{\partial q^i}{\partial Q^k} \frac{\partial p_i}{\partial Q^l} - \frac{\partial q^i}{\partial Q^l} \frac{\partial p_i}{\partial Q^k} \right) + \\ &+ dP_l \wedge dQ^k \left(\frac{\partial q^i}{\partial Q^k} \frac{\partial p_i}{\partial P_l} - \frac{\partial q^i}{\partial P_l} \frac{\partial p_i}{\partial Q^k} \right) + \\ &+ \frac{1}{2} dP_l \wedge dP_k \left(\frac{\partial q^i}{\partial P_k} \frac{\partial p_i}{\partial P_l} - \frac{\partial q^i}{\partial P_l} \frac{\partial p_i}{\partial P_k} \right). \end{aligned} \quad (21.61)$$

Отсюда вытекает первый критерий каноничности преобразований координат. $\square$

Таким образом, при каноническом преобразовании скобки Лагранжа для новых координат обязаны совпадать со скобкой Пуассона.

Матрицы Якоби для канонических преобразований обладают замечательным свойством, которое можно сформулировать в виде второго критерия канонических преобразований.

Теорема 21.11.7. *Преобразование координат фазового пространства является каноническим тогда и только тогда, когда матрица Якоби этого преобразования*

$$J_{\alpha}^{\beta} := \frac{\partial y^{\beta}}{\partial x^{\alpha}} = \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p} & \frac{\partial P}{\partial p} \end{pmatrix} \quad (21.62)$$

является симплектической, $J_{\alpha}^{\beta} \in \mathbb{SP}(n, \mathbb{R})$.

Доказательство. Доказательство проводится прямой проверкой равенства

$$\varpi_{\alpha\beta} = J_{\alpha}^{\gamma} J_{\beta}^{\delta} \varpi_{\gamma\delta} \quad (21.63)$$

с учетом свойств (21.60). Вычисления проще проводить в блочно диагональном виде (21.62). $\square$

Поскольку определитель симплектической матрицы равен единице (19.5), то форма объема фазового пространства сохраняется при каноническом преобразовании. Отсюда также следует теорема Лиувилля 21.11.2.

Третий критерий канонических преобразований формулируется в терминах скобок Пуассона.

Теорема 21.11.8. *Преобразование координат фазового пространства является каноническим тогда и только тогда, когда оно сохраняет скобки Пуассона*

$$[F, G]_{q,p} = [F, G]_{Q,P}. \quad (21.64)$$

где F и G – две произвольные дифференцируемые функции на фазовом пространстве, и скобки Пуассона вычислены, соответственно, в координатах q, p и Q, P с одинаковыми каноническими структурными функциями $\varpi^{-1\alpha\beta}$.

Доказательство. Канонические преобразования и только они сохраняют каноническую симплектическую форму. Сохранение этой формы эквивалентно сохранению матрицы $\varpi^{-1\alpha\beta}$, т.е. канонической пуассоновой структуры. $\square$

Пример 21.11.6. Преобразование

$$Q^i = q^i, \quad P_i = p_i + \frac{\partial F(q)}{\partial q^i}, \quad (21.65)$$

где $F(q) \in \mathcal{C}^2(\mathbb{R}^n)$ – произвольная функция от обобщенных координат q , является каноническим. Действительно, скобки Пуассона

$$[Q^i, Q^j] = 0, \quad [Q^i, P_j] = \delta_j^i$$

остаются прежними. Нетрудно также проверить, что

$$[P_i, P_j] = \left[P_i, \frac{\partial F}{\partial q^j} \right] + \left[\frac{\partial F}{\partial q^i}, P_j \right] = 0.$$

Аналогично, преобразование

$$Q^i = q^i + \frac{\partial G(p)}{\partial p_i}, \quad P_i = p_i,$$

где $G(p) \in \mathcal{C}^2(\mathbb{R}^N)$ – произвольная функция импульсов, также является каноническим. $\square$

Важность канонических преобразований сводится к следующему. Если удастся найти такое каноническое преобразование, при котором гамильтониан упрощается настолько, что уравнения Гамильтона явно интегрируются, то тогда можно построить и решение исходной задачи. Соответствующие канонические преобразования зависят от конкретной задачи, и общего метода их нахождения не существует.

21.12 Производящие функции канонических преобразований

Широкий класс канонических преобразований (но не все) можно описать на языке производящих функций, которые строятся следующим образом. Рассмотрим фазовое пространство $T^*(\mathbb{R}^N) \approx \mathbb{R}^{2N}$ с координатами q, p и канонической пуассоновой структурой. Пусть $2N$ функций $Q^i(q, p)$ и $P_i(q, p)$ от $2N$ переменных задают каноническое преобразование. Тогда 1-форма

$$dq p - dQ P = dS(q, p) \quad (21.66)$$

есть полный дифференциал.

Предположим, что в окрестности некоторой точки старые и новые координаты (q, Q) можно выбрать в качестве координат фазового пространства. Это значит, что

$$\det \frac{\partial(q, Q)}{\partial(q, p)} = \det \begin{pmatrix} \frac{\partial q}{\partial q} & \frac{\partial Q}{\partial q} \\ 0 & \frac{\partial Q}{\partial p} \end{pmatrix} = \det \frac{\partial Q}{\partial p} \neq 0. \quad (21.67)$$

Тогда функцию S можно локально выразить через эти координаты:

$$S(q, p) = S_1(q, Q).$$

Функция $S_1(q, Q)$ называется *производящей функцией* канонического преобразования.

Из определения производящей функции и уравнения (21.66) следует, что

$$p = \frac{\partial S_1}{\partial q}, \quad P = -\frac{\partial S_1}{\partial Q}. \quad (21.68)$$

Конечно, эти формулы являются сокращенной записью равенств

$$p_i = \frac{\partial S_1}{\partial q^i}, \quad P_i = -\frac{\partial S_1}{\partial Q^i}.$$

Полученные уравнения служат для нахождения канонического преобразования $q, p \mapsto Q(q, p), P(q, p)$. При этом выполнено равенство

$$\frac{\partial p_i}{\partial Q^j} = \frac{\partial^2 S_1}{\partial q^i \partial Q^j}.$$

Таким образом, каждая производящая функция $S_1(q, Q)$ такая, что

$$\det \frac{\partial^2 S_1}{\partial q^i \partial Q^j} \neq 0,$$

определяет некоторое каноническое преобразование по формулам (21.68). 21.67).

Для того, чтобы получить гамильтониан $K(Q, P)$ для новых канонических переменных, необходимо просто подставить в старое выражение $H(q, p)$ функции $q = q(Q, P)$ и $p = p(Q, P)$:

$$K(Q, P) := H(q(Q, P), p(Q, P)). \quad (21.69)$$

Пример 21.12.1. Пусть $S_1 = q^i Q_i$. Тогда $Q_i = p_i$ и $P^i = -q^i$. Тем самым координаты и импульсы меняются местами. Этот пример показывает, что в каноническом формализме координаты и импульсы играют совершенно равноправную роль. $\square$

Пусть гамильтониан не зависит от времени явно. Покажем, что если производящая функция $S_1(q, Q)$ сама удовлетворяет уравнению Гамильтона–Якоби по переменным q , то уравнения движения интегрируются в квадратурах. Для этого заметим, что если гамильтониан системы зависит только от новых координат $H = K(Q)$, то канонические уравнения

$$\dot{Q} = 0, \quad \dot{P} = -\frac{\partial K}{\partial Q}$$

просто интегрируются

$$Q = Q_0, \quad P = P_0 - t \left. \frac{\partial K}{\partial Q} \right|_{Q_0}.$$

Это, конечно, соответствует переходу к переменным действие-угол. Теперь будем искать производящую функцию $S_1(q, Q)$ этого канонического преобразования. Из первого условия (21.68) следует, что такая производящая функция должна удовлетворять уравнению

$$H\left(q, \frac{\partial S_1}{\partial q}\right) = K(Q). \quad (21.70)$$

То есть при каждом фиксированном значении Q производящая функция $S_1(q, Q)$ должна удовлетворять укороченному уравнению Гамильтона–Якоби (21.43). Верно также и обратное утверждение.

Теорема 21.12.1 (Якоби). Если найдено решение $S_1(q, Q)$ укороченного уравнения Гамильтона–Якоби (21.70), зависящее от N параметров Q^i и такое, что

$$\det \frac{\partial^2 S_1}{\partial q^i \partial Q^j} \neq 0,$$

то канонические уравнения (21.30) решаются в квадратурах, причем функции $Q(q, p)$, определяемые уравнениями $\partial S_1 / \partial q^i = p_i$, являются N первыми интегралами уравнений Гамильтона.

Доказательство. См., например, [143] глава 9, §47, раздел Б. □

Теорема Якоби сводит решение канонических уравнений к нахождению полного интеграла укороченного уравнения Гамильтона–Якоби. Может показаться удивительным, что сведение более простого к более сложному доставляет эффективный метод решения конкретных задач. Между тем оказывается, что это – наиболее мощный из существующих в настоящее время методов интегрирования уравнений движения, и многие задачи, решенные методом Якоби, вообще не поддаются решению другими способами.

В качестве дополнительного аргумента производящей функции можно рассматривать время, $S_1 = S_1(q, Q, t)$. При этом все предыдущие формулы остаются в силе за исключением выражения (21.69) для нового гамильтониана. Оно получает дополнительное слагаемое

$$K = H + \frac{\partial S_1}{\partial t}, \quad (21.71)$$

что следует из выражения для действия (21.56) в расширенном конфигурационном пространстве.

Если после канонического преобразования новый гамильтониан тождественно равен нулю, то уравнения движения, очевидно, интегрируются. Для этого необходимо выполнение уравнения

$$\frac{\partial S_1}{\partial t} + H(q, p, t) = 0.$$

Учитывая выражение для импульсов через производящую функцию (21.68) получаем, что производящая функция в этом случае должна удовлетворять уравнению Гамильтона–Якоби для всех значений Q . Тем самым доказана

Теорема 21.12.2 (Якоби). *Если найдено решение $S_1(q, Q, t)$ уравнения Гамильтона–Якоби (21.40), зависящее от N параметров Q^i и такое, что*

$$\det \frac{\partial^2 S_1}{\partial q^i \partial Q^j} \neq 0,$$

то канонические уравнения (21.30) решаются в квадратурах, причем координаты $Q^i = \text{const}$ и импульсы $P_i := -\partial S_1 / \partial Q^i = \text{const}$ дают $2N$ первых интегралов уравнений Гамильтона.

Частное решение уравнения Гамильтона–Якоби, зависящее от $N + 1$ параметра (по числу независимых переменных), называется *полным интегралом*. Функция от $N + 1$ параметра $S_1(q, Q, t) + \text{const}$, где S_1 – решение уравнения Гамильтона–Якоби из предыдущей теоремы является полным интегралом уравнения Гамильтона–Якоби. В том случае, когда для уравнения Гамильтона–Якоби можно определить общее решение, оно будет зависеть от произвольных функций. Это значит, что полный интеграл уравнения Гамильтона–Якоби дает лишь незначительную часть всех решений.

Если известен полный интеграл уравнения Гамильтона–Якоби $S_1(q, Q, t) + \text{const}$, то для нахождения траектории частицы, проходящей через заданную точку фазового пространства в начальный момент времени, необходимо решить дополнительные уравнения:

$$\frac{\partial S}{\partial Q^i} = c_i, \quad (21.72)$$

где c_i – некоторые постоянные, определяемые начальным положением частицы.

Продолжим изучение производящих функций. Выше был рассмотрен случай, когда производящая функция зависит от старых и новых координат. Аналогичным образом можно показать, что канонические преобразования генерируются производящими функциями $S_2(q, P, t)$, зависящими от старых координат и новых импульсов. Если

$$\det \frac{\partial^2 S_2}{\partial q^i \partial P_j} \neq 0,$$

то каноническое преобразование задается формулами

$$p := \frac{\partial S_2}{\partial q}, \quad Q := \frac{\partial S_2}{\partial P}, \quad K := H + \frac{\partial S_2}{\partial t}. \quad (21.73)$$

Переход от производящей функции $S_1(q, Q)$ к $S_2(q, P)$ является преобразованием Лежандра по переменной Q при этом

$$S_2(q, P, t) = S_1(q, Q, t) + PQ.$$

Пример 21.12.2. Каноническое преобразование, генерируемое производящей функцией $S_2 = qP := q^i P_i$ является тождественным. $\square$

Пример 21.12.3. Пусть $S_2 = q^i S_i^j P_j$, где S_i^j – произвольная невырожденная матрица. Тогда $P_i = S_i^{-1j} p_j$ и $Q^i = q^j S_j^i$. То есть это преобразование генерирует линейное однородное преобразование фазового пространства. $\square$

Пример 21.12.4. Выберем производящую функцию бесконечно малых канонических преобразований в виде

$$S_2 = qP + \epsilon S(q, P, \epsilon), \quad \epsilon \ll 1,$$

где $S(q, P, \epsilon)$ – произвольная функция, зависящая от параметра ϵ . Тогда

$$Q = q + \epsilon \frac{\partial S}{\partial P}, \quad p = P + \epsilon \frac{\partial S}{\partial q}.$$

Отсюда следует, что бесконечно малые канонические преобразования удовлетворяют каноническим уравнениям

$$\left. \frac{dQ}{d\epsilon} \right|_{\epsilon=0} = \frac{\partial H}{\partial p}, \quad \left. \frac{dP}{d\epsilon} \right|_{\epsilon=0} = -\frac{\partial H}{\partial q}$$

с функцией Гамильтона $H(q, p) = S(q, p, 0)$, т.к. в главном порядке $p = P$. Тем самым гамильтониан произвольной механической системы можно рассматривать, как генератор бесконечно малых канонических преобразований по времени $t = \epsilon$, которые происходят по мере эволюции механической системы. $\square$

На самом деле движение частицы в фазовом пространстве $q(t), p(t)$ можно рассматривать как последовательность канонических преобразований, параметризующихся одним параметром t . Допустим, что решение канонических уравнений движения определено для всех $t \in \mathbb{R}$. Тогда канонические уравнения движения (21.13) определяют однопараметрическую группу преобразований фазового пространства. Для каждого момента времени, по определению, выполнены канонические коммутационные соотношения (21.14). Следовательно, для каждого момента времени t отображение фазового пространства

$$\mathbb{R}^{2N} \ni (q(0), p(0)) \mapsto (q(t), p(t)) \in \mathbb{R}^{2N}$$

представляет собой каноническое преобразование.

В заключение рассмотрим еще два вида производящих функций, часто встречающихся в приложениях. Если производящая функция $S_3(p, Q, t)$ зависит от старых импульсов и новых координат, и при этом

$$\det \frac{\partial^2 S_3}{\partial p_i \partial Q^j} \neq 0,$$

то каноническое преобразование задается формулами:

$$q := -\frac{\partial S_3}{\partial p}, \quad P := -\frac{\partial S_3}{\partial Q}, \quad K := H + \frac{\partial S_3}{\partial t}. \quad (21.74)$$

При этом производящая функция S_3 является преобразованием Лежандра функции S_1 по переменной q :

$$S_3(p, Q, t) = S_1(q, Q, t) - pq.$$

Пример 21.12.5. Каноническое преобразование, генерируемое производящей функцией $S_3 = -p_i Q^i$, является тождественным. $\square$

Производящая функция $S_3(p, Q, t)$ полезна в следующем случае. Допустим, что нам заранее известно, как удобно выбрать координаты в конфигурационном пространстве, т.е. заданы функции $q = q(Q)$. Например, если конфигурационное пространство инвариантно относительно действия некоторой группы преобразований, то новые координаты Q удобно выбрать в соответствии с симметрией задачи. Тогда производящая функция $S_3 = -p_i q^i(Q)$ дает выражения для новых обобщенных импульсов P_i , сопряженных к координатам Q^i .

Точечные преобразования, рассмотренные в разделе 21.3, относятся именно к этому типу. Точечные преобразования – это все преобразования координат, которые допустимы в лагранжевой механике. Рассмотренный выше пример показывает, что точечные преобразования составляют лишь небольшой класс канонических преобразований. Все множество канонических преобразований намного шире, и это является причиной большей гибкости канонического формализма по сравнению с лагранжевым.

Продолжим изучение производящих функций. Если производящая функция $S_4(p, P, t)$ зависит от старых и новых импульсов, и

$$\det \frac{\partial^2 S_4}{\partial p_i \partial P_j} \neq 0,$$

то каноническое преобразование задается формулами:

$$q := -\frac{\partial S_4}{\partial p}, \quad Q := \frac{\partial S_4}{\partial P}, \quad K := H + \frac{\partial S_4}{\partial t}. \quad (21.75)$$

При этом производящая функция S_4 является двойным преобразованием Лежандра функции $S_1(q, Q)$ по переменным q и Q :

$$S_4(p, P, t) = S_1(q, Q, t) - pq + PQ.$$

Пример 21.12.6. Каноническое преобразование, генерируемое производящей функцией $S_4 = p_i P^i$, меняет местами координаты и импульсы: $Q_i = p_i$, $P^i = -q_i$. Это преобразование совпадает с каноническим преобразованием из примера 21.12.1. $\square$

Пример 21.12.7 (Гармонический осциллятор). Продемонстрируем сложные построения последних разделов на примере гармонического осциллятора. Пусть задана функция Лагранжа

$$L = \frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega^2 q^2,$$

где $\omega = \text{const}$ – собственная частота гармонического осциллятора. Конфигурационное пространство представляет собой прямую $q \in \mathbb{R}$. Импульс, сопряженный координате q , равен скорости

$$p = \frac{\partial L}{\partial \dot{q}} = \dot{q}.$$

Преобразование Лежандра приводит к положительно определенному гамильтониану:

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2.$$

Фазовым пространством для гармонического осциллятора является евклидова плоскость $(q, p) \in \mathbb{T}^*(\mathbb{R}) \approx \mathbb{R}^2$ с канонической пуассоновой структурой. Соответствующие уравнения движения имеют вид

$$\left. \begin{aligned} \dot{q} &= p, \\ \dot{p} &= -\omega^2 q, \end{aligned} \right\} \Leftrightarrow \ddot{q} + \omega^2 q = 0. \quad (21.76)$$

Общее решение этих уравнений параметризуется двумя произвольными постоянными: амплитудой A_0 и фазой φ ,

$$\begin{aligned} q &= A_0 \cos(\omega t + \varphi), \\ p &= -A_0 \omega \sin(\omega t + \varphi). \end{aligned} \quad (21.77)$$

Фаза φ соответствует выбору начала отсчета времени. Полученное решение соответствует решению задачи Коши с начальными данными

$$\begin{aligned} q_0 &= A_0 \cos \varphi, \\ p_0 &= -A_0 \omega \sin \varphi. \end{aligned}$$

В терминах начальных условий общее решение (21.77) перепишется в виде

$$\begin{aligned} q &= q_0 \cos \omega t + \frac{p_0}{\omega} \sin \omega t, \\ p &= -\omega q_0 \sin \omega t + p_0 \cos \omega t. \end{aligned}$$

Траекториями гармонического осциллятора в фазовом пространстве являются эллипсы. Как видим, эволюция во времени представляет собой эллиптическое вращение фазового пространства $(q_0, p_0) \mapsto (q(t), p(t))$. Если $\mathbb{V}_0 \subset \mathbb{R}^2$ – произвольная ограниченная область фазового пространства, то в соответствии с теоремой Лиувилля (21.11.2) объем этой области сохраняется во времени:

$$\int_{\mathbb{V}_0} dq_0 dp_0 = \int_{\mathbb{V}(t)} dq dp = \int_{\mathbb{V}(t)} dq_0 dp_0,$$

т.к. якобиан преобразования координат равен единице,

$$\det \frac{\partial(q, p)}{\partial(q_0, p_0)} = 1.$$

Поскольку гамильтониан не зависит от времени явно, то энергия на каждой траектории постоянна $E = H = \text{const}$.

Найдем функцию действия $S(q, t)$. Чтобы вычислить интеграл (21.38) необходимо исключить скорости $\dot{q}$ из лагранжиана. Поскольку интегрирование ведется вдоль траектории частицы, то в лагранжиан необходимо подставить

$$\dot{q} = p = \sqrt{2E - \omega^2 q^2}.$$

В этом выражении E рассматривается, как некоторая постоянная. После несложных вычислений получаем функцию действия

$$\begin{aligned} S(q, t) &= - \int dt E + \int dq \sqrt{2E - \omega^2 q^2} = \\ &= -Et + \frac{E}{\omega} \arcsin \frac{\omega q}{\sqrt{2E}} + \frac{q}{2} \sqrt{2E - \omega^2 q^2} \end{aligned}$$

с точностью до несущественной постоянной, определяемой начальными данными. Нетрудно проверить, что

$$\frac{\partial S}{\partial q} = \sqrt{2E - \omega^2 q^2} = p.$$

Поэтому уравнение Гамильтона–Якоби (21.40) выполняется. Укороченное действие имеет вид

$$W(q, E) = \frac{E}{\omega} \arcsin \frac{\omega q}{\sqrt{2E}} + \frac{q}{2} \sqrt{2E - \omega^2 q^2}.$$

Таким образом, нам известно укороченное действие, зависящее от параметра E , причем

$$\frac{\partial^2 W}{\partial q \partial E} \neq 0.$$

Поэтому его можно выбрать в качестве производящей функции канонического преобразования $q, p \rightarrow Q, P$

$$S_2(q, P) = \frac{P}{\omega} \arcsin \frac{\omega q}{\sqrt{2P}} + \frac{q}{2} \sqrt{2P - \omega^2 q^2}.$$

Воспользовавшись формулами (21.73), получаем явный вид канонических преобразований

$$\begin{aligned} q &= \frac{\sqrt{2P}}{\omega} \sin(\omega Q), \\ p &= \sqrt{2P} \cos(\omega Q). \end{aligned}$$

Отсюда вытекает связь дифференциалов:

$$\begin{aligned} dq &= \sqrt{2P} \cos(\omega Q) dQ + \frac{1}{\omega \sqrt{2P}} \sin(\omega Q) dP, \\ dp &= -\omega \sqrt{2P} \sin(\omega Q) dQ + \frac{1}{\sqrt{2P}} \cos(\omega Q) dP. \end{aligned}$$

Теперь нетрудно убедиться в том, что построенное преобразование действительно является каноническим

$$dp \wedge dq = dP \wedge dQ.$$

В новых переменных гамильтониан имеет вид $H = P$ и уравнения движения

$$\dot{Q} = P, \quad \dot{P} = 0$$

тривиально интегрируются. Таким образом, укороченная функция действия для гармонического осциллятора является производящей функцией к переменным “действие-угол”. Фазовым пространством в переменных “действие-угол” является полуплоскость $Q \in \mathbb{R}$, $P \geq 0$, а траекториями – прямые линии $P = \text{const}$. $\square$

В настоящем разделе мы рассмотрели широкий класс канонических преобразований, которые генерируются четырьмя видами производящих функций: $S_1(q, Q, t)$, $S_2(q, P, t)$, $S_3(p, Q, t)$ и $S_4(p, P, t)$. Существуют также производящие функции более общего вида, которые мы сейчас опишем.

Предложение 21.12.1. Пусть заданы две системы координат q, p и Q, P в фазовом пространстве $\mathbb{R}^{2N}$ с канонической пуассоновой структурой. Тогда среди $4N$ координатных функций q, p, Q, P всегда можно выбрать $2N$ независимых функций так, чтобы среди них не было ни одной пары канонически сопряженных функций q^i, p_i или Q^i, P_i .

Доказательство. См., например, [145], глава IV, лемма из §29. $\square$

Из данного утверждения следует, что в качестве координат на фазовом пространстве $\mathbb{R}^{2N}$ всегда можно выбрать набор функций

$$q^1, \dots, q^L, p_{L+1}, \dots, p_N, Q^1, \dots, Q^M, P_{M+1}, \dots, P^N, \quad 0 \leq L, M \leq N,$$

состоящий из N старых и N новых координат. Введем обозначения:

$$\begin{aligned} \{z^i\} &:= \{q^1, \dots, q^L, p_{L+1}, \dots, p_N\}, \\ \{Z^i\} &:= \{Q^1, \dots, Q^M, P_{M+1}, \dots, P^N\}. \end{aligned}$$

Допустим, что на фазовом пространстве задана функция $S(z, Z)$ такая, что

$$\det \frac{\partial^2 S}{\partial z^i \partial Z^j} \neq 0. \quad (21.78)$$

Тогда нетрудно проверить, что функция $S(z, Z)$ определяет каноническое преобразование:

$$\begin{aligned} p_i &:= \frac{\partial S}{\partial q^i}, & i &= 1, \dots, L, \\ q^i &:= -\frac{\partial S}{\partial p_i}, & i &= L+1, \dots, N, \\ P_i &:= -\frac{\partial S}{\partial Q^i}, & i &= 1, \dots, M, \\ Q^i &:= \frac{\partial S}{\partial P_i}, & i &= M+1, \dots, N. \end{aligned}$$

Рассмотренное каноническое преобразование включает в себя четыре предыдущих при $L = 0, N$ и $M = 0, N$. В общем случае производящая функция $S(z, Z)$ зависит от старых координат и импульсов и новых координат и импульсов. При этом среди них нет ни одной пары канонически сопряженных переменных. Примеры 21.12.1 и 21.12.6 показывают, что различные производящие функции могут приводить к одинаковым каноническим преобразованиям. Во всех случаях от производящей функции требуется отличие от нуля определителя (21.78). Поэтому производящие функции описывают широкий класс канонических преобразований, но не все.

Пример 21.12.8. В примере 21.11.6 из предыдущего раздела было рассмотрено каноническое преобразование, которое не охватывается производящими функциями, рассмотренными в настоящем разделе. Действительно, для преобразования (21.65) формула (21.66) принимает вид

$$-dq \frac{\partial F(q)}{\partial q} = dS(q, p) = dq \frac{\partial S}{\partial q} + dp \frac{\partial S}{\partial p}.$$

Откуда следуют равенства:

$$\frac{\partial S}{\partial q} = -\frac{\partial F}{\partial q}, \quad \frac{\partial S}{\partial p} = 0.$$

Таким образом, для канонических преобразований вида (21.65) функция $S(q, p)$ имеет вид

$$S(q, p) = -F(q) + \text{const}$$

и не относится ни к одному типу производящих функций настоящего раздела. $\square$

21.13 Разделение переменных в уравнении Гамильтона–Якоби

Теорема Якоби 21.12.2 утверждает, что если найден полный интеграл уравнения Гамильтона–Якоби, то канонические уравнения решаются в квадратурах. По-видимому, все известные в настоящее время полные интегралы могут быть найдены путем *разделения переменных* в уравнении Гамильтона–Якоби. Суть метода состоит в следующем.

Допустим, что одна из координат – обозначим ее q^N – и соответствующая ей производная $\partial S / \partial q^N$ входят в уравнение Гамильтона–Якоби только в виде комбинации $f(q^N, \partial S / \partial q^N)$, где f – некоторая функция двух переменных, не зависящая от других координат, соответствующих производных действия и времени. Тогда будем искать решение уравнения Гамильтона–Якоби (21.40) в виде

$$S(q, t) = \check{S}(\check{q}, t) + S_N(q^N),$$

где $\check{S}(\check{q}, t)$ – функция оставшихся переменных $\check{q} = (q^1, \dots, q^{N-1})$ и времени, а $S_N(q^N)$ зависит только от отделяемой координаты q^N . Тогда уравнение Гамильтона–Якоби принимает вид

$$\frac{\partial \check{S}}{\partial t} + H \left(\check{q}, \frac{\partial \check{S}}{\partial \check{q}}, f \left(q^N, \frac{dS_N}{dq^N} \right), t \right) = 0. \quad (21.79)$$

Допустим, что решение этого уравнения найдено. Тогда, после подстановки найденного решения в уравнение (21.79) мы получаем тождество, справедливое для всех значений координат. При изменении координаты q^N меняется только функция f . Поэтому для выполнения уравнения (21.79) необходимо, чтобы функция f сама была равна постоянной, которую мы обозначим Q^N , для произвольного решения. Таким образом, уравнение Гамильтона–Якоби (21.79) распадается на два уравнения:

$$f \left(q^N, \frac{dS_N}{dq^N} \right) = Q^N, \quad (21.80)$$

$$\frac{\partial \check{S}}{\partial t} + H \left(\check{q}, \frac{\partial \check{S}}{\partial \check{q}}, Q^N, t \right) = 0. \quad (21.81)$$

Первое уравнение представляет собой обыкновенное дифференциальное уравнение первого порядка, из которого можно определить функцию $S_N = S_N(q^N; Q^N)$. Если функция f достаточно проста, то это уравнение иногда удается проинтегрировать явно. Постоянная Q^N является первым интегралом исходного уравнения Гамильтона–Якоби. После этого остается решить оставшееся уравнение (21.81), которое также имеет вид уравнения Гамильтона–Якоби, но для системы с меньшим числом степеней свободы. Эта система описывается гамильтонианом, который получается из исходного путем замены функции f для отделенной степени свободы на некоторую константу. Единственное ограничение на постоянную Q^N состоит в том, чтобы уравнение (21.80) имело решение.

Иногда этот процесс отделения переменных можно продолжить. Если таким образом можно разделить все N координат, то тогда полный интеграл уравнения Гамильтона–Якоби находится в квадратурах. В этом случае мы говорим, что уравнение Гамильтона–Якоби допускает *разделение переменных*.

Замечание. Разделение переменных в уравнении Гамильтона–Якоби для одной и той же механической системы возможно не во всякой системе координат конфигурационного пространства. Как правило системы координат, в которых переменные делятся полностью или частично связаны с симметрией системы. Процедура нахождения такой системы координат, в которой переменные для интегрируемой системы разделяются, представляет собой отдельную задачу. $\square$

Допустим, что механическая система консервативна и переменные делятся. Тогда полный интеграл уравнения Гамильтона–Якоби имеет вид

$$S(q, t) = -E(Q^1, \dots, Q^N)t + \sum_{k=1}^N S_k(q^k; Q^k, \dots, Q^N), \quad (21.82)$$

где каждая функция S_k зависит от одной координаты q^k , а энергия E как функция первых интегралов $Q^1, \dots, Q^N$ получается подстановкой укороченного действия

$$W = \sum_{k=1}^N S_k(q^k, Q^k, \dots, Q^N)$$

в укороченное уравнение Гамильтона–Якоби (21.43).

Пример 21.13.1 (Циклические координаты). Пусть q^N – циклическая координата, т.е. гамильтониан от нее вообще не зависит. В этом случае функция $f(q^N, \partial S/\partial q^N)$ сводится просто к $\partial S/\partial q^N$. Тогда из уравнения (21.80) следует, что $S_N = q^N Q_N$ (никакого суммирования). Тогда функция действия принимает вид

$$S(q, t) = \check{S}(\check{q}, t) + q^N Q_N.$$

В этом случае первый интеграл $Q_N = \partial S/\partial q^N = p_N$ имеет смысл обобщенного импульса, сопряженного циклической координате q^N .

Отметим, что если время t рассматривать, как обобщенную координату, то отделение времени в виде слагаемого $-Et$ в функции действия (21.42) для консервативной системы соответствует методу разделения переменных для циклической координаты t . При этом энергия E сохраняется и соответствует импульсу для обобщенной координаты t . $\square$

Разделение переменных для консервативной системы возможно также в следующем более общем случае. Пусть гамильтониан механической системы имеет вид

$$H(q, p) = \frac{f_1(q^1, p_1) + \dots + f_N(q^N, p_N)}{g_1(q^1, p_1) + \dots + g_N(q^N, p_N)},$$

где $f_i(q^i, p_i)$ и $g_i(q^i, p_i)$ – некоторые функции только от указанных аргументов. Предыдущий случай возникает, например, при $g_i = \text{const}$, $i = 1, \dots, N$, $\sum_i g_i \neq 0$. Будем искать решение уравнения Гамильтона–Якоби в виде

$$S(q, t) = -Et + \sum_{i=1}^N W_i(q^i).$$

Тогда укороченное уравнение Гамильтона–Якоби (21.43) примет вид

$$\sum_{i=1}^N \left[f_i \left(q^i, \frac{dW_i}{dq_i} \right) - E g_i \left(q^i, \frac{dW_i}{dq_i} \right) \right] = 0.$$

Поскольку координаты q_i меняются независимо, то каждое слагаемое в данной сумме должно быть равно некоторой постоянной,

$$f_i \left(q^i, \frac{dW_i}{dq_i} \right) - E g_i \left(q^i, \frac{dW_i}{dq_i} \right) = c_i. \quad (21.83)$$

При этом не все постоянные c_i являются независимыми, т.к. должно быть выполнено равенство

$$c_1 + \dots + c_N = 0.$$

Допустим, что все уравнения (21.83) можно разрешить относительно производных:

$$\frac{dW_i}{dq^i} = F_i(q^i, c_i).$$

Тогда полный интеграл уравнения Гамильтона–Якоби равен сумме

$$S(q, t) = -Et + \sum_{i=1}^N \int dq^i F_i(q^i, c_i, E).$$

Пример 21.13.2 (Задача Кеплера). Рассмотрим частицу массы m , которая движется по плоскости с полярными координатами r, φ под действием центральной силы с потенциалом $U(r)$. Гамильтониан такой системы имеет вид

$$H(r, \varphi) = \frac{1}{2m} \left(p_r^2 + \frac{p_\varphi^2}{r^2} \right) + U(r). \quad (21.84)$$

К этой задаче сводится задача движения двух тел, находящихся под действием гравитационного взаимодействия с потенциалом $U(r) = -m_1 m_2 / r$ после отделения движения центра масс. Поскольку гамильтониан не зависит от угла, то координата φ является циклической, и переменные делятся. Поэтому функция действия имеет вид

$$S = -Et + W(r, E, M) + \varphi M, \quad (21.85)$$

где $p_\varphi = M = \text{const}$ – сохраняющийся угловой импульс частицы (момент количества движения). В этом случае уравнение Гамильтона–Якоби сводится к одному обыкновенному дифференциальному уравнению первого порядка

$$\left(\frac{dW}{dr}\right)^2 = 2m(E - U) - \frac{M^2}{r^2}. \quad (21.86)$$

Произвол в решении данного уравнения фиксируется положением частицы в начальный момент времени. В соответствии с общими утверждениями, функция действия (21.85), где W – решение уравнения (21.86), зависит от двух постоянных: энергии E и момента количества движения M , и представляет собой полный интеграл уравнения Гамильтона–Якоби. Таким образом, задача двух тел, взаимодействующих посредством центральных сил с произвольным потенциалом $U(r)$, решается в квадратурах. Чтобы найти траекторию частицы, проходящую через заданную точку фазового пространства в начальный момент времени, в соответствии с теоремой Якоби (21.12.2) необходимо решить дополнительные уравнения:

$$\frac{\partial S}{\partial M} = \frac{\partial W}{\partial M} + \varphi = c_1, \quad \frac{\partial S}{\partial E} = -t + \frac{\partial W}{\partial E} = c_2,$$

где $c_{1,2}$ – некоторые постоянные (21.73), определяемые начальным положением частицы.

Заметим, что переменные в задаче Кеплера на плоскости делятся в полярных координатах (в трехмерном пространстве – в сферических), которые связаны со сферической симметрией потенциала взаимодействия. В декартовой системе координат переменные в уравнении Гамильтона–Якоби не делятся. $\square$

Глава 22

Гамильтонова динамика частиц со связями

В современной математической физике большую роль играют модели, инвариантные относительно действия локальных групп преобразований, которые принято называть калибровочными. С точки зрения гамильтонова формализма такие модели соответствуют системам со связями. В калибровочных моделях связи возникают не как дополнительные условия на канонические переменные, наложенные извне, а при переходе от вырожденного инвариантного лагранжиана к гамильтониану. Основы гамильтонова формализма для калибровочных моделей были заложены П. Дираком на примере точечных частиц [146, 147, 148]. Обобщение этого формализма на теорию поля представляет значительные трудности, которые мы кратко обсудим в разделе 22.6. Несмотря на важность, эта задача до сих пор не решена. В теории поля до настоящего времени, как правило, используют методы, развитые для точечных частиц. В простейших случаях это приводит к разумным результатам.

В последующих разделах мы обсудим вопрос о гамильтоновой динамике со связями для систем точечных частиц [149, 150]. Квантование калибровочных моделей с помощью функционального интеграла в фазовом пространстве было развито в [151, 152].

22.1 Связи в гамильтоновом формализме

Рассмотрим систему N точечных частиц с действием (21.7). Предположим, что фазовое пространство представляет собой дифференцируемое многообразие $\mathbb{N}$ класса $\mathcal{C}^k$, $\dim \mathbb{N} = 2N$. Для простоты предположим, что фазовое пространство топологически тривиально $\mathbb{N} \approx \mathbb{R}^{2N}$ и покрывается одной картой. Координаты на фазовом пространстве $\mathbb{N}$, как и раньше, обозначим $(q, p) = (q^1, \dots, q^N, p_1, \dots, p_N)$. Мы предполагаем, что на фазовом пространстве задана каноническая пуассонова структура $[q^i, p_j] = \delta_j^i$.

Пусть на канонические переменные наложены *связи* в виде $2M$ алгебраических условий:

$$\Phi^\mu(q, p) = 0, \quad \Phi^\mu \in \mathcal{C}^k(\mathbb{N}), \quad \mu = 1, \dots, 2M < 2N. \quad (22.1)$$

Мы пишем индекс связи сверху, поскольку в дальнейшем связи будут частью новых координатных функций. Условия (22.1) представляют собой систему уравнений на канонические переменные. Будем считать, что все Φ^μ функционально независимы, т.е. ранг матрицы $\partial\Phi^\mu/\partial x^\alpha$, $\alpha = 1, \dots, 2N$, где $x = (q, p)$, максимален и равен $2M$ в окрестности всех точек, координаты которых удовлетворяют системе уравнений

(22.1). Если все связи принадлежат классу $\mathcal{C}^k$, то их совокупность, согласно теореме 2.10.4, определяет подмногообразие

$$\mathbb{M} = \{x \in \mathbb{N} : \Phi = 0\}$$

того же класса гладкости и размерности $\dim \mathbb{M} = 2(N - M)$. Это подмногообразие называется *поверхностью связей*, которую будем обозначать равенством $\Phi = 0$ без индекса.

Замечание. Вообще говоря, $\Phi^\mu(q, p)$ не являются функциями на фазовом пространстве, т.к. могут иметь тензорные индексы. Например, в электродинамике первичная связь имеет вид $p_0 = 0$. $\square$

Условие функциональной независимости связей накладывает жесткие ограничения на вид функций Φ^μ .

Пример 22.1.1. Пусть на фазовом пространстве $\mathbb{N}$ задана одна связь $\Phi(q, p) = 0$. Условие функциональной независимости означает, что ранг матрицы $\partial_\alpha \Phi$, которая в рассматриваемом случае представляет собой строку, равен единице, т.е. по крайней мере одна из частных производных в каждой точке поверхности связей $\mathbb{M}$ отлична от нуля. Если возвести связь в некоторую степень, то уравнение $\Phi^n = 0$, $n = 2, 3, \dots$, определяет ту же поверхность связей $\mathbb{M} \subset \mathbb{N}$ (как множество точек). Однако связь $\Phi^n = 0$ не является функционально независимой. Действительно,

$$\partial_\alpha \Phi^n|_{\Phi=0} = n\Phi^{n-1}\partial_\alpha \Phi|_{\Phi=0} = 0.$$

Поэтому ранг матрицы $\partial_\alpha \Phi^n$ равен нулю, и связь $\Phi^n = 0$ не является функционально независимой на $\mathbb{M} \subset \mathbb{N}$, хотя и определяет ту же поверхность связей как множество точек. $\square$

Рассмотренный пример показывает, что условие функциональной независимости связей означает не только определение поверхности связей $\mathbb{M} \subset \mathbb{N}$ как множества точек, но и возможность выбора функций Φ^μ в качестве трансверсальных координат к $\mathbb{M}$ в некоторой окрестности поверхности связей.

Уравнения, задающие поверхность связей, определены неоднозначно. Например, невырожденные линейные комбинации исходных связей

$$\Psi^\mu(q, p) := \Phi^\nu(q, p)S_\nu^\mu(q, p), \quad (22.2)$$

где матрица S_ν^μ невырождена на всем фазовом пространстве $\mathbb{N}$, определяют ту же поверхность связей $\mathbb{M} \subset \mathbb{N}$ и функционально независимы.

Определение. Системы связей $\Phi = 0$ и $\Psi = 0$ называются *эквивалентными*, если они связаны между собой невырожденным линейным преобразованием (22.2) во всем фазовом пространстве. $\square$

Замечание. Если потребовать выполнения равенства (22.2) не во всем фазовом пространстве, а только на поверхности связей, то этого недостаточно для эквивалентности. Действительно, пусть поверхность связей $\mathbb{M} \subset \mathbb{N}$ определена уравнениями $\Phi = 0$. Определим новые связи $\Psi = 0$ на поверхности связей с помощью некоторой невырожденной матрицы S , заданной на $\mathbb{M}$. После этого продолжим функции Ψ на все фазовое пространство таким образом, чтобы уравнения $\Psi = 0$ определяли некоторое

подмногообразие вида $\mathbb{M} \cup \mathbb{U} \subset \mathbb{N}$, где $\mathbb{U}$ – некоторое собственное подмногообразие $\mathbb{N}$ такое, что $\mathbb{U} \cap \mathbb{M} = \emptyset$. Никаких препятствий для этого нет. Таким образом, для новых связей условие (22.2) будет выполнено на $\mathbb{M}$, однако условия $\Psi = 0$ выделяют в фазовом пространстве $\mathbb{N}$ другое подмногообразие. Это означает, что матрица S будет вырождена на $\mathbb{U}$. $\square$

Предложение 22.1.1. Любая функция $f(q, p) \in \mathcal{C}^k(\mathbb{N})$, обращающаяся в нуль на поверхности связей $\mathbb{M} \subset \mathbb{N}$, в некоторой окрестности $\mathbb{U}_0 \subset \mathbb{N}$ произвольной точки $(q_0, p_0) \in \mathbb{M}$ представима в виде

$$f(q, p) = \Phi^\mu(q, p) f_\mu(q, p), \quad (q, p) \in \mathbb{U}_0, \quad (22.3)$$

с достаточно гладкими коэффициентами $f_\mu(q, p) \in \mathcal{C}^k(\mathbb{U}_0)$.

Доказательство. Поскольку связи Φ^μ функционально независимы, то в некоторой окрестности произвольной точки на поверхности связей $(q_0, p_0) \in \mathbb{M}$ их можно дополнить до системы координат $\{x^a\} \rightarrow \{y^a, \Phi^\mu\}$, где y^a , $a = 1, \dots, 2(n-m)$, – координаты на поверхности связей. Тогда согласно предложению 2.6.2 произвольная достаточно гладкая функция $f \in \mathcal{C}^k(\mathbb{U}_0)$ представима в виде

$$f(q, p) = f(q_0, p_0) + y^a(q, p) f_a(q, p) + \Phi^\mu(q, p) f_\mu(q, p)$$

с достаточно гладкими коэффициентами f_μ , f_a . Так как функция f равна нулю на поверхности связей $\mathbb{M}$, то $f(q_0, p_0) = 0$ и $f_a(q, p) = 0$ для всех a и $(q, p) \in \mathbb{U}_0$. Поэтому справедливо представление (22.3). $\square$

Из этого предложения следует, что если любые два набора связей определяют одно и то же подмножество $\mathbb{M} \subset \mathbb{N}$, то они связаны линейным преобразованием (22.2). При этом невырожденность матрицы S необходима для функциональной независимости.

22.2 Гамильтонова динамика частиц со связями II рода

Начнем рассмотрение гамильтоновой динамики частиц со связями с обсуждения связей второго рода, т.к. она значительно проще. Связи первого рода будут рассмотрены в следующем разделе. Там будет показано, что динамика частиц со связями I рода после наложения канонической калибровки сводится к рассмотрению частиц со связями II рода.

Определение. Связи Φ^μ , $\mu = 1, \dots, 2m$, называются *связями II рода*, если определитель матрицы, составленной из скобок Пуассона связей между собой, отличен от нуля на поверхности связей:

$$\det [\Phi^\mu, \Phi^\nu]_{\Phi=0} \neq 0. \quad \square \quad (22.4)$$

Отметим, что элементы матрицы, составленной из скобок Пуассона $[\Phi^\mu, \Phi^\nu]$, в общем случае зависят от точки фазового пространства. Условие (22.4) означает отличие определителя от нуля в каждой точке подмногообразия $\mathbb{M} \subset \mathbb{N}$, определяемого связями (22.1). Из непрерывности функций Φ^μ следует, что определитель (22.4) отличен от нуля не только на $\mathbb{M}$, но и в некоторой окрестности подмногообразия $\mathbb{M}$.

Во многих случаях эта окрестность оказывается большой и даже совпадает со всем $\mathbb{N}$.

Поскольку матрица, составленная из скобок Пуассона,

$$J^{\mu\nu} = -J^{\nu\mu} := [\Phi^\mu, \Phi^\nu], \quad (22.5)$$

антисимметрична, то ее определитель может быть отличен от нуля только при четном числе связей. Поэтому мы с самого начала предположили наличие четного числа связей $2M$. Это значит, что поверхность связей также имеет четную размерность $2(N - M)$.

С точки зрения вариационного принципа при наличии связей на фазовом пространстве мы имеем задачу на условный экстремум, которую можно решать методом неопределенных множителей Лагранжа. С этой целью рассмотрим *обобщенный* (*extended*) *гамильтониан*, который получается после добавления к исходному гамильтониану всех связей второго рода:

$$H_E = H + \lambda_\mu \Phi^\mu, \quad (22.6)$$

где $\lambda_\mu = \lambda_\mu(q, p, t)$ – неопределенные множители Лагранжа. Этому гамильтониану соответствует *обобщенное действие*

$$S_E = \int_{t_1}^{t_2} dt (p\dot{q} - H - \lambda_\mu \Phi^\mu). \quad (22.7)$$

При варьировании этого действия мы считаем вариации координат на границе нулевыми, $\delta q(t_{1,2}) = 0$, а вариации импульсов и множителей Лагранжа могут быть произвольны, т.к. они входят в действие без производных. Канонические уравнения движения для обобщенного гамильтониана имеют вид

$$\begin{aligned} \dot{q}^i &= [q^i, H_E] = [q^i, H] + \lambda_\mu [q^i, \Phi^\mu] + [q^i, \lambda_\mu] \Phi^\mu, \\ \dot{p}_i &= [p_i, H_E] = [p_i, H] + \lambda_\mu [p_i, \Phi^\mu] + [p_i, \lambda_\mu] \Phi^\mu. \end{aligned} \quad (22.8)$$

Эти уравнения необходимо дополнить уравнениями связей (22.1), возникающими при варьировании обобщенного действия по множителям Лагранжа. В уравнениях движения (22.8) последние слагаемые можно отбросить, т.к. они равны нулю на поверхности связей (22.1). Таким образом, мы имеем $2(N + M)$ уравнений (22.8), (22.1) на $2(N + M)$ переменных q, p, λ , из которых $2N$ уравнений являются дифференциальными. Решение этой задачи можно провести следующим образом.

Предложение 22.2.1. *Для того, чтобы фазовая траектория гамильтоновой системы (22.7), проходящая через произвольную точку поверхности связей, целиком лежала на этой поверхности необходимо и достаточно, чтобы производная по времени от всех связей обращалась в нуль на поверхности связей,*

$$\dot{\Phi}^\mu = [\Phi^\mu, H_E]_{\Phi=0} = [\Phi^\mu, H]_{\Phi=0} + \lambda_\nu [\Phi^\mu, \Phi^\nu]_{\Phi=0} = 0. \quad (22.9)$$

Доказательство. Достаточность. Допустим, что в начальный момент времени траектория находилась на поверхности связей, т.е. $\Phi^\mu(q_0, p_0) = 0$, где $q_0 := q(0)$ и $p_0 := p(0)$. Тогда с течением времени уравнения $\Phi^\mu(q(t), p(t)) = 0$ будут выполнены на решении задачи Коши для системы уравнений (22.9).

Необходимость. Допустим, что каждая траектория, имеющая хотя бы одну точку на поверхности связей, целиком принадлежит этой поверхности. Тогда производная вдоль траектории от каждой связи должна быть равна нулю на поверхности связей, поскольку мы можем произвольно менять исходную точку на поверхности. Это означает выполнение условий (22.9) для всех значений индекса μ . $\square$

Заметим, что условие $\dot{\Phi}^\mu = 0$ не может привести к новым связям. Действительно, наличие новых связей означает, что система уравнений (22.9) не определяет все множители Лагранжа однозначно. А это противоречит условию (22.4).

Уравнения (22.9) можно рассматривать как уравнения на множители Лагранжа на поверхности связей. При этом условие, определяющее связи второго рода (22.4), является необходимым и достаточным для однозначного определения множителей Лагранжа $\lambda_\mu(x)$ на поверхности связей $\mathbb{M}$. Вне поверхности связей множители Лагранжа можно продолжить любым достаточно гладким образом, поскольку это не влияет на динамику частицы на поверхности связей. Для определенности, положим

$$\lambda_\mu = -J_{\mu\nu}^{-1}[\Phi^\nu, H] \quad (22.10)$$

в некоторой окрестности поверхности связей. Если продолжить множители Лагранжа вне поверхность связей каким либо иным образом, то это изменит только траектории, лежащие вне поверхности связей, которые нас не интересуют. Множители Лагранжа (22.10) подставляем в уравнения (22.8) и решаем задачу Коши для координат и импульсов. Тогда, если в начальный момент времени точка фазового пространства находилась на поверхности связей, то она там и останется при эволюции системы. Это означает, что метод неопределенных множителей Лагранжа позволяет из вариационного принципа для обобщенного действия (22.7) определить множители Лагранжа и эволюцию динамических переменных.

Решение (22.10) для множителей Лагранжа означает, что, если связи выполняются в начальный момент времени для некоторой траектории, то множители Лагранжа всегда можно подобрать таким образом, что связи будут выполнены и во все последующие моменты времени.

Исключим множители Лагранжа (22.10) из уравнений движения (22.8):

$$\begin{aligned} \dot{q}^i &= [q^i, H] - [q^i, \Phi^\mu] J_{\mu\nu}^{-1}[\Phi^\nu, H], \\ \dot{p}_i &= [p_i, H] - [p_i, \Phi^\mu] J_{\mu\nu}^{-1}[\Phi^\nu, H]. \end{aligned} \quad (22.11)$$

Если подставить множители Лагранжа (22.10) в обобщенное действие (22.7), то получим действие, не содержащее множителей Лагранжа:

$$S = \int dt (p\dot{q} - H + [H, \Phi^\mu] J_{\mu\nu}^{-1} \Phi^\nu). \quad (22.12)$$

Это действие приводит к правильным уравнениям движения (22.11) для канонических переменных q, p , если “руками” наложить условия $\Phi = 0$, однако уравнения связей при этом не будут следовать из принципа наименьшего действия. Это связано с тем, что решены не уравнения связей $\Phi = 0$, а уравнения $\dot{\Phi} = 0$. Следовательно, действия (22.12) и (22.7) не эквивалентны. Тем не менее пространство решений уравнений Эйлера–Лагранжа для действия (22.7) совпадает с тем подпространством решений для действия (22.12), которые проходят через какую либо точку поверхности связей и поэтому целиком лежат на ней.

Отметим, что для обобщенного действия (22.7) связи являются частью уравнений Эйлера–Лагранжа. В соответствии с общим правилом (см. раздел 20.4), если их можно однозначно разрешить относительно каких либо переменных, то это решение можно подставить в исходное действие (22.7) и получить при этом эквивалентное эффективное действие для меньшего числа переменных.

Пример 22.2.1. Пусть механическая система описывается двумя парами канонически сопряженных переменных: q^1, p_1 и q^2, p_2 . Предположим, что одна из связей имеет вид

$$\Phi := p_1 + p_2 + f(q) = 0,$$

где $f(q)$ – некоторая функция координат. Тогда решения уравнения связи

$$p_1 = -p_2 - f(q) \quad \text{или} \quad p_2 = -p_1 - f(q)$$

можно подставить в действие, уменьшив тем самым число независимых переменных на единицу. Если связь сложнее, то необходим более тщательный анализ. $\square$

Уравнения (22.11) можно записать в компактном виде с помощью нового важного понятия, которое было введено П. Дираком [148].

Определение. Для любых функций $f, g \in \mathcal{C}^1(\mathbb{N})$ билинейная операция

$$[f, g]_{\text{D}} := [f, g] - [f, \Phi^\mu] J_{\mu\nu}^{-1} [\Phi^\nu, g], \quad (22.13)$$

называется *скобкой Дирака*. $\square$

Скобка Дирака, очевидно, антисимметрична и билинейна. Кроме того, нетрудно проверить, что для нее справедливо правило Лейбница и тождества Якоби. Следовательно, скобка Дирака определяет на фазовом пространстве $\mathbb{N}$ новую пуассонову структуру (см. раздел 19.3). Напомним, что на фазовом пространстве уже существует каноническая пуассонова структура для координат и сопряженных импульсов $[f, g]$. Скобка Дирака определяет на $\mathbb{N}$ вторую *пуассонову структуру Дирака*, которая обладает рядом замечательных свойств для гамильтоновых систем со связями II рода.

Сначала заметим, что уравнения движения (22.8) можно записать в компактном виде

$$\dot{q}^i = [q^i, H]_{\text{D}}, \quad \dot{p}_i = [p_i, H]_{\text{D}}. \quad (22.14)$$

В эти уравнения не входят явно множители Лагранжа, потому что достаточная информация о связях содержится в определении скобки Дирака.

Из определения (22.13) следует, что скобка Дирака каждой связи (22.1) с произвольной функцией $F \in \mathcal{C}^1(\mathbb{N})$ равна нулю:

$$[F, \Phi^\mu]_{\text{D}} = 0.$$

Отсюда вытекает, что диракова пуассонова структура вырождена и все связи второго рода являются функциями Казимира для скобки Дирака. Покажем, что других функционально независимых функций Казимира не существует.

Предложение 22.2.2. На поверхности связей $\Phi = 0$ существует такая система координат y^a , $a = 1, \dots, 2(N - M)$, что выполнены следующие условия:

$$\det [y^a, y^b]_{\Phi=0} \neq 0, \quad [y^a, \Phi^\mu]_{\Phi=0} = 0.$$

Доказательство. Поскольку связи функционально независимы, то их можно выбрать в качестве части координатных функций новой системы координат в некоторой окрестности поверхности связей. Выберем какие либо координаты $\tilde{y}^a$, $a = 1, \dots, 2(N - M)$, на поверхности связей. Тогда совокупность функций $\{\tilde{y}^a, \Phi^\mu\}$ образует систему координат в окрестности поверхности связей. В общем случае $[\tilde{y}^a, \Phi^\mu] \neq 0$.

Тогда введем новые координаты на поверхности связей $y^a = \tilde{y}^a + \lambda_\mu^a \Phi^\mu$, где λ_μ^a – некоторые функции. На поверхности связей $\Phi = 0$ и эти координаты совпадают со старыми. Выберем неизвестные функции λ_μ^a таким образом, чтобы на поверхности связей выполнялись уравнения:

$$[y^a, \Phi^\mu]_{\Phi=0} = [\tilde{y}^a, \Phi^\mu]_{\Phi=0} + \lambda_\nu^a [\Phi^\nu, \Phi^\mu]_{\Phi=0} = 0.$$

Это всегда можно сделать, поскольку связи Φ^μ второго рода. В новой системе координат структурные функции канонической пуассоновой структуры на поверхности связей примут вид

$$J = \begin{pmatrix} [y^a, y^b]_{\Phi=0} & 0 \\ 0 & [\Phi^\mu, \Phi^\nu]_{\Phi=0} \end{pmatrix}.$$

Поскольку пуассонова структура не может быть вырожденной, то $\det [y^a, y^b]_{\Phi=0} \neq 0$. $\square$

В системе координат $\{y^a, \Phi^\mu\}$ из только что доказанного предложения структурные функции дираковой пуассоновой структуры примут вид

$$J_D = \begin{pmatrix} [y^a, y^b]_{\Phi=0} & 0 \\ 0 & 0 \end{pmatrix},$$

поскольку $[y^a, y^b]_{D, \Phi=0} = [y^a, y^b]_{\Phi=0}$. Отсюда следует, что в окрестности связей ранг J_D равен $2(N - M)$. Таким образом, мы доказали следующее утверждение.

Теорема 22.2.1. *Связи второго рода Φ^μ являются функциями Казимира для пуассоновой структуры Дирака (22.13). Все остальные функции Казимира имеют вид $c = c(\Phi)$.*

Используя скобку Дирака, для каждой связи Φ^μ можно построить векторное поле с компонентами

$$X_{\Phi^\mu}^\alpha = [\Phi^\mu, x^\alpha]_D,$$

где $x = (q, p)$ – координаты на фазовом пространстве $\mathbb{N}$. Согласно предложению 19.3.1 коммутаторы этих векторных полей равны нулю:

$$[X_{\Phi^\mu}, X_{\Phi^\nu}] = X_{[\Phi^\mu, \Phi^\nu]_D} = 0.$$

Следовательно, распределение векторных полей $\{X_{\Phi^\mu}\}$, $\mu = 1, \dots, 2M$, находится в инволюции и определяет интегральные подмногообразия в $\mathbb{N}$. Таким образом, постоянные значения функций Казимира слоют фазовое пространство $\mathbb{N}$ на симплектические сечения, одним из которых является поверхность связей $\mathbb{M} \subset \mathbb{N}$.

Каноническая пуассонова структура на всем фазовом пространстве $\mathbb{N}$ не приспособлена для описания систем со связями, потому что сужение скобки Пуассона двух функций на поверхность связей нельзя проводить до вычисления самой скобки:

$$[f, g]_{\Phi=0} \neq [f_{\Phi=0}, g_{\Phi=0}].$$

Например, мы сразу приходим к противоречию, если $f = \Phi^\mu$ и $g = \Phi^\nu$ для некоторых μ и ν . В то же время сужение скобки Дирака на поверхность связей можно проводить до вычисления самой скобки:

$$[f, g]_D|_{\Phi=0} = [f_{\Phi=0}, g_{\Phi=0}]_D. \quad (22.15)$$

Так как поверхность связей является подмногообразием $\mathbb{M} \subset \mathbb{N}$, то ее можно рассматривать, как вложение $\mathbb{M} \hookrightarrow \mathbb{N}$. Тогда равенство (22.15) означает, что вложение $\mathbb{M} \hookrightarrow \mathbb{N}$ является пуассоновым отображением для скобки Дирака, но не для канонической пуассоновой структуры на $\mathbb{N}$.

Поскольку скобки Дирака достаточно для описания эволюции всех переменных, то ее использование приносит существенные упрощения в описание динамики систем со связями II рода.

Скобка Дирака (22.13) определена с помощью канонической пуассоновой структуры на фазовом пространстве $\mathbb{N}$. В разделе 21.11 было показано, что канонические преобразования сохраняют вид канонической скобки Пуассона. Поэтому скобка Дирака двух функций также инвариантна относительно канонических преобразований, т.е. она имеет одинаковый вид в любых координатах на $\mathbb{N}$, связанных между собой каноническим преобразованием.

В физических приложениях важную роль играет специальная система координат на фазовом пространстве, которая строится с учетом вложения $\mathbb{M} \hookrightarrow \mathbb{N}$. Ее существование обеспечивается следующим важным утверждением.

Теорема 22.2.2. *Пусть задано фазовое пространство $\mathbb{N}$ и набор функционально независимых связей второго рода (22.1). Тогда существует такое каноническое преобразование*

$$\{q^i, p_i\} \mapsto \{q^{*a}, p_a^*, Q^A, P_A\}, \quad a = 1, \dots, N - M, \quad A = 1, \dots, M,$$

что набор связей $\Phi = 0$ эквивалентен связям

$$Q = 0, \quad P = 0.$$

При этом координаты q^*, p^* являются канонически сопряженными координатами и импульсами на поверхности связей $\mathbb{M} \subset \mathbb{N}$.

Доказательство. Поскольку связи Φ^μ являются связями второго рода, то согласно теореме Дарбу существуют такие координаты Q^A, P_A , что $[Q^A, P_B] = \delta_B^A$ и уравнения $\Phi = 0$ эквивалентны уравнениям $Q = 0, P = 0$. После этого из теоремы Дарбу 19.5.3 следует существование системы координат q^{*a}, p_a^*, Q^A, P_A . Поскольку скобка Пуассона в новых координатах имеет канонический вид, то преобразование координат является каноническим. $\square$

Эта теорема локальна. Если найдены координаты q^*, p^*, Q, P в явном виде, то ограничение функций на поверхность связей особенно просто: нужно просто положить $Q = 0$ и $P = 0$. В частности,

$$H_E|_{\Phi=0} = H|_{Q=0, P=0}.$$

Поскольку

$$[q^*, Q] = 0, \quad [q^*, P] = 0 \quad \text{и} \quad [p^*, Q] = 0, \quad [p^*, P] = 0,$$

то уравнения движения (22.14) в новой системе координат примут вид

$$\begin{aligned} \dot{q}^* &= [q^*, H_{\text{ph}}], & \dot{p}^* &= [p^*, H_{\text{ph}}] \\ \dot{Q}^A &= 0, & \dot{P}_A &= 0, \end{aligned}$$

где

$$H_{\text{ph}}(q^*, p^*) = H_{\text{e}}|_{\Phi=0} = H(q^*, p^*, Q, P)|_{Q=0, P=0}.$$

Мы видим, что динамика системы N частиц, на которую наложено $2M$ связей второго рода свелась к обычной динамике системы из $N - M$ частиц, на которую уже не наложено никаких связей. По этой причине мы говорим, что система имеет $N - M$ *физических степеней свободы* q^*, p^* . Координаты Q, P описывают *нефизические степени свободы*, поскольку определяют связи. Динамика физических степеней свободы задается *эффективным (физическим) гамильтонианом* H_{ph} , зависящим только от физических координат и импульсов. Координаты на поверхности связей определены неоднозначно. Например, на $\mathbb{M}$ всегда можно совершить каноническое преобразование. Однако размерность поверхности связей (удвоенное число физических степеней свободы) фиксирована и всегда равна $2(N - M)$. Мы говорим, что система N частиц с $2M$ связями II рода описывает $N - M$ физических степеней свободы.

Замечание. На практике найти в явном виде координаты q^*, p^* для физических степеней свободы удастся только в простейших случаях. Кроме того, в теории поля связи часто представляют собой дифференциальные уравнения по пространственным координатам. В этих случаях переход к координатам q^*, p^* задается нелокальными выражениями, как, например, в свободной электродинамике (см. раздел 25.3.3). Поэтому, вычисления обычно проводят в исходных координатах q^i, p_i , используя теорему 22.2.2 для доказательства общих утверждений. $\square$

Таким образом, если на фазовом пространстве $\mathbb{N}$ с канонической пуассоновой структурой задана система связей $\{\Phi^\mu\}$ второго рода, то она определяет некоторое подмногообразие $\mathbb{M} \subset \mathbb{N}$, на котором естественным образом определена каноническая пуассонова структура. Посмотрим на эту задачу с обратной точки зрения. Пусть задано вложение $\mathbb{M} \hookrightarrow \mathbb{N}$, и мы знаем, что подмногообразие $\mathbb{M}$ является фазовым пространством некоторой механической системы с канонической пуассоновой структурой. Возникает естественный вопрос может ли пространство-мишень также быть фазовым пространством? Ниже мы покажем, что ответ на этот вопрос положительный.

Пусть задано вложение

$$\varphi: \mathbb{M} \hookrightarrow \mathbb{N}, \quad \dim \mathbb{M} = 2(N - M), \quad \dim \mathbb{N} = 2N, \quad M < N,$$

фазового пространства $\mathbb{M}$ с канонической пуассоновой структурой. Тогда в окрестности каждой точки $\mathbb{M}$ существует такая система координат y^a , $a = 1, \dots, 2(N - M)$, в которой пуассонова структура имеет канонический вид

$$[y^a, y^b] = \varpi^{-1ab}.$$

Эта пуассонова структура определяет пуассонову структуру на образе $\varphi(\mathbb{M})$ с помощью дифференциала отображения φ_* . В дальнейшем образ отображения вложения мы отождествим с самим многообразием, т.е. положим $\mathbb{M} = \varphi(\mathbb{M}) \subset \mathbb{N}$. Если x^α , $\alpha = 1, \dots, 2N$, – система координат на $\mathbb{N}$, то индуцированная пуассонова структура на $\mathbb{M}$ в координатах $x^\alpha \in \mathbb{N}$ задается антисимметричной матрицей

$$J^{\alpha\beta} = \varpi^{-1ab} \partial_a x^\alpha \partial_b x^\beta, \quad (22.16)$$

где функции $x^\alpha(y)$ описывают вложение. Из свойств произведения матриц следует, что ранг этой пуассоновой структуры равен $2(N - M)$, и поэтому она всегда вырождена. Следовательно, для индуцированной на $\mathbb{N}$ пуассоновой структуры существует

2M независимых функций Казимира c^μ , $\mu = 1, \dots, 2M$. Выберем координаты в окрестности $M \subset \mathbb{N}$ следующим образом

$$\{x^\alpha\} = \{y^a, c^\mu\}.$$

Тогда подмногообразие M задается постоянными значениями функций Казимира $c^\mu = \text{const}$. Отсюда следует, что уравнения $c^\mu = \text{const}$ эквивалентны исходной системе связей (22.1).

Проведенное построение показывает, что на фазовом пространстве $\mathbb{N}$ существует такая система координат, в которой индуцированная пуассонова структура имеет вид

$$J^{\alpha\beta} = \begin{pmatrix} \varpi^{-1ab} & 0 \\ 0 & 0 \end{pmatrix}.$$

Тем самым можно отождествить индуцированную пуассонову структуру с дираковой пуассоновой структурой, а функции Казимира c^μ со связями Φ^μ .

Ранее скобка Дирака была определена через каноническую пуассонову структуру на исходном многообразии $\mathbb{N}$ с помощью связей. В обратную сторону однозначного рецепта определения невырожденной пуассоновой структуры на $\mathbb{N}$ не существует, поскольку ранг индуцированной пуассоновой структуры $2(N - M)$ меньше размерности многообразия $\dim \mathbb{N} = 2N$. Здесь существует много возможностей. Например, можно положить $J^{\alpha\beta} = \varpi^{-1\alpha\beta}$ в координатах $\{y^a, c^\mu\}$. Тогда получим каноническую пуассонову структуру на $\mathbb{N}$.

Если обозначить всю совокупность координат фазового пространства одной буквой $x = \{x^\alpha\} = \{q^i, p_i\}$, $\alpha = 1, \dots, 2N$, то для скобки Дирака справедливо представление

$$[f, g]_D = \frac{\partial f}{\partial x^\alpha} \frac{\partial g}{\partial x^\beta} [x^\alpha, x^\beta]_D.$$

Это равенство проверяется прямой проверкой и соответствует общей формуле для пуассоновой структуры (19.22).

В заключение раздела рассмотрим ситуацию, когда из всей совокупности связей второго рода выделяется некоторое подмножество, которые также представляет собой систему связей второго рода. Если связей достаточно много, то это всегда возможно. Разобьем совокупность связей второго рода $\Phi := \{\Phi^\mu\}$ на две группы:

$$\Phi = (\Psi, \Theta) \quad \Leftrightarrow \quad \{\Phi^\mu\} = \{\Psi^a, \Theta^A\},$$

где группа связей Ψ сама представляет собой набор связей второго рода. Тогда можно считать, что на фазовом пространстве $\mathbb{N}$ задано три пуассоновых структуры: исходная каноническая пуассонова структуры и две вырожденных пуассоновых структуры, соответствующих скобкам Дирака для связей Φ и Ψ . При этом между скобками Дирака вырожденных пуассоновых структур существует связь.

Предложение 22.2.3. *Связи (Ψ, Θ) всегда можно выбрать таким образом, чтобы матрица скобок Пуассона была блочно диагональной*

$$([\Phi^\mu, \Phi^\nu]) = \begin{pmatrix} [\Psi^a, \Psi^b] & 0 \\ 0 & [\Theta^A, \Theta^B] \end{pmatrix} \quad (22.17)$$

в некоторой окрестности U_Φ исходной поверхности связей. То есть связи Ψ и Θ являются связями второго рода по отдельности.

Доказательство. В общем случае матрица скобок Пуассона для связей имеет вид

$$([\Phi^\mu, \Phi^\nu]) = \begin{pmatrix} [\Psi^a, \Psi^b] & [\Psi^a, \Theta^b] \\ [\Theta^a, \Psi^b] & [\Theta^a, \Theta^b] \end{pmatrix}.$$

Сдвинем второй набор связей: $\Theta^A \mapsto \tilde{\Theta}^A := \Theta^A + \lambda_b^A \Psi^b$, где функции λ_b^A удовлетворяют системе уравнений

$$[\Psi^a, \tilde{\Theta}^b] = [\Psi^a, \Theta^b] + \lambda_a^b [\Psi^a, \Psi^b] = 0.$$

Эта система однозначно разрешима в некоторой окрестности $\mathbb{U}_\Phi$ поверхности исходных связей, где $\det [\Phi^\mu, \Phi^\nu] \neq 0$. Тогда в этой окрестности выполнено равенство (22.17), где опущен знак тильды. Теперь заметим, что поверхность связей $\Phi = 0$ является пересечением поверхностей связей $\Psi = 0$ и $\Theta = 0$. Следовательно $\mathbb{U}_\Phi \subseteq \mathbb{U}_\Psi$. Поскольку

$$\det [\Phi^\mu, \Phi^\nu]_{\mathbb{U}_\Phi} = \det [\Psi^a, \Psi^b]_{\mathbb{U}_\Phi} \times \det [\Theta^a, \Theta^b]_{\mathbb{U}_\Phi} \neq 0,$$

то

$$\det [\Psi^a, \Psi^b]_{\mathbb{U}_\Phi} \neq 0 \quad \text{и} \quad \det [\Theta^a, \Theta^b]_{\mathbb{U}_\Phi} \neq 0.$$

То есть связи Ψ и Θ являются связями второго рода в некоторой окрестности исходных связей $\Phi = 0$. $\square$

Предложение 22.2.4. Для скобок Дирака $[f, g]_\Phi$ и $[f, g]_\Psi$ двух функций на фазовом пространстве, определяемых связями II рода Φ и $\Psi \subset \Phi$, в некоторой окрестности поверхностей связей справедливо равенство

$$[f, g]_\Phi = [f, g]_\Psi - [f, \Theta^A]_\Psi J_{AB}^{-1} [\Theta^B, g]_\Psi, \quad (22.18)$$

где

$$J^{AB} := [\Theta^A, \Theta^B].$$

Доказательство. Из равенства (22.17) следует, что в окрестности $\mathbb{U}_\Phi$ скобка Дирака имеет вид

$$[f, g]_\Phi = [f, g] - [f, \Psi^a] J_{ab}^{-1} [\Psi^b, g] - [f, \Theta^A] J_{AB}^{-1} [\Theta^B, g].$$

Отсюда следует равенство (22.18). $\square$

22.3 Гамильтонова динамика частиц со связями I рода

Пусть действие для N точечных частиц в фазовом пространстве $\mathbb{N}$ имеет обычный вид (21.7). Будем искать решение канонических уравнений движения (21.30) при наличии $M < N$ связей на канонические переменные:

$$G_A(q, p) = 0, \quad A = 1, \dots, M < N. \quad (22.19)$$

Другими словами, будем считать, что частицы не могут покинуть $(2N - M)$ -мерное подмногообразие фазового пространства $\mathbb{U} \subset \mathbb{N}$ (поверхность связей), определенного уравнениями (22.19). Из дальнейшего рассмотрения будет ясно, почему мы выбрали число связей первого рода в два раза меньшим числа связей второго рода и почему индекс у связей первого рода пишется внизу.

Предположим, что связи являются достаточно гладкими функциями и функционально независимы на поверхности связей (22.19), т.е. ранг матрицы $\partial G_A / \partial x^\alpha$, где $x = (q^1, \dots, q^N, p_1, \dots, p_N)$ – координаты фазового пространства, максимален и равен M . Будем считать, что рассматриваемая гамильтонова система со связями находится в *инволюции*:

$$[G_A, G_B] = f_{AB}^C G_C \approx 0, \quad (22.20)$$

$$[G_A, H] = v_A^B G_B \approx 0, \quad (22.21)$$

где $f_{AB}^C(q, p) = -f_{BA}^C(q, p)$ и $v_A^B(q, p)$ – некоторые функции от точки фазового пространства $(q, p) \in \mathbb{N}$. Волнистый знак равенства $\approx$ обозначает, что данная функция обращается в нуль на поверхности связей:

$$f \approx 0 \quad \Leftrightarrow \quad f|_{G=0} = 0.$$

При этом говорят, что скобки Пуассона связей между собой и с гамильтонианом *слабо равны нулю*.

Второе условие (22.21) необходимо для того, чтобы связи сохранялись во времени, и гамильтонова динамика была самосогласованной.

Для выполнения равенств (22.20) необходимо, чтобы количество связей не превосходило половины размерности фазового пространства $M \leq N$, что мы потребовали с самого начала (22.19). Действительно, при преобразованиях координат фазового пространства каноническая пуассонова структура не может вырождаться. В силу функциональной независимости функций G_A их можно выбрать в качестве части новых координат. Если $M > N$, то каноническая пуассонова структура оказалась бы вырожденной, что невозможно для фазового пространства.

Определение. Связи (22.19), наложенные на канонические переменные механической системы с гамильтонианом $H(q, p)$, которые удовлетворяют условиям (22.20), (22.21), называются *связями I рода*. $\square$

Мы предполагаем, что все связи I рода учтены в системе уравнений (22.19), т.е. для заданной гамильтоновой системы не существует большего числа функционально независимых соотношений между каноническими переменными, для которых выполнены условия (22.20) и (22.21).

Скобки Пуассона (22.20) по виду напоминают коммутатор базисных векторных полей в алгебре Ли. Однако в рассматриваемом случае допускается нетривиальная зависимость *структурных функций* от точки фазового пространства: $f_{AB}^C = f_{AB}^C(q, p)$.

Напомним, что связи определены неоднозначно. А именно, невырожденные линейные комбинации связей определяют ту же поверхность связей. Это приведет к изменению структурных функций f_{AB}^C и v_A^B , что важно при решении уравнений движения.

Функции f_{AB}^C и v_A^B не могут быть произвольными. По определению, скобки Пуассона удовлетворяют тождеству Якоби. Рассмотрев скобки Пуассона $[[G_A, G_B], G_C]$ и $[[G_A, G_B], H]$ и их циклические перестановки, получим ограничения на структурные функции:

$$\begin{aligned} f_{AB}^D f_{DC}^E + f_{BC}^D f_{DA}^E + f_{CA}^D f_{DB}^E + [f_{AB}^E, G_C] + [f_{BC}^E, G_A] + [f_{CA}^E, G_B] &= 0, \\ f_{AB}^D v_D^C + v_B^D f_{DA}^C - v_A^D f_{DB}^C + [f_{AB}^C, H] + [v_B^C, G_A] - [v_A^C, G_B] &= 0. \end{aligned} \quad (22.22)$$

Тождество Якоби для двойной скобки Пуассона $[[G_A, H], H]$ удовлетворяются автоматически в силу уравнения (22.21).

Если структурные функции постоянны, $f_{AB}^C = \text{const}$, то множество всех линейных комбинаций $a^A G_A$, $a^A \in \mathbb{R}$, образует алгебру Ли с базисом G_A . Тогда уравнение (22.20) задает коммутатор базисных векторов, а соотношение (22.22) сводится к тождеству Якоби для структурных констант.

Предложение 22.3.1. *Для того, чтобы фазовые траектории для гамильтониана H , проходящие через произвольную точку на поверхности связей, целиком лежали на этой поверхности необходимо и достаточно, чтобы были выполнены условия (22.21).*

Доказательство. Совпадает с доказательством предложения 22.2.1. Достаточно только заметить, что

$$\dot{G}_A = [G_A, H] \approx 0, \quad (22.23)$$

и воспользоваться предложением 22.1.1. $\square$

С точки зрения вариационного принципа решение канонических уравнений (21.30) при наличии связей (22.19) является задачей на условный экстремум. Применим к решению этой задачи метод неопределенных множителей Лагранжа. С этой целью построим *полный (total) гамильтониан*, добавив к исходному гамильтониану линейную комбинацию связей первого рода:

$$H_T = H + \lambda^A G_A, \quad (22.24)$$

где $\lambda^A = \lambda^A(q, p, t)$ – неопределенные множители Лагранжа. Соответствующее действие, которое мы назовем *полным*, имеет вид

$$S_T = \int_{t_1}^{t_2} dt (p\dot{q} - H - \lambda^A G_A).$$

Из него вытекают уравнения Эйлера–Лагранжа:

$$\frac{\delta S_T}{\delta p_i} = \dot{q}^i - \frac{\partial H}{\partial p_i} - \lambda^A \frac{\partial G_A}{\partial p_i} = 0, \quad (22.25)$$

$$\frac{\delta S_T}{\delta q^i} = -\dot{p}_i - \frac{\partial H}{\partial q^i} - \lambda^A \frac{\partial G_A}{\partial q^i} = 0,$$

$$\frac{\delta S_T}{\delta \lambda^A} = -G_A = 0, \quad (22.26)$$

где в правых частях уравнений движения (22.25) мы отбросили слагаемые, пропорциональные связям. При варьировании действия мы считаем вариации координат на границе равными нулю, $\delta q(t_{1,2}) = 0$, а вариации импульсов и множителей Лагранжа могут быть произвольны.

Предложение 22.3.2. *Для того, чтобы фазовые траектории для полного действия S_T , проходящие через произвольную точку на поверхности связей, целиком лежали на этой поверхности достаточно, чтобы были выполнены условия (22.20) и (22.21).*

Доказательство. Повторяет доказательство предложения 22.2.1. Достаточно заметить, что

$$\dot{G}_A = [G_A, H] + \lambda^B [G_A, G_B] \approx 0. \quad \square$$

Уравнения движения для систем со связями первого рода (22.25) существенно отличаются от уравнений для систем со связями второго рода. Дело в том, что они не позволяют определить ни одного множителя Лагранжа. Это связано с тем, что уравнения (22.9), из которых находились множители Лагранжа для связей второго рода, в рассматриваемом случае (22.23) на поверхности связей вовсе не содержат множителей Лагранжа. Тем самым любое решение уравнений движения для систем со связями первого рода содержит произвольные функции $\lambda^A(q, p, t)$, число которых совпадает с числом связей. Причиной функционального произвола в решении уравнений движения является калибровочная инвариантность.

Теорема 22.3.1. *Полное действие S_T инвариантно относительно локальных инфинитезимальных преобразований, генерируемых каждой связью первого рода:*

$$\begin{aligned}\delta q^i &:= \epsilon^A [q^i, G_A] = \epsilon^A \frac{\partial G_A}{\partial p_i}, \\ \delta p_i &:= \epsilon^A [p_i, G_A] = -\epsilon^A \frac{\partial G_A}{\partial q^i}, \\ \delta \lambda^A &:= \dot{\epsilon}^A + \epsilon^B v_B^A + \epsilon^B \lambda^C f_{BC}^A,\end{aligned}\tag{22.27}$$

где $\epsilon^A = \epsilon^A(q, p, t)$ – малый параметр преобразований и $\dot{\epsilon} = \partial \epsilon / \partial t$. Параметр преобразований может быть произвольной достаточно гладкой функцией координат, импульсов и времени с нулевыми граничными условиями $\epsilon^A(t_{1,2}) = 0$ для всех q и p .

Доказательство. Вариация действия имеет вид

$$\delta S_T = \int dt \left(-\epsilon^A \frac{\partial G_A}{\partial q^i} \dot{q}^i - \dot{p}_i \epsilon^A \frac{\partial G_A}{\partial p_i} + \epsilon^A [G_A, H] + \epsilon^A \lambda^B [G_A, G_B] - \delta \lambda^A G_A \right).$$

Подставляя сюда вариацию множителей Лагранжа и интегрируя слагаемое $\dot{\epsilon}^A G_A$ по частям с учетом уравнений (22.20) и (22.21), получаем $\delta S_T = 0$. Условия на параметры калибровочных преобразований $\epsilon^A(t_{1,2}) = 0$ достаточны для того, чтобы интегрирование слагаемого $\dot{\epsilon}^A G_A$ по частям было возможно без появления граничных вкладов. $\square$

Согласно второй теореме Нетер локальная инвариантность приводит к линейной зависимости уравнений движения:

$$\frac{\delta S_T}{\delta q} \frac{\partial G_A}{\partial p} - \frac{\delta S_T}{\delta p} \frac{\partial G_A}{\partial q} - \frac{\partial}{\partial t} \left(\frac{\delta S_T}{\delta \lambda^A} \right) + \frac{\delta S_T}{\delta \lambda^B} (v_A^B + \lambda^C f_{AC}^B) = 0,\tag{22.28}$$

в чем нетрудно убедиться и прямой проверкой. В этом случае для любого решения системы уравнений (22.25) будет автоматически выполнено уравнение

$$\dot{G}_A = (v_A^B + \lambda^C f_{AC}^B) G_B.$$

Для этой системы уравнений точка $G = 0$ является неподвижной. Поэтому, если в начальный момент времени фазовая траектория находится на поверхности связей, то для любого решения системы уравнений (22.25) при любых множителях Лагранжа $\lambda^A(q, p, t)$ связи $G_A = 0$ будут автоматически удовлетворены. Это и является причиной возникновения функционального произвола в решениях уравнений движения.

Замечание. Преобразования координат фазового пространства q, p (22.27) с малым постоянным параметром $\epsilon = \text{const}$ являются инфинитезимальной формой канонического преобразования, описанного в примере 21.12.4,

$$q^i(t) \mapsto q^i(t, \epsilon), \quad p_i(t) \mapsto p_i(t, \epsilon),$$

определяемого уравнениями

$$\frac{\partial q^i}{\partial \epsilon^A} = [q^i, G_A], \quad \frac{\partial p_i}{\partial \epsilon^A} = [p_i, G_A]. \quad \square$$

Определение. Модели, функционал действия которых инвариантен относительно локальных преобразований, называют *калибровочными*, а сами преобразования – *калибровочными*. $\square$

Замечание. Это название пришло из физики, где все калибровочные модели, в частности, электродинамика и поля Янга–Миллса обладают этим свойством. $\square$

Тем самым мы доказали, что каждой связи первого рода соответствует локальная инвариантность полного действия, а сама связь является генератором калибровочных преобразований. По аналогии с генераторами группы Ли мы пишем индексы у связей первого рода внизу.

Проанализируем уравнения движения подробнее. Поскольку среди $2N + M$ уравнений Эйлера–Лагранжа (22.25) и (22.26) только $2N$ являются независимыми, то в отличие от задачи на условный экстремум, рассмотренной в разделе 20.1, этих уравнений недостаточно для определения всех неизвестных функций $q(t)$, $p(t)$ и $\lambda(t)$. Допустим, что систему уравнений (22.25) можно решить относительно канонических переменных, для которых поставлена задача Коши: $q(t_1) = q_1$, $p(t_1) = p_1$. Тогда решение этой задачи будет зависеть от M произвольных функций времени, которыми являются множители Лагранжа, и $2N$ постоянных интегрирования. Все постоянные интегрирования определяются начальными данными. В этом случае через одну точку фазового пространства проходит множество фазовых траекторий, которые параметризуется множителями Лагранжа.

С физической точки зрения это означает следующее. Пусть некоторое физическое явление описывается калибровочной моделью. Тогда начальное состояние системы не определяет однозначно последующую эволюцию, что противоречит экспериментальным данным, если не принимать во внимание квантовомеханическую неопределенность. Тем не менее калибровочные модели в настоящее время широко используются в теоретической физике: классическим примером служит электродинамика.

Выход из этого противоречия прост. Для калибровочных моделей вводится постулат: все физические наблюдаемые калибровочно инвариантны, т.е. не зависят от произвольных функций, которые могут содержаться в решении уравнений движения. Отсюда следует, что физические наблюдаемые описываются функциями $f \in \mathcal{C}^k(\mathbb{N})$ на фазовом пространстве $\mathbb{N}$, которые являются калибровочно инвариантными. А именно, мы требуем, чтобы скобка Пуассона каждой наблюдаемой функции со связями обращалась в нуль на поверхности связей:

$$[f, G_A] = d_A^B G_B \approx 0, \quad (22.29)$$

где $d_A^B(q, p)$ – некоторые достаточно гладкие функции канонических переменных. Тогда в уравнении движения для калибровочно инвариантной функции

$$\dot{f} = [f, H] + \lambda^A [f, G_A] \approx [f, H],$$

все слагаемые с множителями Лагранжа обратятся в нуль на поверхности связей и, следовательно, на поверхности связей никакого произвола в эволюции калибровочно инвариантной функции нет.

Проведенное рассмотрение требует комментария, потому что каждой физической наблюдаемой соответствует не одна, а целый класс функций. Поскольку мы рассматриваем динамику частиц на поверхности связей (22.19), то физические наблюдаемые определяются значениями калибровочно инвариантных функций на поверхности связей. Пусть две калибровочно инвариантные функции f_1 и f_2 совпадают на поверхности связей $\mathbb{U}$. Тогда они могут отличаться только на линейную комбинацию связей:

$$f_2 = f_1 + \mu^A G_A, \quad (22.30)$$

где $\mu^A(q, p)$ – некоторые достаточно гладкие функции. Следовательно, все множество калибровочно инвариантных функций разбивается на классы эквивалентности (22.30). При этом каждый класс эквивалентности соответствует одной физической наблюдаемой.

Другими словами, калибровочно инвариантная функция задается на поверхности связей $\mathbb{U} \subset \mathbb{N}$, а затем продолжается на все фазовое пространство в значительной степени произвольным образом. Степень произвола описывается произвольными функциями $\mu^A(q, p)$. При этом физические наблюдаемые не зависят от способа продолжения.

Оценим произвол, с которым калибровочно инвариантная функция может быть задана на поверхности связей $\mathbb{U} \subset \mathbb{N}$. Условие калибровочной инвариантности (22.29) представляет собой систему M дифференциальных уравнений в частных производных первого порядка, которую можно переписать в виде

$$X_{G_A} f = 0,$$

где X_{G_A} – векторное поле, соответствующее связи G_A (19.29). Для этой системы уравнений условиями совместности являются уравнения (22.20). Действительно,

$$[X_{G_A}, X_{G_B}]f = X_{[G_A, G_B]}f = f_{AB}{}^C X_{G_C}f = 0,$$

где мы воспользовались равенством (19.30). Поэтому функция f однозначно определяется начальными данными на некотором собственном подмногообразии $\mathbb{M} \subset \mathbb{U} \subset \mathbb{N}$ размерности $2N - M - M = 2(N - M)$ и существенно зависит только от части координат на $\mathbb{U}$. Подмногообразие $\mathbb{M}$ можно задать с помощью M дополнительных функционально независимых соотношений между координатами и импульсами:

$$F^A(q, p) = 0, \quad (22.31)$$

которые называются *калибровочными условиями*. Эти условия должны удовлетворять неравенству

$$\det [F^A, G_B] \Big|_{F=0, G=0} \neq 0, \quad (22.32)$$

т.к. только в этом случае на $\mathbb{M}$ можно задать начальные данные для системы уравнений (22.29). Условия (22.31) при выполнении (22.32) называются *канонической калибровкой*. Для точечных частиц отличие от нуля определителя (22.32) является необходимым и достаточным условием однозначного определения множителей Лагранжа в полном гамильтониане (22.24). В общем случае функции F^A , определяющие калибровочные условия, могут зависеть также от времени и множителей Лагранжа λ^A .

После фиксирования канонической калибровки на рассматриваемую гамильтонову систему будет наложено $2M$ связей (22.19) и (22.31). Введем для полной совокупности связей и калибровочных условий следующее обозначение

$$\{\Phi^\mu\} = (F^1, \dots, F^M, G_1, \dots, G_M), \quad \mu = 1, \dots, 2M. \quad (22.33)$$

Очевидно, что

$$\det [\Phi^\mu, \Phi^\nu]_{\Phi=0} = \det \begin{pmatrix} [F^A, F^B] & [F^A, G_B] \\ [G_A, F^B] & [G_A, G_B] \end{pmatrix}_{\Phi=0} = \det^2 [F^A, G_B]_{\Phi=0}, \quad (22.34)$$

т.к. $[G_A, G_B] \approx 0$. Поскольку определитель скобок Пуассона для канонических калибровочных условий со связями первого рода, по построению, отличен от нуля (22.32), то полная совокупность связей Φ^μ представляет собой систему связей второго рода, рассмотренную в предыдущем разделе. Таким образом, калибровочные модели в канонической калибровке сведены к гамильтоновым системам со связями второго рода, для которых метод множителей Лагранжа применим в полном объеме. Заметим, что значение скобок Пуассона для канонических калибровочных условий между собой $[F^A, F^B]$ несущественно.

После наложения калибровочных условий возникает обобщенное действие

$$S_E = \int dt (p_i \dot{q}^i - H - \lambda^A G_A - \pi_A F^A) = \int dt (p_i \dot{q}^i - H - \lambda_\mu \Phi^\mu),$$

которое совпадает с выражением (22.7) для систем со связями II рода.

Каноническая калибровка выделяет в фазовом пространстве N единственную траекторию, проходящую через данную точку физического подпространства M . Это происходит благодаря тому, что каноническая калибровка однозначно определяет множители Лагранжа на поверхности связей $\Phi = 0$. Верно также обратное утверждение: произвольный выбор множителей Лагранжа эквивалентен некоторой канонической калибровке. Действительно, при фиксированных множителях Лагранжа граничная задача для уравнений (22.25) имеет единственное решение $x(t) = \{q(t), p(t)\}$. Перепишем для данного решения уравнения (22.25), но теперь уже с неопределенными множителями Лагранжа. В результате получим переопределенную, но совместную систему $2N$ линейных алгебраических уравнений на M множителей Лагранжа. Тогда соответствующие соотношения $x^A = x^A(t)$ можно принять в качестве канонических калибровочных условий. Это следует из того, что, поскольку условия $F^A = x^A - x^A(t)$ позволяют однозначно определить множители Лагранжа, то условие (22.32) выполнено. Напомним, что это условие является необходимым и достаточным для однозначного определения множителей Лагранжа λ^A на поверхности связей, как следует из анализа, проведенного в предыдущем разделе, и (22.34).

Замечание. В приложениях часто рассматриваются неканонические калибровки. Например, лоренцева или временная калибровки в электродинамике не являются каноническими. Они не фиксируют калибровочную свободу полностью, и это создает определенные трудности, например, при квантовании. В электродинамике неканонические калибровочные условия можно использовать, т.к. уравнения движения линейны и их можно проанализировать. В общем случае существенно нелинейных моделей формализм для неканонических калибровок не развит и мы на них останавливаться не будем. $\square$

Таким образом, динамика частиц со связями I рода в канонических калибровках сводится к гамильтоновым моделям со связями II рода. Это сведение не является однозначным, т.к. от калибровочных условий требуется выполнение только неравенств (22.32). Выбор той или иной системы калибровочных условий диктуется рассматриваемой задачей и соображениями простоты. Как правило, исследование калибровочных моделей проводится в различных калибровках, каждая из которых имеет свои преимущества и недостатки.

22.4 Смешанный набор связей

В математической физике часто встречаются модели, содержащие одновременно связи и первого, и второго родов. Общая схема исследования таких моделей следующая.

Допустим, что на фазовом пространстве $\mathbb{N}$, задан некоторый набор связей

$$\Psi^r(q, p) = 0, \quad r = 1, \dots, R. \quad (22.35)$$

Будем считать, что этот набор связей согласован с некоторой механической системой, эволюция которой определяется гамильтонианом $H(q, p)$, и других связей в модели не возникает.

Рассмотрим матрицу скобок Пуассона

$$[\Psi^r, \Psi^s]. \quad (22.36)$$

В общем случае ранг этой матрицы может меняться от точки к точке фазового пространства. Мы предположим, что ранг матрицы (22.36) постоянен на поверхности связей $\Psi = 0$ и равен $2M$. В силу непрерывности он также равен $2M$ в некоторой окрестности $\mathbb{U}_\Psi$ поверхности связей. С помощью линейного преобразования (22.2) перейдем к системе эквивалентных связей, такой, чтобы максимальный минор матрицы (22.36) с отличным от нуля определителем стоял в верхнем левом углу. Тогда, не ограничивая общности, можно считать, что все связи делятся на два класса:

$$\{\Psi^r\} = \{\Phi^\mu, G_A\}, \quad \mu = 1, \dots, 2M, \quad A = 1, \dots, L, \quad 2M + L = R.$$

Матрица скобок Пуассона в общем случае имеет вид

$$([\Psi^r, \Psi^s]) = \begin{pmatrix} [\Phi^\mu, \Phi^\nu] & [\Phi^\mu, G_B] \\ [G_A, \Phi^\nu] & [G_A, G_B] \end{pmatrix},$$

где $\det [\Phi^\mu, \Phi^\nu]_{\mathbb{U}_\Psi} \neq 0$.

Теперь сдвинем связи

$$G_A \mapsto \tilde{G}_A := G_A + \lambda_{A\nu} \Phi^\nu,$$

где $\lambda_{A\nu} = \lambda_{A\nu}(q, p)$ – некоторые функции. Подберем эти функции так, чтобы было выполнено равенство

$$[\tilde{G}_A, \Phi^\mu] = [G_A, \Phi^\mu] + \lambda_{A\nu} [\Phi^\nu, \Phi^\mu] = 0.$$

Эти уравнения для $\lambda_{A\nu}$ имеют однозначное решение, т.к. $\det [\Phi^\mu, \Phi^\nu]_{\mathbb{U}_\Psi} \neq 0$. Поэтому, не ограничивая общности, можно считать, что в окрестности $\mathbb{U}_\Psi$ матрица скобок Пуассона имеет блочно диагональный вид

$$([\Psi^r, \Psi^s])_{\mathbb{U}_\Psi} = \begin{pmatrix} [\Phi^\mu, \Phi^\nu] & 0 \\ 0 & [G_A, G_B] \end{pmatrix}_{\mathbb{U}_\Psi},$$

где у связей G_A мы отбросили знак тильды. По построению, ранг матрицы $[\Phi^\mu, \Phi^\nu]$ максимален и равен $2M$ на поверхности связей. С учетом антисимметрии матрицы $[G_A, G_B]$ отсюда следует, что на поверхности связей $\Psi = 0$ все элементы матрицы $[G_A, G_B]$ обращаются в нуль. Поэтому они имеют вид (22.20). Следовательно, связи G_A являются связями первого рода. Таким образом, всю совокупность связей можно разделить на связи второго и первого родов: $\{\Psi\} = \{\Phi, G\}$. Далее можно действовать с этой совокупностью связей независимо. Каждой связи первого рода соответствует калибровочная инвариантность. Поэтому мы накладываем ровно L канонических калибровочных условий $F^A = 0$. В результате полная система связей и калибровочных условий превращается в систему связей второго рода. Расширенный гамильтониан принимает вид

$$H_E = H + \lambda_\mu \Phi^\mu + \lambda^A G_A + \pi_A F^A. \quad (22.37)$$

Далее, из условия сохранения связей и калибровочных условий во времени однозначно находятся все множители Лагранжа, и таким образом гамильтонова динамика однозначно определена.

Для того, чтобы корректно поставить задачу Коши для гамильтоновых уравнений движения, необходимо задать точку на поверхности связей и калибровочных условий, т.е. такие q и p , которые удовлетворяют $2(M + L)$ условиям:

$$\Phi = 0, \quad G = 0, \quad F = 0.$$

Таким образом, при постановке задачи Коши можно задать произвольно только $N - M - L$ обобщенных координат и столько же обобщенных импульсов. Напомним, что те степени свободы, для которых начальные данные Коши могут быть заданы произвольным образом, называются физическими. Это означает, что гамильтонова система со связями описывает

$$\frac{1}{2} \left\{ \begin{array}{l} \text{Размерность фазового} \\ \text{пространства} \end{array} \right\} - \frac{1}{2} \left\{ \begin{array}{l} \text{Число связей} \\ \text{второго рода} \end{array} \right\} - \left\{ \begin{array}{l} \text{Число связей} \\ \text{первого рода} \end{array} \right\}$$

физических, т.е. распространяющихся степеней свободы при условии, что ранг матрицы (22.36) постоянен на поверхности связей (22.35). Обратим внимание на отсутствие множителя $1/2$ у связей первого рода, что связано с калибровочной инвариантностью.

Возможна ситуация, когда совокупное число связей и калибровочных условий равно размерности фазового пространства, и, следовательно, число физических степеней свободы равно нулю. Если число степеней свободы конечно, то связи и калибровочные условия определяют в фазовом пространстве только некоторый набор отдельных точек. В теории поля связи могут представлять собой дифференциальные уравнения по пространственным переменным и иметь решения, представляющие физический интерес. Такие модели в теории поля часто называют *топологическими*.

Расширенному гамильтониану (22.37) соответствует действие

$$S_E = \int dt (p\dot{q} - H - \lambda_\mu \Phi^\mu - \lambda^A G_A - \pi_A F^A),$$

которое приводит к гамильтоновым уравнениям для q, p и связям. Иногда часть связей и калибровочных условий можно однозначно разрешить относительно нефизических степеней свободы. В этом случае решения можно подставить непосредственно в

действие (а не в уравнения движения) и получить эффективное действие для меньшего числа переменных. Это действие будет эквивалентно исходному действию, т.к. связи и калибровочные условия являются частью уравнений Эйлера–Лагранжа для расширенного действия (см. раздел 20.4).

22.5 Переход от калибровочно инвариантного лагранжиана к гамильтониану

Гамильтонова динамика полей со связями нашла широкое применение в квантовой теории поля, поскольку калибровочные модели составляют основу современных моделей элементарных частиц. На практике исходным пунктом построения соответствующих моделей является лагранжева формулировка, в которой калибровочная инвариантность, а также инвариантность относительно преобразований Лоренца и других групп симметрии более прозрачна. В гамильтоновом формализме время выделено явно, и инвариантность модели далеко не очевидна. Кроме того, многие вычисления, особенно в рамках теории возмущений, удобнее проводить также в лагранжевом формализме, где легче следить за инвариантностью. Поэтому важным моментом в исследовании калибровочных моделей является переход от лагранжевой к гамильтоновой форме теории, которого мы вообще не касались в предыдущих разделах.

В разделе 21.2 был описан переход от лагранжева формализма к гамильтонову с помощью преобразования Лежандра. Этот переход содержит существенное ограничение: функция Лагранжа $L(q, \dot{q}, t)$ должна быть выпукла по скоростям, т.е. ее гессиан $\partial^2 L / \partial \dot{q}^i \partial \dot{q}^j$ положительно определен. В то же время в калибровочных моделях математической физики гессиан вырожден, и с формальной точки зрения переход от лагранжиана к гамильтониану не определен. Ниже мы покажем, что лагранжевы модели допускают гамильтонову формулировку независимо от вида гессиана.

Общая схема перехода от лагранжиана к гамильтониану для механической системы с N степенями свободы следующая. Пусть в конфигурационном пространстве $\mathbb{N}$ задан лагранжиан $L(q, \dot{q}, t)$. По заданному лагранжиану находятся импульсы

$$p_i := \frac{\partial L}{\partial \dot{q}^i}, \quad i = 1, \dots, N. \quad (22.38)$$

Если гессиан вырожден, то эти равенства нельзя разрешить относительно скоростей. Однако справедливы следующие утверждения, в которых q , $\dot{q}$ и p рассматриваются, как независимые функции времени. Для краткости, индекс i там, где это не вызывает недоразумений, мы опустим.

Предложение 22.5.1. *Уравнения Эйлера–Лагранжа*

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0 \quad (22.39)$$

эквивалентны уравнениям Гамильтона

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}, \quad (22.40)$$

если выполнены следующие равенства:

$$p = \frac{\partial L}{\partial \dot{q}}, \quad (22.41)$$

$$H = p\dot{q} - L. \quad (22.42)$$

Доказательство. Допустим, что выполнены уравнения Гамильтона (22.40) и условия (22.41), (22.42). Из определения (22.42) следует равенство

$$\frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q}. \quad (22.43)$$

Теперь подставим определение импульсов (22.41) во второе уравнение (22.40) и получим уравнения Эйлера–Лагранжа (22.39).

Обратно. Пусть выполнены уравнения Эйлера–Лагранжа (22.39) и условия (22.41), (22.42). Подстановка равенства (22.43) и определения обобщенных импульсов (22.41) в уравнения Эйлера–Лагранжа (22.39) дает второе уравнение Гамильтона (22.40). Из определения гамильтониана (22.42) следует равенство

$$\frac{\partial H}{\partial p} = \dot{q} - \frac{\partial L}{\partial p} = \dot{q},$$

которое совпадает с первым уравнением Гамильтона (22.40). $\square$

Предложение 22.5.2. *Гамильтониан (22.42), в котором обобщенные импульсы и координаты связаны соотношением (22.41), не зависит от скоростей $\dot{q}$, т.е. $H = H(q, p, t)$.*

Доказательство. Продифференцируем определение (22.42) по скоростям:

$$\frac{\partial H}{\partial \dot{q}} = p - \frac{\partial L}{\partial \dot{q}} = 0,$$

т.к. выполнено условие (22.41). $\square$

В доказательствах приведенных утверждений вид гессиана никак не использован, т.к. мы нигде не разрешали равенства (22.38) относительно скоростей. Поэтому гамильтонова формулировка лагранжевой теории возможна также для законечномерных и вырожденных гессианов. Если гессиан положительно или отрицательно определен, то существует преобразование Лежандра и определение обобщенных импульсов (22.38) можно, в принципе, разрешить относительно скоростей $\dot{q}$. Если скорости как функции координат и импульсов находятся явно, то, подставляя их в определение (22.42), найдем явный вид гамильтониана. В случае вырожденного гессиана в общем случае скорости нельзя выразить через координаты и импульсы. Поэтому регулярный способ построения гамильтониана отсутствует. Тем не менее гамильтониан существует и, согласно предложению 22.5.2, не зависит от скоростей. В таких случаях гамильтониан (22.42) находится “методом пристального всматривания”. Во всех известных автору случаях этот метод работает и позволяет найти явный вид гамильтониана. Для вырожденных гессианов в теории возникают связи, т.е. некоторые соотношения между координатами и импульсами в фиксированный момент времени. Эти связи являются частью уравнений Эйлера–Лагранжа, и их решения можно подставлять в действие. Поэтому гамильтониан в лагранжевых моделях с вырожденным гессианом определен не однозначно, а с точностью до связей.

Определение. Модель с вырожденным гессианом,

$$\det \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i} = 0, \quad (22.44)$$

называется *вырожденной*. $\square$

Пример 22.5.1. Пусть лагранжиан зависит от двух полей $\{q^i\} = \{\varphi, A\}$ и имеет вид

$$L := \frac{1}{2}(\dot{\varphi} - A)^2, \quad S := \int dt L. \quad (22.45)$$

Гессиан модели равен

$$\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^i} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

и, очевидно, вырожден. Уравнения Эйлера–Лагранжа для рассматриваемой вырожденной модели

$$\begin{aligned} \frac{\delta S}{\delta \varphi} &= -\ddot{\varphi} + \dot{A} = 0, \\ \frac{\delta S}{\delta A} &= -\dot{\varphi} + A = 0. \end{aligned}$$

зависимы, т.к. первое уравнение получается из второго дифференцированием по времени.

Рассмотрим задачу Коши для полученной системы уравнений. Из второго уравнения следует, что данные Коши (имеются в виду значения функций и их производных) не могут быть заданы произвольным образом. Пусть заданы допустимые (т.е. совместимые с уравнениями Эйлера–Лагранжа) начальные условия:

$$\varphi(0) = a, \quad A(0) = b, \quad \dot{\varphi}(0) = b, \quad \dot{A}(0) = c,$$

где a, b, c – произвольные постоянные. Нетрудно проверить, что общее решение задачи Коши имеет вид

$$\varphi = a + bt + \frac{1}{2}ct^2 + \int_0^t dt' f(t'), \quad A = b + ct + f(t), \quad (22.46)$$

где $f(t)$ – произвольная функция, удовлетворяющая начальным условиям $f(0) = 0$ и $\dot{f}(0) = 0$. Таким образом, решение задачи Коши для рассматриваемой модели содержит функциональный произвол. Это произошло потому, что среди уравнений Эйлера–Лагранжа для двух неизвестных функций $\varphi(t)$ и $A(t)$ только одно является независимым. В свою очередь зависимость уравнений движения

$$\frac{\delta S}{\delta \varphi} - \frac{d}{dt} \frac{\delta S}{\delta A} = 0$$

является следствием второй теоремы Нетер, поскольку и действие, и лагранжиан инвариантны относительно калибровочных преобразований

$$\varphi \mapsto \varphi + \alpha, \quad A \mapsto A + \dot{\alpha},$$

где $\alpha(t)$ – произвольная функция времени.

Рассмотренная модель имеет простой геометрический смысл. Функция $A(t)$ определяет локальную форму связности на тривиальном главном расслоении $\mathbb{P}(\mathbb{R}, \pi, \mathbb{R}) \approx \mathbb{R} \times \mathbb{R}$, где $t \in \mathbb{R}$ – база, а структурной группой является группа сдвигов $\mathbb{R}$ вещественной прямой. Кривизна этой связности равна нулю, т.к. база одномерна. Поле $\varphi(t)$ – это сечение тривиального ассоциированного расслоения $\mathbb{E}(\mathbb{R}, \pi_{\mathbb{E}}, \mathbb{R}, \mathbb{R}, \mathbb{P}) \approx \mathbb{R} \times \mathbb{R}$. Типичным слоем этого расслоения является прямая $\mathbb{R}$, на которой структурная группа

действует сдвигами. Лагранжиан (22.45) представляет собой квадрат ковариантной производной, где $dtA(t)$ – локальная форма связности.

Перейдем к гамильтонову формализму. Импульсы, сопряженные полям φ и A , равны

$$p := \frac{\partial L}{\partial \dot{\varphi}} = \dot{\varphi} - A, \quad P := \frac{\partial L}{\partial \dot{A}} = 0.$$

Данные соотношения позволяют определить только одну скорость $\dot{\varphi} = p + A$, а равенство $P = 0$ является связью. При этом скорость $\dot{A}$ остается произвольной. Гамильтониан системы с точностью до интегрирования по частям имеет вид

$$H = p\dot{\varphi} + P\dot{A} - L = \frac{1}{2}p^2 + pA - \dot{A}P,$$

где последнее слагаемое можно отбросить ввиду наличия связи. Поскольку $[P, H] = 0$ при $P = 0$ и $[P, P] = 0$, то равенство $P = 0$ представляет собой первичную связь первого рода. Других связей в теории не возникает. Полный гамильтониан получается из H после добавления связи первого рода, $H_T := H + \lambda P$. Таким образом, скорость $\dot{A}$ в рассматриваемой модели играет роль множителя Лагранжа.

Для однозначного фиксирования решения задачи Коши необходимо наложить каноническую калибровку, например, $A = 0$. Тогда произвольная функция однозначно фиксируется $f(t) = -b - ct$, и решение (22.46) задачи Коши примет вид $\varphi = b = \text{const}$. $\square$

Сначала схематично рассмотрим вырожденные модели в лагранжевом формализме, не вдаваясь в анализ условий, при которых описанные операции имеют смысл. Распишем уравнения Эйлера–Лагранжа (22.39) подробнее:

$$\ddot{q}^j \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i} + \dot{q}^j \frac{\partial^2 L}{\partial q^j \partial \dot{q}^i} + \frac{\partial^2 L}{\partial t \partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0. \quad (22.47)$$

Поскольку гессиан вырожден, то данную систему уравнений нельзя разрешить относительно вторых производных $\ddot{q}^i$, и теорема о существовании и единственности решения задачи Коши (в классе вещественно аналитических функций) неприменима.

Допустим, что ранг гессиана постоянен во всем конфигурационном пространстве и равен $N - M_1 < N$. Тогда у гессиана существует M_1 линейно независимых в каждой точке конфигурационного пространства собственных векторов $\theta^{(\mu_1)i}(q, p, t)$, где индекс $\mu_1 = 1, \dots, M_1$ нумерует эти векторы, соответствующих нулевому собственному значению:

$$\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \theta^{(\mu_1)j} = 0.$$

Умножив уравнения Эйлера–Лагранжа (22.47) на $\theta^{(\mu_1)i}$, получим M_1 независимых условий

$$\Theta^{(1)} = \{\Theta^{(\mu_1)}(q, \dot{q}, t)\} = 0, \quad \mu_1 = 1, \dots, M_1, \quad (22.48)$$

где

$$\Theta^{(\mu_1)} := \theta^{(\mu_1)i} \left(\dot{q}^j \frac{\partial^2 L}{\partial q^j \partial \dot{q}^i} + \frac{\partial^2 L}{\partial t \partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right), \quad (22.49)$$

на обобщенные координаты и скорости, которые должны быть удовлетворены во все моменты времени. Наличие этих условий говорит нам о том, что при постановке задачи Коши (имеются в виду значения всех координат и скоростей) для системы

уравнений (22.47) нельзя задать все координаты и скорости произвольным образом. Для того, чтобы задача Коши имела решение, начальные данные обязаны удовлетворять условиям (22.49). Эти условия называются *первичными связями в лагранжевом формализме*. Однако выполнения первичных связей недостаточно для существования решения задачи Коши. Действительно, допустим, что первичные связи удовлетворены в начальный момент времени, тогда нет никакой гарантии в том, что условия (22.49) будут выполнены для всех $t > 0$.

Для нахождения полного набора необходимых условий на начальные данные необходимо продифференцировать условия (22.49) по времени и исключить вторые производные с помощью уравнений Эйлера–Лагранжа. В результате может получиться тождество $0=0$ или некоторая линейная комбинация первичных связей (22.49). В этом случае никаких дополнительных соотношений между скоростями и импульсами не возникает. Однако в общем случае возможно появление новых функционально независимых от первичных связей (22.49) условий на координаты и скорости, которые называются вторичными связями. Обозначим их совокупность

$$\Theta^{(2)} = \{\Theta^{(\mu_2)}(q, \dot{q}, t)\} = 0, \quad \mu_2 = 1, \dots, M_2.$$

Для корректной постановки задачи Коши этот процесс необходимо продолжить, и продифференцировать вторичные связи по времени. В результате возможно появление третичных функционально независимых связей, которые мы обозначим $\Theta^{(\mu_3)}$, $\mu_3 = 1, \dots, M_3$. Этот процесс рано или поздно оборвется, скажем, на шаге k , т.к. фазовое пространство имеет размерность $2N$ и, следовательно, допускает не более $2N$ функционально независимых условий. В дальнейшем, для краткости, все связи, за исключением первичных, будем называть вторичными.

В принципе, возможно, что на некотором этапе вычисления связей возникнет противоречие типа $1 = 0$. В этом случае уравнения Эйлера–Лагранжа являются противоречивыми, и мы их рассматривать не будем.

После нахождения всех функционально независимых связей в лагранжевом формализме можно ставить задачу Коши. Итак, для того, чтобы задача Коши для системы уравнений Эйлера–Лагранжа (22.47) имела решение необходимо, чтобы начальные данные удовлетворяли всем связям

$$\Theta := \{\Theta^{(1)}, \Theta^{(2)}, \dots, \Theta^{(k)}\} = 0,$$

число которых равно $M_1 + \dots + M_k$, т.к. они должны быть выполнены во все моменты времени. Мы видим, что с точки зрения постановки задачи Коши все связи равноправны, и их деление на первичные и вторичные отражает лишь процесс их получения. Если найден полный набор связей, то об их делении на первичные и вторичные можно забыть. Нахождение всех связей в какой либо модели зависит от конкретного вида лагранжиана и содержит элементы искусства.

Перейдем к анализу вырожденных моделей в гамильтоновом формализме. Для простоты предположим, что лагранжиан и, следовательно, гамильтониан не зависят от времени явно. Рассмотрим определение импульсов (22.38) как систему уравнений на скорости. По теореме о неявной функции система уравнений (22.38) локально разрешима относительно скоростей тогда и только тогда, когда гессиан невырожден (22.44). Поэтому для вырожденных моделей не все скорости $\dot{q}^i$ можно выразить через координаты и сопряженные импульсы. Допустим, что ранг гессиана постоянен и равен $N - M_1 < N$ во всем касательном расслоении к конфигурационному пространству $(q, \dot{q}) \in T(M)$. Тогда в теории возникают M_1 функционально независимых *первичных*

связей, совокупность которых обозначим

$$\Phi^{(1)} := \{\Phi^{\mu_1}(q, p)\} = 0, \quad \mu_1 = 1, \dots, M_1.$$

Отметим, что, если лагранжиан квадратичен по скоростям, то все первичные связи линейны по импульсам.

Теперь рассмотрим эволюцию первичных связей во времени. Для постановки вариационной задачи для системы со связями в гамильтоновом формализме используем метод неопределенных множителей Лагранжа. Поэтому модифицируем гамильтониан, добавив к нему первичные связи,

$$H^{(1)} := H + \lambda_{\mu_1} \Phi^{\mu_1}.$$

Тогда условия сохранения первичных связей во времени примут вид

$$\dot{\Phi}^{(1)} = [\Phi^{(1)}, H^{(1)}] = 0. \quad (22.50)$$

Для того, чтобы модель была самосогласована, первичные связи должны сохраняться, и, следовательно, правая часть (22.50) должна быть равна нулю. В результате получится некоторая система уравнений на канонические переменные и множители Лагранжа. Если эта система уравнений определяет все множители Лагранжа, число которых совпадает с числом уравнений, то никаких дополнительных связей на координаты и импульсы не возникает (связи второго рода). Возможна также ситуация, когда уравнения (22.50) определяют только часть лагранжевых множителей (возможно, ни одного), и возникает некоторое соотношение между каноническими переменными, которое обращается в нуль на поверхности первичных связей. Тогда также не возникает никаких дополнительных ограничений на канонические переменные. Однако не исключен случай, когда на поверхности первичных связей (22.50) возникают новые дополнительные ограничения. Выделим среди них функционально независимые между собой и с первичными связями условия

$$\Phi^{(2)} := \{\Phi^{\mu_2}(q, p)\} = 0, \quad \mu_2 = 1, \dots, M_2,$$

которые называются *вторичными* связями. Как и ранее, модифицируем гамильтониан, добавив вторичные связи,

$$H^{(2)} := H + \lambda_{\mu_1} \Phi^{\mu_1} + \lambda_{\mu_2} \Phi^{\mu_2}.$$

Условие сохранения вторичных связей во времени

$$\dot{\Phi}^{(2)} = [\Phi^{(2)}, H^{(2)}] = 0$$

может привести к возникновению M_3 третичных связей $\Phi^{(3)} = 0$, и т.д. Поскольку фазовое пространство $\mathbb{N}$ конечномерно, то эта процедура оборвется на конечном числе шагов, которое мы обозначим через $k \leq 2N$. В результате мы получим полную систему связей в гамильтоновом формализме

$$\{\Phi^\mu\} = \{\Phi^{(1)}, \Phi^{(2)}, \dots, \Phi^{(k)}\}, \quad \mu = 1, \dots, M_1 + \dots + M_k, \quad (22.51)$$

которые необходимо добавить к полному гамильтониану

$$H_T := H + \lambda_\mu \Phi^\mu.$$

Процесс нахождения всех функционально независимых связей зависит от вида гамильтониана и первичных связей, вид которых определен неоднозначно. Поэтому общего алгоритма получения связей в настоящее время нет. Эта задача бывает сложна и требует перебора многих вариантов.

Поскольку теория содержит связи, которые должны удовлетворяться во все моменты времени, то начальные данные также должны удовлетворять связям (22.51). Это является необходимым условием существования решения задачи Коши. В общем случае мы не можем гарантировать существование и единственность решения этой задачи, но необходимое условие получено. С этой точки зрения деление связей на первичные и вторичные несущественно и, как и в лагранжевом формализме, только отражает последовательность появления связей. После того, как все связи получены, об их делении на первичные и вторичные можно забыть. Следующий шаг состоит в анализе связей и их разделении на связи первого и второго рода, как это было описано в предыдущем разделе.

22.6 Гамильтонова динамика в теории поля

Рассмотрим лоренцево многообразие $(\mathbb{M}, g)$ с координатами

$$\{x^\alpha\} = \{x^0, \mathbf{x}\} = \{x^0, x^\mu\}, \quad \alpha = 0, 1 \dots n-1, \quad \mu = 1, \dots, n-1.$$

Предположим, что координата $x^0 = t$ является временем и все сечения $t = \text{const}$ пространственноподобны. Для простоты будем считать, что пространство-время представляет собой прямое произведение вещественной прямой $t \in \mathbb{R}$ (время) на $(n-1)$ -мерное многообразие $\mathbf{x} \in \mathbb{U}$ (пространство). Рассмотрим действие (20.1) с лагранжианом, зависящим от некоторого набора полей $\varphi = \{\varphi^a(t, \mathbf{x})\}$, $a = 1 \dots N$, и их первых производных

$$S[\varphi] = \int_{\mathbb{M}} dt d\mathbf{x} \mathcal{L}(\varphi, \dot{\varphi}, \hat{\partial}\varphi) = \int dt L, \quad (22.52)$$

где точка обозначает дифференцирование по времени ($\dot{\varphi}^a := \partial\varphi^a/\partial t$), шляпка обозначает частные производные полей только по пространственным координатам ($\hat{\partial}\varphi := \{\partial_\mu\varphi\}$), и явно выделено интегрирование по времени. Функция

$$L := \int_{\mathbb{U}} d\mathbf{x} \mathcal{L} \quad (22.53)$$

называется *лагранжианом*, а $\mathcal{L}$ – *лагранжевой плотностью*. Для дальнейшего рассмотрения возможная явная зависимость лагранжевой плотности от координат несущественна, и мы ее опускаем. Кроме этого, лагранжеву плотность в дальнейшем, для краткости, мы будем также называть лагранжианом, поскольку из контекста всегда ясно о чем идет речь.

В общем случае уравнения движения (уравнения Эйлера–Лагранжа) имеют второй порядок по времени. В теории поля для уравнений движения ставится обычно смешанная задача. Для полевых переменных решается задача Коши по времени при заданных граничных условиях на пространственной границе:

$$\varphi|_{t=0} = \varphi_0, \quad \dot{\varphi}|_{t=0} = \varphi_1, \quad \varphi|_{x \in \partial\mathbb{U}} = \varphi_{\partial\mathbb{U}}.$$

Такая формулировка полевой динамики называется *лагранжевой*.

Помимо лагранжевой важную роль в математической физике играет *каноническая* или *гамильтонова* формулировка теории поля. Каноническая формулировка явно нарушает лоренц и общекоординатную инвариантность, но в то же время делает теорию гибче при исследовании задачи Коши. Помимо этого гамильтонова формулировка является основой для канонического квантования, которое в настоящее время рассматривается, как наиболее последовательный способ перехода от классической к квантовой теории поля.

Для канонической формулировки теории поля нам понадобится пространственная $(n - 1)$ -мерная δ -функция и ее производные, определенные на пространственных сечениях $\mathbb{U}$. Для этого нам достаточно рассматривать класс полей, компоненты которых являются достаточно гладкими функциями на $\mathbb{U}$ для всех моментов времени $t \in \mathbb{R}$. Кроме того, мы предполагаем, что поля ведут себя таким образом, что все встречающиеся функционалы определены. В дальнейшем мы не будем останавливаться на этом вопросе, т.к. функциональные пространства зависят от выбора лагранжиана и постановки задачи. Этот вопрос требует исследования в каждом конкретном случае.

Переход от лагранжевой формулировки к гамильтоновой представляет собой (обобщенное) преобразование Лежандра (см. раздел 21.1) лагранжиана (22.53) по скоростям $\dot{\varphi}$. Гамильтоновы уравнения движения имеют первый порядок по времени, при этом второй порядок уравнений по пространственным производным сохраняется. По аналогии с механикой функции φ^a называются *обобщенными координатами*, а $\dot{\varphi}^a$ – *обобщенными скоростями*. Обобщенные импульсы, сопряженные к φ^a , определим соотношением

$$p_a(x) = p_a(t, \mathbf{x}) := \frac{\delta L}{\delta \dot{\varphi}^a} = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}^a}. \quad (22.54)$$

При этом функции φ^a и $\dot{\varphi}^a$, от которых зависит лагранжиан, рассматриваются, как независимые переменные. Здесь и в дальнейшем все вариации полей считаются финитными функциями на $\mathbb{U}$, и граничные члены не рассматриваются. Это соответствует не зависящим от времени граничным условиям

$$\varphi|_{\mathbf{x} \in \partial \mathbb{U}} := \varphi_{\partial \mathbb{U}}$$

на пространственной границе $\partial \mathbb{U}$.

Иными словами, полевую теорию мы рассматриваем как механическую систему с бесконечным числом переменных (обобщенных координат), которые параметризуются точками пространства $\mathbf{x} \in \mathbb{U}$ и индексом a . Этому соответствует замена частных производных на вариационные, а суммирование по переменным включает в себя интегрирование по пространству.

В дальнейшем, для краткости, мы часто будем писать $\varphi(\mathbf{x})$ вместо $\varphi(t, \mathbf{x})$, подразумевая, что все поля зависят от времени t , как от параметра и рассматриваются в один и тот же момент времени.

Преобразование Лежандра по скоростям приводит к следующей *функции Гамильтона* или *гамильтониану*

$$H(\varphi, \hat{\partial}\varphi, p) := \int d\mathbf{x} \left[p_a \dot{\varphi}^a - \mathcal{L}(\varphi, \hat{\partial}\varphi, \dot{\varphi}) \right] = \int d\mathbf{x} \mathcal{H}(\varphi, \hat{\partial}\varphi, p). \quad (22.55)$$

Функция $\mathcal{H}(\varphi, \hat{\partial}\varphi, p)$ называется *гамильтоновой плотностью*. В дальнейшем гамильтонову плотность мы также будем называть гамильтонианом, т.к. из контекста всегда ясно, о чем идет речь.

Преобразование Лежандра определено в том случае, если квадратичная форма

$$\int d\mathbf{x} d\mathbf{x}' \frac{\delta^2 L}{\delta \dot{\varphi}^a(\mathbf{x}) \delta \dot{\varphi}^b(\mathbf{x}')} f^a(\mathbf{x}) f^b(\mathbf{x}') = \int d\mathbf{x} \frac{\partial^2 \mathcal{L}}{\partial \dot{\varphi}^a(\mathbf{x}) \partial \dot{\varphi}^b(\mathbf{x})} f^a(\mathbf{x}) f^b(\mathbf{x}) \quad (22.56)$$

положительно определена для всех φ^a и $\partial_\mu \varphi^a$. Предыдущее равенство справедливо, т.к. гессиан, который равен второй вариационной производной от лагранжиана, пропорционален δ -функции:

$$\frac{\delta^2 L}{\delta \dot{\varphi}^a(\mathbf{x}) \delta \dot{\varphi}^b(\mathbf{x}')} = \frac{\partial^2 \mathcal{L}}{\partial \dot{\varphi}^a(\mathbf{x}) \partial \dot{\varphi}^b(\mathbf{x})} \delta(\mathbf{x}' - \mathbf{x}).$$

Пусть задано два функционала

$$F := \int_{\mathbb{U}} d\mathbf{x} \mathcal{F}(\varphi, \hat{\partial}\varphi, p, \hat{\partial}p) \quad \text{и} \quad G := \int_{\mathbb{U}} d\mathbf{x} \mathcal{G}(\varphi, \hat{\partial}\varphi, p, \hat{\partial}p). \quad (22.57)$$

Предположим, что все поля и их производные достаточно быстро убывают при приближении к границе $\partial\mathbb{U}$, так, чтобы в выражениях для функционалов можно было интегрировать по частям, отбрасывая все граничные члены. Определим для них скобку Пуассона следующим равенством

$$[F, G] := \int d\mathbf{x} \left(\frac{\delta F}{\delta \varphi^a(x)} \frac{\delta G}{\delta p_a(x)} - \frac{\delta F}{\delta p_a(x)} \frac{\delta G}{\delta \varphi^a(x)} \right), \quad (22.58)$$

где

$$\begin{aligned} \frac{\delta F}{\delta \varphi^a} &= \frac{\partial \mathcal{F}}{\partial \varphi^a} - \partial_\mu \frac{\partial \mathcal{F}}{\partial (\partial_\mu \varphi^a)}, \\ \frac{\delta F}{\delta p_a} &= \frac{\partial \mathcal{F}}{\partial p_a} - \partial_\mu \frac{\partial \mathcal{F}}{\partial (\partial_\mu p_a)}, \end{aligned}$$

и аналогичные равенства для функциональных производных от G . Эта скобка, очевидно, билинейна и антисимметрична. Нетрудно проверить, что она удовлетворяет правилу Лейбница и тождеству Якоби. Тем самым скобка Пуассона определяет пуассонову структуру на множестве функционалов вида (22.57).

Поскольку

$$\varphi^a(\mathbf{x}) = \int_{\mathbb{U}} d\mathbf{y} \varphi^a(\mathbf{y}) \delta(\mathbf{y} - \mathbf{x}) \quad \text{и} \quad p_a(\mathbf{x}) = \int_{\mathbb{U}} d\mathbf{y} p_a(\mathbf{y}) \delta(\mathbf{y} - \mathbf{x}),$$

то справедливы канонические одновременные скобки Пуассона между полями и сопряженными импульсами:

$$\begin{aligned} [\varphi^a(\mathbf{x}), p_b(\mathbf{y})] &= \delta_b^a \delta(\mathbf{x} - \mathbf{y}), \\ [\varphi^a(\mathbf{x}), \varphi^b(\mathbf{y})] &= 0, \\ [p_a(\mathbf{x}), p_b(\mathbf{y})] &= 0. \end{aligned}$$

Скобки Пуассона между полевыми переменными, взятыми в различные моменты времени, не определены.

Нетрудно проверить, что уравнения Эйлера–Лагранжа для действия (22.52) эквивалентны уравнениям Гамильтона:

$$\dot{\varphi}^a = [\varphi^a, H], \quad \dot{p}_a = [p_a, H]. \quad (22.59)$$

В общем случае эволюция во времени произвольного функционала вида (22.57) задается его скобкой Пуассона с гамильтонианом:

$$\dot{F} = [F, H].$$

Выше мы обобщили гамильтонову механику систем точечных частиц на теорию поля в случае, когда гессиан модели невырожден и определено преобразование Лежандра по скоростям. Примеры дают вещественное и комплексное скалярные поля, рассмотренные в разделах 25.1.1 и 25.2.1.

Определение. Модель теории поля называется *сингулярной*, если гессиан

$$\frac{\partial^2 \mathcal{L}}{\partial \dot{\varphi}^a(\mathbf{x}) \partial \dot{\varphi}^b(\mathbf{x})}$$

вырождена хотя бы в одной точке $\mathbf{x} \in \mathbb{U}$. □

В современной теории поля важную роль играют калибровочные модели, которые являются вырожденными. В этом случае в гамильтоновой формулировке возникают связи на канонические переменные, которые выделяют в бесконечномерном фазовом пространстве некоторое подпространство. В настоящее время гамильтонов анализ полевых моделей со связями развит недостаточно полно. На практике действуют по аналогии с механикой точечных частиц. Как и ранее, связи делятся на связи первого и второго рода, а также на первичные и вторичные. В большинстве случаев такой подход дает удовлетворительную гамильтонову формулировку теории поля. Примерами со связями первого рода являются электромагнитное поле (раздел 25.3.3) и поле Янга–Миллса (раздел 25.5.2). К связям первого рода приводит также общая теория относительности, гамильтонова формулировка которой дана в главе 24. Для поля Прокá (раздел 25.4.1) и спинорного поля (раздел 27.1.5) возникает гамильтонова формулировка со связями второго рода.

Существенным отличием вырожденной теории поля от механики точечных частиц является то обстоятельство, что связи в общем случае являются не алгебраическими уравнениями, а уравнениями в частных производных по пространственным координатам. В этих случаях возникают важные вопросы о разрешимости связей на пространственных сечениях $\mathbb{U}$ при заданных граничных условиях. Особенно, если пространственные сечения топологически нетривиальны.

Сравнение гамильтониана (22.55) с определением тензора энергии-импульса (23.170) показывает, что гамильтонова плотность совпадает с $(0, 0)$ компонентой энергии-импульса: $\mathcal{H} = T_0^0$. Численное значение гамильтоновой плотности $\mathcal{H}$ хотелось бы назвать плотностью энергии системы полей. При этом возникает неоднозначность. Уравнения движения не изменятся, если к $\mathcal{H}$ добавить частную производную от произвольной функции полей. При фиксированных граничных условиях это меняет полную энергию на постоянную величину, что допустимо. Однако плотность энергии при этом может существенно меняться.

При гамильтоновой формулировке теории поля может возникнуть еще одна трудность. В случае конечного числа частиц все связи делятся на связи первого и второго родов. В теории поля могут появиться связи, которые нельзя отнести ни к одному из указанных типов. С одной стороны, скобка Пуассона связей не обращается в нуль на поверхности связей, а, с другой стороны, “матрица” скобок Пуассона для связей необратима. В разделе 22.9 будет рассмотрен пример, где скобка Пуассона двух связей равна производной δ -функции по пространственной координате. Она отлична от

нуля на поверхности связей и, следовательно, не относится к связям I рода. С другой стороны, если рассматривать δ' как оператор в функциональном пространстве, то он необратим, поскольку имеет нетривиальное ядро (все постоянные функции отображаются в нуль). Поэтому такие связи нельзя отнести к связям II рода, поскольку они не позволяют однозначно определить множители Лагранжа из условия сохранения связей во времени. Общая теория для моделей такого типа в настоящее время не развита.

22.7 Калибровочная модель нерелятивистской частицы

В настоящем разделе мы рассмотрим динамику точечной частицы с точки зрения калибровочных моделей и покажем трудности в определении энергии, которые при этом возникают.

Рассмотрим движение точечной частицы в фазовом пространстве $\mathbb{T}^*(\mathbb{R}^N)$ с координатами $q^* = \{q^{*a}\}$ и $p^* = \{p_a^*\}$, где $a = 1, \dots, N$. Пусть динамика частицы задана гамильтонианом $H^*(q^*, p^*)$, не зависящим явно от времени. Действие такой частицы имеет обычный вид

$$S = \int_{t_1}^{t_2} dt (\dot{q}^{*a} p_a^* - H^*), \quad (22.60)$$

где точка обозначает дифференцирование по времени t . Будем считать, что все координаты имеют фиксированные значения на границе $t = t_{1,2}$. Для каждой траектории частицы $q^*(t), p^*(t)$, которая определяется каноническими уравнениями движения

$$\dot{q}^* = \frac{\partial H^*}{\partial p^*}, \quad \dot{p}^* = -\frac{\partial H^*}{\partial q^*},$$

выполнено равенство $\dot{H}^* = 0$. Это значит, что на каждой траектории гамильтониан имеет постоянное численное значение, которое называется энергией точечной частицы.

Переформулируем модель точечной частицы, как калибровочную. С этой целью расширим фазовое пространство $\mathbb{T}^*(\mathbb{R}^N) \rightarrow \mathbb{T}^*(\mathbb{R}^{N+1})$ путем введения дополнительной пары сопряженных канонических переменных Q, P , и рассмотрим новое действие

$$S_T = \int_{\tau_1}^{\tau_2} d\tau (\dot{Q}P + \dot{q}^* p_* - \lambda G), \quad G := -Q + H^*, \quad (22.61)$$

где λ – множитель Лагранжа, а точка обозначает дифференцирование по некоторому параметру τ , играющему роль времени. Уравнения движения для этой модели имеют вид

$$\dot{Q} = 0, \quad (22.62)$$

$$\dot{P} = \lambda, \quad (22.63)$$

$$\dot{q}^* = \lambda \frac{\partial H^*}{\partial p_*}, \quad (22.64)$$

$$\dot{p}_* = -\lambda \frac{\partial H^*}{\partial q^*}, \quad (22.65)$$

$$G = 0. \quad (22.66)$$

Последнее уравнение представляет собой уравнение связи. Эта связь, очевидно, является связью первого рода, которая определяет гамильтониан системы.

Согласно общей теории действие (22.61) инвариантно относительно инфинитезимальных калибровочных преобразований с параметром $\epsilon(\tau)$, генерируемых связью G ,

$$\begin{aligned}\delta Q &= \epsilon[Q, G] = 0, \\ \delta P &= \epsilon[P, G] = \epsilon, \\ \delta q^* &= \epsilon[q^*, G] = \epsilon \frac{\partial H^*}{\partial p_*}, \\ \delta p_* &= \epsilon[p_*, G] = -\epsilon \frac{\partial H^*}{\partial q^*}.\end{aligned}\tag{22.67}$$

При этом множитель Лагранжа преобразуется по правилу (22.27)

$$\delta \lambda = \dot{\epsilon}.\tag{22.68}$$

Для того, чтобы вариации δq^* на границе были равны нулю, необходимо предположить, что $\epsilon(\tau_{1,2}) = 0$. Вариация множителя Лагранжа на границе несущественна, т.к. он входит в действие без производных. Отметим, что эволюцию во времени канонических переменных можно рассматривать, как калибровочное преобразование с параметром $\epsilon = \lambda d\tau$.

Приведенные выше калибровочные преобразования представляют собой бесконечно малые преобразования, соответствующие инвариантности действия (22.61) относительно перепараметризации времени. При произвольном преобразовании временного параметра $\tau' = \tau'(\tau)$ мы постулируем, что координаты фазового пространства преобразуются, как скаляры:

$$Q'(\tau') = Q(\tau), \quad P'(\tau') = P(\tau)$$

(такие же формулы преобразования постулируются для q^* и p_*). При этом множитель Лагранжа преобразуется, как 1-форма:

$$\lambda'(\tau') = \frac{d\tau}{d\tau'} \lambda(\tau).$$

Нетрудно проверить, что действие (22.61) инвариантно относительно произвольной перепараметризации времени τ . Это и есть калибровочная инвариантность. При бесконечно малом преобразовании $\tau' = \tau + u(\tau)$ для вариаций формы функций имеем следующие формулы:

$$\begin{aligned}\delta Q &= -u\dot{Q} = 0, \\ \delta P &= -u\dot{P} = -u\lambda, \\ \delta q^* &= -u\dot{q}^* = -u\lambda \frac{\partial H^*}{\partial p_*}, \\ \delta p_* &= -u\dot{p}_* = u\lambda \frac{\partial H^*}{\partial q^*}, \\ \delta \lambda &= -u\dot{\lambda} - \dot{u}\lambda = -\frac{d}{d\tau}(u\lambda),\end{aligned}\tag{22.69}$$

где в первых четырех уравнениях были использованы уравнения движения. Полученные преобразования совпадают с инфинитезимальными калибровочными преобразованиями (22.67), (22.68) при $\epsilon = -u\lambda$.

Таким образом, мы показали, что действие (22.61) инвариантно относительно произвольной перепараметризации времени. В модели имеется одна связь первого рода, и для фиксирования соответствующего произвола в решениях уравнений движения необходимо наложить одно калибровочное условие. Поскольку множитель Лагранжа λ является произвольной функцией времени, то из уравнения движения (22.63) следует, что импульс P также произволен. Чтобы устранить этот произвол, зафиксируем калибровку

$$F_1 := P - \text{const} = 0.$$

Тогда из уравнений движения определяется множитель Лагранжа $\lambda = 0$. При этом для действия получаем следующее выражение

$$S_{\text{т}}|_{F_1=0, G=0} = \int d\tau \dot{q}^* p^*.$$

Это значит, что в выбранной калибровке мы получили “замороженную” теорию, в которой не происходит никакой эволюции. При этом вся эволюция заменяется калибровочным преобразованием.

Можно рассмотреть класс калибровок, явно зависящих от времени. Пусть калибровочное условие имеет вид

$$F_2 := P - \tau = 0.$$

В этом случае из уравнений движения следует $\lambda = 1$, и эффективный гамильтониан для физических степеней свободы становится нетривиальным. Поскольку

$$\dot{Q}P = \frac{d}{d\tau}(QP) - Q\dot{P},$$

и

$$Q\dot{P}|_{F_2=0, G=0} = H^*(q^*, p^*),$$

то эффективное действие для физических степеней свободы равно

$$S_{\text{т}}|_{F_2=0, G=0} = \int d\tau (\dot{q}^* p^* - H^*),$$

что совпадает с исходным действием (22.60) для точечной частицы. Мы видим, что в калибровке $F_2 = 0$ нетривиальный эффективный гамильтониан возникает из кинетического слагаемого $\dot{Q}P$ для нефизической степени свободы.

Можно рассмотреть более общий класс калибровок

$$F_3 := P - f(\tau) = 0,$$

где $f(\tau)$ – некоторая функция времени с положительной производной, $\dot{f} > 0$. Для этой калибровки $\lambda = \dot{f}$, и полное действие принимает вид

$$S_{\text{т}}|_{F_3=0, G=0} = \int d\tau (\dot{q}^* p^* - \dot{f} H^*).$$

После перепараметризации траектории $dt := d\tau \dot{f}$, мы возвращаемся к исходному действию для точечной частицы (22.60).

Таким образом, действие (22.61) калибровочно инвариантно и после наложения калибровочного условия (из достаточно широкого класса калибровок) эквивалентно обычному действию для точечной частицы. В рассмотренном примере нефизическую

степень свободы удалось в явном виде исключить из теории после решения связи и калибровочного условия. В подавляющем большинстве моделей математической физики это сделать не удастся. Даже в электродинамике связи и калибровочные условия нельзя решить в явном виде. Ситуация в моделях Янга–Миллса и гравитации еще более сложная. Поэтому для проведения вычислений в калибровочных моделях используют методы, учитывающие как физические, так и нефизические степени свободы.

А теперь обратимся к вопросу об определении энергии в калибровочно инвариантных теориях. Во многих моделях математической физики исходное действие в гамильтоновой форме имеет вид (22.61). В таком виде гамильтониан системы при выполнении уравнений движения тождественно равен нулю, и принимать его за энергию системы не имеет никакого смысла. В рассмотренной модели за энергию частицы естественно принять численное значение гамильтониана H^* для физических степеней свободы. Для того, чтобы его построить, исходя из действия (22.61), необходимо сначала зафиксировать калибровку, зависящую явно от времени, а затем решить уравнения движения для нефизических степеней свободы и связь. При этом, выбирая различные функции времени в калибровочном условии $F_3 = 0$, можно получить, что множитель Лагранжа и, следовательно, эффективный гамильтониан будут явно зависеть от времени. Для простоты картины следует выбрать такую функцию времени, чтобы эта зависимость исчезла. Если это возможно, то построенный таким образом гамильтониан следует принять за определение энергии, а соответствующий ему временной параметр назвать временем $\tau = t$.

22.8 Частица в псевдоримановом пространстве

Рассмотрим точечную частицу постоянной массы $m > 0$, которая движется в произвольном псевдоримановом пространстве $(\mathbb{M}, g)$ размерности n , на котором задана достаточно гладкая метрика лоренцевой сигнатуры. Пусть x^α , $\alpha = 0, 1, \dots, n-1$, – локальная система координат в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$. Будем считать, что координаты выбраны таким образом, что x^0 является временной координатой, т.е. $g_{00} > 0$, и все сечения $x^0 = \text{const}$ пространственноподобны, т.е. пространственная часть метрики $g_{\mu\nu}$, $\mu, \nu = 1, \dots, n-1$ отрицательно определена или $N^2 > 0$, где N – функция хода в АДМ параметризации метрики (см. раздел 24.2).

Поскольку на многообразии задана метрика лоренцевой сигнатуры, то в каждой точке заданы световые конусы прошлого и будущего. Будем считать, что на $\mathbb{M}$ выбрана ориентация во времени, т.е. световые конусы будущего непрерывно зависят от точки многообразия.

Рассмотрим времениподобную кривую $\gamma = q(t)$, $t \in [t_0, t_1] \subset \mathbb{R}$, соединяющую две причинно связанные точки выбранной координатной окрестности $q(t_0), q(t_1) \in \mathbb{U}$ и целиком лежащую в $\mathbb{U}$. Действие для точечной частицы массы m , по определению, пропорционально длине траектории (16.20) и имеет вид

$$S = \int_{\gamma} dt L(q, \dot{q}) := -m \int_{\gamma} dt \sqrt{g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta}. \quad (22.70)$$

Поскольку траектория частицы предполагается времениподобной, т.е.

$$\dot{q}^2 := g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta > 0,$$

то подынтегральное выражение определено. В рассматриваемом действии метрика $g_{\alpha\beta}(q(t))$ является внешним заданным полем и по ней варьирование не производится пока не включено взаимодействие с гравитационным полем, т.е. не добавлено, например, действие Гильберта–Эйнштейна.

Обозначения выбраны таким образом, чтобы производная, например, от метрики вдоль траектории частицы записывалась в виде

$$\dot{g}_{\alpha\beta} := \frac{dg_{\alpha\beta}}{dt} = \dot{q}^\gamma \frac{\partial g_{\alpha\beta}}{\partial x^\gamma} \Big|_{x=q},$$

что следует из правила дифференцирования сложной функции.

Интеграл в выражении (22.70) имеет размерность длины. Следовательно, для безразмерности действия размерность массы равна $[m] = l^{-1}$. Такую же размерность имеют массы скалярного (раздел 25.1.1) и спинорного (раздел 27.1.5) полей.

Рассмотрим случай как положительных функций хода $N > 0$, так и отрицательных $N < 0$. Если в некоторой области функция хода меняет знак, то из непрерывности следует, что она где-то обращается в нуль. В таких точках метрика становится вырожденной, и этот вопрос требует отдельного рассмотрения. Пока же предположим, что в области $\mathbb{U}$ функция хода либо положительна, либо отрицательна.

Мы рассматриваем оба возможных случая знака производных $\dot{q}^0 > 0$ и $\dot{q}^0 < 0$. В дальнейшем мы увидим, что в случае $\dot{q}^0 < 0$ действие (22.70) описывает античастицу, т.е. частицу той же массы m , но противоположного электрического заряда.

Следуя общим правилам, построим гамильтонов формализм для точечной массивной частицы, описываемой действием (22.70). Обобщенные импульсы, сопряженные координатам точечной частицы q^α имеют вид

$$p_\alpha := \frac{\partial L}{\partial \dot{q}^\alpha} = -m \frac{g_{\alpha\beta} \dot{q}^\beta}{\sqrt{\dot{q}^2}}. \quad (22.71)$$

Отсюда следует, что импульсы удовлетворяют квадратичному соотношению

$$p^2 := g^{\alpha\beta}(q) p_\alpha p_\beta = m^2, \quad (22.72)$$

которое должно быть выполнено для всех траекторий, вдоль которых частица может двигаться. Поэтому соотношение

$$\tilde{G} := p^2 - m^2 = 0 \quad (22.73)$$

является первичной связью для точечной частицы.

Пример 22.8.1. В пространстве Минковского связь (22.73) зависит только от импульсов:

$$p_0^2 + \eta^{\mu\nu} p_\mu p_\nu - m^2 = 0,$$

где $\eta^{\mu\nu} = \text{diag}(-\dots-)$, и выделяет в фазовом пространстве $(q, p) \in \mathbb{T}^*(\mathbb{R}^{1,n-1})$ двуполостный гиперboloид в кокасательном пространстве, умноженный на пространство Минковского $\mathbb{R}^{1,n-1}$, которое соответствует координатам q . Поверхность связей является несвязным подмногообразием в фазовом пространстве $\mathbb{T}^*(\mathbb{R}^{1,n-1})$ и состоит из двух компонент связности, определяемых неравенствами $p_0 > 0$ и $p_0 < 0$. Топологически поверхность связей в произвольном псевдоримановом многообразии устроена также. $\square$

Используя АДМ параметризацию метрики (см. раздел 24.2), связь (22.73) перепишем в виде произведения двух сомножителей:

$$p^2 - m^2 = \left[\frac{1}{N}(p_0 - N^\mu p_\mu) + \sqrt{\hat{p}^2 + m^2} \right] \left[\frac{1}{N}(p_0 - N^\mu p_\mu) - \sqrt{\hat{p}^2 + m^2} \right] = 0, \quad (22.74)$$

где введено обозначение $\hat{p}^2 := -\hat{g}^{\mu\nu} p_\mu p_\nu > 0$ для положительно определенного квадрата пространственных компонент импульса частицы. Напомним, что индексы из середины греческого алфавита пробегают только пространственные значения: $\mu, \nu, \dots = 1, \dots, n-1$ и $\hat{g}^{\mu\nu}$ – отрицательно определенная матрица, обратная к $g_{\mu\nu}$. В дальнейшем для подъема пространственных индексов всегда используется обратная метрика $\hat{g}^{\mu\nu}$.

Равенство нулю одного из сомножителей выделяет в фазовом пространстве одну полу “гиперболоида”. То, на какой именно полё “гиперболоида” находится частица, определяется начальными данными. Если в начальный момент времени частица находилась, скажем, на полё, определяемой первым сомножителем в (22.74), то из непрерывности следует, что она на ней и останется в процессе эволюции.

Из определения импульсов (22.71) следует равенство

$$p_0 - N^\mu p_\mu = -\frac{mN^2}{\sqrt{\dot{q}^2}} \dot{q}^0. \quad (22.75)$$

Мы видим, что знак производной $\dot{q}^0$ всегда противоположен знаку функции $p_0 - N^\mu p_\mu$. Поэтому, если $\dot{q}^0 N > 0$, то в нуль обращается первый сомножитель в формуле (22.74). В противном случае, $\dot{q}^0 N < 0$, равен нулю второй сомножитель. Следовательно, связь (22.73) эквивалентна связи

$$G := \sqrt{\hat{p}^2 + m^2} - \left| \frac{p_0 - N^\mu p_\mu}{N} \right| = 0, \quad (22.76)$$

где мы использовали знак модуля, чтобы объединить оба случая. В таком виде первичная связь будет удобна для дальнейших вычислений.

Мы предполагаем, что если в начальный момент времени частица находилась на какой-то одной полё гиперboloида, то в процессе движения она на ней и останется. В противном случае траектория в фазовом пространстве не будет непрерывной. Это означает, что если в некоторой области пространства-времени функция хода меняет знак, то с точки зрения внешнего наблюдателя частица будет восприниматься как частица в области $N > 0$ и античастица в области $N < 0$, т.к. $\dot{q}^0$ меняет знак.

Перейдем к вычислению гамильтониана $H := p_\alpha \dot{q}^\alpha - L$. Поскольку в теории есть первичная связь, то гессиан модели вырожден и из определения обобщенных импульсов (22.71) нельзя определить все скорости, как функции импульсов и координат. Ранг гессиана в рассматриваемом случае равен $n-1$, что позволяет определить пространственные компоненты скорости. Чтобы их найти, заметим, что

$$\dot{q}^2 = N^2(\dot{q}^0)^2 + g_{\mu\nu}(\dot{q}^\mu + N^\mu \dot{q}^0)(\dot{q}^\nu + N^\nu \dot{q}^0). \quad (22.77)$$

Далее, из определения импульсов (22.71) следует равенство

$$p_\mu \sqrt{\dot{q}^2} = -m(N_\mu \dot{q}^0 + g_{\mu\nu} \dot{q}^\nu).$$

Возведение этого равенства в квадрат с помощью метрики $\hat{g}^{\mu\nu}$ позволяет найти квадрат пространственных компонент скорости:

$$g_{\mu\nu}(\dot{q}^\mu + N^\mu \dot{q}^0)(\dot{q}^\nu + N^\nu \dot{q}^0) = -\frac{N^2 \hat{p}^2}{\hat{p}^2 + m^2} (\dot{q}^0)^2.$$

Подстановка полученного выражения в формулу (22.77) приводит к равенству

$$\sqrt{\dot{q}^2} = \frac{m|\dot{q}^0 N|}{\sqrt{\hat{p}^2 + m^2}},$$

которое позволяет решить выражение для импульсов (22.71) относительно компонент скорости:

$$\frac{\dot{q}^0}{|\dot{q}^0|} = -\frac{p_0 - N^\nu p_\nu}{|N|\sqrt{\hat{p}^2 + m^2}}, \quad (22.78)$$

$$\frac{\dot{q}^\mu}{|\dot{q}^0|} = -\frac{N^2 \hat{g}^{\mu\nu} p_\nu - N^\mu (p_0 - N^\nu p_\nu)}{|N|\sqrt{\hat{p}^2 + m^2}}, \quad (22.79)$$

Мы видим, что уравнения (22.71) определяют только знак производной $\dot{q}^0$ и пространственные компоненты скорости $\dot{q}^\mu$. При этом модуль временной компоненты скорости $|\dot{q}^0|$ является произвольной функцией. Позже мы увидим, что она соответствует перепараметризации мировой линии частицы.

Поскольку мы нашли все $n - 1$ пространственные компоненты скоростей, то это доказывает, что гессиан модели имеет ранг $n - 1$, и других первичных связей в теории нет.

Теперь нетрудно вычислить гамильтониан

$$\begin{aligned} H &:= p\dot{q} - L = p_0\dot{q}^0 + p_\mu\dot{q}^\mu + m\sqrt{\dot{q}^2} = \\ &= \frac{|\dot{q}^0|}{|N|\sqrt{\hat{p}^2 + m^2}} \left[p_0|N|\sqrt{\hat{p}^2 + m^2} + N^\mu p_\mu (p_0 - N^\nu p_\nu) + N^2(\hat{p}^2 + m^2) \right] = \\ &= \frac{|\dot{q}^0|}{|N|\sqrt{\hat{p}^2 + m^2}} \left[-(p_0 - N^\mu p_\mu)^2 + N^2(\hat{p}^2 + m^2) \right] = -\frac{|\dot{q}^0 N|}{\sqrt{\hat{p}^2 + m^2}} \tilde{G}, \end{aligned} \quad (22.80)$$

где мы использовали связь (22.76) во второй строке. Таким образом, гамильтониан определен и пропорционален первичной связи $\tilde{G}$. Он определен неоднозначно, т.к. в теории есть связь, которую мы используем.

Перепишем гамильтониан в более удобной форме. С этой целью в предпоследнем выражении (22.80) используем связь (22.76):

$$H = \frac{|\dot{q}^0|}{|N|\sqrt{\hat{p}^2 + m^2}} \left[-|(p_0 - N^\mu p_\mu)N|\sqrt{\hat{p}^2 + m^2} + N^2(\hat{p}^2 + m^2) \right] = |\dot{q}^0 N|G. \quad (22.81)$$

Теперь гамильтониан пропорционален связи в форме (22.76).

Действие (22.70) в гамильтоновой форме принимает вид

$$S = \int dt (p_\alpha \dot{q}^\alpha - H) = \int dt \left(p_\mu \dot{q}^\mu - |\dot{q}^0 N|\sqrt{\hat{p}^2 + m^2} + \dot{q}^0 N^\mu p_\mu \right). \quad (22.82)$$

Отметим сокращение кинетических слагаемых $p_0\dot{q}^0$. Как и исходное действие полученное выражение не зависит от знака функции хода и параметризационно инвариантно, если перепараметризация не меняет ориентацию кривой.

Действие (22.82) приводит к следующим уравнениям движения для пространственных координат и импульсов:

$$\dot{q}^\mu = - \frac{|\dot{q}_0 N|}{\sqrt{\hat{p}^2 + m^2}} \Big|_{x=q} p^\mu - \dot{q}^0 N^\mu \Big|_{x=q}, \quad (22.83)$$

$$\dot{p}_\mu = -\partial_\mu \left[|\dot{q}^0 N| \sqrt{\hat{p}^2 + m^2} - \dot{q}^0 N^\nu p_\nu \right]_{x=q} = \quad (22.84)$$

$$= - |\dot{q}^0| \sqrt{\hat{p}^2 + m^2} \partial_\mu |N| \Big|_{x=q} - \frac{|\dot{q}^0 N| p^\nu p^\rho \hat{\Gamma}_{\mu\nu\rho}}{\sqrt{\hat{p}^2 + m^2}} \Big|_{x=q} + \dot{q}^0 \partial_\mu N^\nu p_\nu \Big|_{x=q}, \quad (22.85)$$

где $\hat{\Gamma}_{\mu\nu\rho}$ – символы Кристоффеля для пространственной метрики $g_{\mu\nu}$ и $p^\mu := \hat{g}^{\mu\nu} p_\nu$. В таком виде эквивалентность гамильтоновых и лагранжевых уравнений движения совсем не очевидна. В дальнейшем эквивалентность гамильтоновых и лагранжевых уравнений движения будет доказана для уравнений, записанных в другой форме.

Поскольку исходное действие инвариантно относительно произвольной перепараметризации мировой линии частицы (локальные преобразования), то рассматриваемая модель является калибровочной. Этому обстоятельству соответствует наличие одной связи первого рода (22.73). Поэтому частица имеет физические и нефизическую степень свободы. В качестве физических степеней свободы, для которых можно поставить задачу Коши, выберем пространственные компоненты координат и импульсов q^μ, p_μ , а нефизической степени свободы – q^0, p_0 . Посмотрим с этой точки зрения на действие (22.82). Для определенности предположим, что $\dot{q}_0 N > 0$. Тогда действие можно переписать в виде

$$S = \int dt (p_\mu \dot{q}^\mu - \dot{q}^0 H_{\text{eff}}), \quad (22.86)$$

где

$$H_{\text{eff}} := N \sqrt{\hat{p}^2 + m^2} - N^\mu p_\mu$$

– эффективный гамильтониан для физических степеней свободы. При этом уравнения движения примут вид

$$\begin{aligned} \dot{q}^\mu &= \dot{q}_0 \frac{\partial H_{\text{eff}}}{\partial p_\mu}, \\ \dot{p}_\mu &= -\dot{q}^0 \frac{\partial H_{\text{eff}}}{\partial q^\mu}. \end{aligned} \quad (22.87)$$

Легко проверить, что если выполнены уравнения движения и метрика $g_{\alpha\beta}$ не зависит от времени $x^0 = q^0$, то энергия сохранится во времени,

$$E := H_{\text{eff}} = \text{const},$$

для произвольной функции $q^0(t)$. Мы видим, что произвольная функция $q^0(t)$ не определяется уравнениями движения и соответствует свободе в выборе параметра вдоль мировой линии. От нее всегда можно избавиться, переопределив параметр, что соответствует выбору калибровки. Таким образом, нефизическая степень свободы убирается из модели путем решения связи (22.76) относительно временной компоненты импульса p_0 и наложения калибровочного условия на произвольную функцию $q^0(t)$.

Заметим, что вариационная производная действия (22.82) по q_0 имеет вид

$$\frac{\delta S}{\delta q^0} = \frac{dH_{\text{eff}}}{dt} = \dot{q}^\mu \frac{\partial H_{\text{eff}}}{\partial q^\mu} + \dot{p}_\mu \frac{\partial H_{\text{eff}}}{\partial p_\mu}.$$

Эта вариационная производная равна нулю на уравнениях движения (22.87). Следовательно, вариация действия (22.86) по q^0 не дает никаких новых уравнений движения.

Если предположить, что допускаются только те перепараметризации, которые сохраняют ориентацию мировой линии, т.е. $\dot{q}^0(t) > 0$, то отрицательной функции хода будет соответствовать эффективный гамильтониан

$$H_{\text{eff}} := -N\sqrt{\hat{p}^2 + m^2} - N^\mu p_\mu.$$

В любом случае эффективный гамильтониан для физических степеней свободы массивной точечной частицы положительно определен при $N^\mu = 0$.

В соответствии с общим методом, описанном в разделе 22.3, полный гамильтониан на первом этапе получается путем добавления первичной связи. Так как исходный гамильтониан уже пропорционален связи, то полный гамильтониан равен

$$H_T = \lambda G, \quad (22.88)$$

где $\lambda = \lambda(t)$ – неопределенный множитель Лагранжа.

Поскольку скобка Пуассона связи с собой равна нулю,

$$[G, G] = 0$$

(она должна быть антисимметрична, что невозможно для одной связи), то вторичных связей не возникает и связь $G = 0$ является связью первого рода. Поэтому система находится в инволюции, т.е. выполнены условия (22.20), (22.21). Отсюда вытекает, что полный гамильтониан системы (22.24) определяется единственной связью первого рода (22.88).

Гамильтониан (22.88) приводит к уравнениям движения, которые эквивалентны уравнениям движения для гамильтониана

$$\tilde{H}_T = \mu \tilde{G}, \quad (22.89)$$

где $\mu = \mu(t)$ – множитель Лагранжа, на поверхности связей, т.к. связи $\tilde{G}$ и G эквивалентны. Действительно, связи G и $\tilde{G}$ отличаются на отличный от нуля множитель (для каждой полы гиперboloида):

$$\tilde{G} = fG, \quad f(q, p) \neq 0.$$

Поэтому

$$\dot{q}^\alpha = \mu[q^\alpha, \tilde{G}] = \mu[q^\alpha, f]G + \mu f[q^\alpha, G].$$

На поверхности связи $G = 0$, и поэтому первое слагаемое в правой части исчезает. Такой же вид имеет уравнение для импульсов. Следовательно, замена связи в полном гамильтониане H_T приводит к переопределению множителя Лагранжа: $\lambda = \mu f$.

В дальнейшем, из соображений удобства, мы будем выбирать тот или иной вид полного гамильтониана.

Рассмотрим гамильтоновы уравнения движения (22.25) для гамильтониана (22.89)

$$\begin{aligned}\dot{q}^\alpha &= 2\mu g^{\alpha\beta} p_\beta|_{x=q}, \\ \dot{p}_\alpha &= \mu p^\beta p^\gamma \partial_\alpha g_{\beta\gamma}|_{x=q}.\end{aligned}\quad (22.90)$$

Для сравнения полученных уравнений движения с уравнениями для экстремалей (16.21), продифференцируем первое уравнение по t и воспользуемся вторым уравнением для исключения $\dot{p}$. В результате получим уравнение для координат частицы

$$\ddot{q}^\alpha = \frac{\dot{\mu}}{\mu} \dot{q}^\alpha - \Gamma_{\beta\gamma}{}^\alpha \dot{q}^\beta \dot{q}^\gamma. \quad (22.91)$$

С точностью до первого слагаемого в правой части оно совпадает с уравнением для экстремалей (16.21). Дополнительное слагаемое связано с произволом в выборе параметризации мировой линии частицы. Действительно, после преобразования $t \mapsto t'(t)$, где функция $t'(t)$ удовлетворяет дифференциальному уравнению

$$\frac{dt'}{dt} = \mu(t), \quad \mu > 0,$$

гамильтоновы уравнения движения будут иметь вид (22.91), но с $\mu = 1$. В этом случае уравнения движения совпадут с уравнениями для экстремалей, и, значит, параметр t' является каноническим параметром вдоль экстремали.

Обратное утверждение также верно. Если выполнены уравнения для экстремалей

$$\ddot{q}^\alpha = -\Gamma_{\beta\gamma}{}^\alpha \dot{q}^\beta \dot{q}^\gamma,$$

то из них вытекают гамильтоновы уравнения (22.90) при $\mu = 1$. Таким образом мы доказали эквивалентность гамильтоновых и лагранжевых уравнений движения.

Поскольку модель содержит связь первого рода, то ей соответствует калибровочная инвариантность полного действия

$$S_T = \int_\gamma dt (p_\alpha \dot{q}^\alpha - \mu \tilde{G}). \quad (22.92)$$

Чтобы найти соответствующие преобразования симметрии, рассмотрим инфинитезимальные преобразования, которые генерируются связью (22.27):

$$\begin{aligned}\delta q^\alpha &= \epsilon [q^\alpha, \tilde{G}] = 2\epsilon g^{\alpha\beta} p_\beta, \\ \delta p_\alpha &= \epsilon [p_\alpha, \tilde{G}] = \epsilon p^\beta p^\gamma \partial_\alpha g_{\beta\gamma}, \\ \delta \mu &= \dot{\epsilon}.\end{aligned}$$

Сравнение этих преобразований с уравнениями движения (22.90) показывает, что эволюция во времени канонических переменных представляет собой последовательность калибровочных преобразований, где $\epsilon = \mu dt$. Эта калибровочная симметрия описывает произвол в выборе параметра t вдоль траектории частицы. Уравнения движения (22.90) вместе со связью (22.76) не определяют множитель Лагранжа μ . Для его определения необходимо зафиксировать калибровку.

Если учесть выражение для импульсов (22.71), то кинетический член в действии (22.92) примет вид

$$p_\alpha \dot{q}^\alpha = -m \sqrt{\dot{q}^2}.$$

После интегрирования кинетического члена получается исходное действие (22.70). Таким образом, на поверхности связей гамильтониан равен нулю, а исходное действие определяется только кинетическим слагаемым.

22.8.1 Временная калибровка

Продолжим исследование модели в соответствии с общей схемой. Зафиксируем *временную калибровку*

$$F_T := q^0 - bt = 0, \quad b = \text{const} \neq 0. \quad (22.93)$$

Модуль свободного параметра b связан с выбором единицы измерения времени, что не существенно. Поэтому, без ограничения общности, положим $|b| = 1$, т.е. $b = 1$, если время q^0 для внешнего наблюдателя увеличивается при увеличении параметра t вдоль мировой линии частицы, и $b = -1$, если при увеличении t время q^0 уменьшается.

Скобка Пуассона этого калибровочного условия со связью равна

$$[F_T, G] = -\frac{1}{|N|} [q^0, |p_0 - N^\mu p_\mu|] = \begin{cases} 1/|N|, & b > 0, \\ -1/|N|, & b < 0, \end{cases} \quad (22.94)$$

где мы использовали равенство (22.75) для определения знака выражения, стоящего под модулем. Поскольку функция хода отлична от нуля, $N \neq 0$, то скобка Пуассона (22.94) отлична от нуля во всем фазовом пространстве и, в частности, на поверхности связи $G = 0$. Следовательно, условие $q^0 = bt$ определяет каноническую калибровку.

Расширенный гамильтониан с учетом связи и калибровочного условия имеет вид

$$H_E = \lambda G + \pi F_T,$$

где λ и π – множители Лагранжа. Из условий сохранения связи и калибровочного условия во времени,

$$\begin{aligned} \dot{G} &= \lambda[G, G] + \pi[G, F_T] = \pi[G, F_T] \approx 0, \\ \dot{F}_T &= \frac{\partial F_T}{\partial t} + \lambda[F_T, G] + \pi[F_T, F_T] = -b + \lambda[F_T, G] \approx 0, \end{aligned}$$

находим множители Лагранжа:

$$\lambda = |N|, \quad \pi = 0,$$

где мы объединили оба случая, $b > 0$ и $b < 0$, и учли равенство $|b| = 1$. Таким образом, в данной канонической калибровке полный гамильтониан (22.88) совпадает с исходным гамильтонианом (22.81), т.к. $|\dot{q}^0| = |b| = 1$.

Связь (22.76) и калибровочное условие (22.93) можно решить относительно нефизических переменных q^0 и p_0 :

$$\begin{aligned} q^0 &= bt, \\ p_0 &= -b|N|\sqrt{\hat{p}^2 + m^2} + N^\mu p_\mu. \end{aligned}$$

Так как решения уравнений Эйлера–Лагранжа можно подставлять в действие (см. раздел 20.4), то эффективное действие для физических переменных q^μ, p_μ принимает вид

$$S_{\text{eff}} = \int dt (p_0 \dot{q}^0 + p_\mu \dot{q}^\mu) = \int dt (p_\mu \dot{q}^\mu - H_{\text{eff}}),$$

где эффективный гамильтониан для физических степеней свободы имеет вид

$$H_{\text{eff}} = |N|\sqrt{\hat{p}^2 + m^2} - bN^\mu p_\mu. \quad (22.95)$$

Как и в случае нерелятивистской точечной частицы эффективный гамильтониан для физических степеней свободы полностью определяется кинетическим слагаемым $p_0 \dot{q}^0$ для нефизических степеней свободы. Эффективный гамильтониан зависит только от физических степеней свободы, которыми являются пространственные координаты точечной частицы и соответствующие импульсы, $\{q^\mu, p_\mu\}$, $\mu = 1, \dots, n-1$. Компоненты метрики N , N^μ и $g_{\mu\nu}$ входят в гамильтониан в качестве внешних полей. Уравнения движения для физических степеней свободы во временной калибровке (22.93) имеют вид

$$\dot{q}^\mu = -\frac{|N|p^\mu}{\sqrt{\hat{p}^2 + m^2}} - bN^\mu, \quad (22.96)$$

$$\dot{p}_\mu = -\sqrt{\hat{p}^2 + m^2}\partial_\mu|N| + \frac{|N|\partial_\mu\hat{g}^{\nu\rho}p_\nu p_\rho}{2\sqrt{\hat{p}^2 + m^2}} + b\partial_\mu N^\nu p_\nu, \quad (22.97)$$

где $p^\mu := \hat{g}^{\mu\nu}p_\nu$. Конечно, первое уравнение (22.96) совпадает с равенством (22.79) при $|\dot{q}^0| = 1$ и учете связи (22.76).

Чтобы дать физическую интерпретацию двух возможных ориентаций мировой линии $b = \pm 1$ рассмотрим следующий

Пример 22.8.2. В пространстве Минковского $\mathbb{R}^{1,n-1}$ компоненты метрики имеют вид

$$N = 1, \quad N^\mu = 0, \quad g_{\mu\nu} = \eta_{\mu\nu} = \text{diag}(-\dots-),$$

и уравнения движения (22.96), (22.97) существенно упрощаются:

$$\dot{q}^\mu = \frac{p^\mu}{\sqrt{\hat{p}^2 + m^2}}, \quad \dot{p}_\mu = 0.$$

Эти уравнения имеют хорошо известный в специальной теории относительности вид. Возводя первое уравнение в квадрат с помощью пространственной метрики $\eta_{\mu\nu}$, получим выражение для пространственного импульса частицы через ее скорость:

$$p^\mu = \frac{m\dot{q}^\mu}{\sqrt{1 - \mathbf{u}^2}}, \quad (22.98)$$

где мы ввели квадрат пространственной скорости частицы $\mathbf{u}^2 := -\dot{q}^\mu \dot{q}^\nu \eta_{\mu\nu} > 0$. Тогда уравнения движения для свободной точечной частицы в пространстве Минковского во временной калибровке сводятся просто к условию сохранения импульса: $p_\mu = \text{const}$.

Для того, чтобы дать физическую интерпретацию двум возможным ориентациям $\dot{q}^0 = \pm 1$ мировой линии частицы относительно временной координаты q^0 внешнего наблюдателя, рассмотрим взаимодействие частицы с внешним электромагнитным полем в четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$. Такое взаимодействие описывается дополнительным слагаемым в лагранжиане:

$$L \mapsto L + eA_\alpha \dot{q}^\alpha,$$

где $A_\alpha(x)$ – компоненты потенциала электромагнитного поля и $e = \text{const}$ – заряд частицы. С точки зрения внешнего наблюдателя, для которого временем является координата q^0 , добавочное слагаемое в лагранжиане имеет вид

$$gA_\alpha \frac{dq^\alpha}{dq^0},$$

где $g := be = \pm e$ – наблюдаемый заряд частицы. Таким образом, для внешнего наблюдателя частица имеет либо положительный, либо отрицательный заряд. В остальном частицы совпадают. Такие частицы в физике принято называть частицей и античастицей. Например, электрон и позитрон (они, правда, имеют спин $1/2$, в то время как в рассматриваемом случае спин частиц равен нулю).

Таким образом, свобода в выборе ориентации мировой линии частицы относительно временной координаты внешнего наблюдателя соответствует двум возможным зарядам частицы $\pm e$. Это означает, что исходное действие (22.70) описывает частицу и античастицу, у которых массы совпадают, а заряды при включении внешнего электромагнитного поля имеют противоположный знак. Если заряд равен нулю, $e = 0$, то действие (22.70) описывает одну нейтральную частицу.

Заметим, что и частица, и античастица движутся вперед по времени x^0 с точки зрения внешнего наблюдателя. Однако собственное время античастицы t движется в обратную сторону.

Гамильтоновы уравнения движения для точечной заряженной частицы принимают особенно простой вид при движении в постоянном магнитном поле. В этом случае временная компонента потенциала равна нулю, а пространственные компоненты зависят только от пространственных координат:

$$\{A_\alpha\} = \{A_0 = 0, A_\mu(\mathbf{x})\}.$$

Тогда все изменения в гамильтоновом формализме сводятся к переопределению пространственных компонент импульсов:

$$p_\mu \mapsto \tilde{p}_\mu := \frac{\partial(L + eA_\nu \dot{q}^\nu)}{\partial \dot{q}^\mu} = p_\mu + eA_\mu.$$

При этом связь (22.72) остается прежней, а скобка Пуассона имеет вид $[q^\mu, \tilde{p}_\nu] = \delta_\nu^\mu \delta(\mathbf{x} - \mathbf{x}')$. В частности, гамильтониан для физических степеней свободы и уравнения движения принимают вид

$$\begin{aligned} H_{\text{eff}} &= \sqrt{\hat{p}^2 + m^2}, & \hat{p}^2 &:= -\eta^{\mu\nu}(\tilde{p}_\mu - eA_\mu)(\tilde{p}_\nu - eA_\nu), \\ \dot{q}^\mu &= -\frac{\tilde{p}^\mu - eA^\mu}{\sqrt{\hat{p}^2 + m^2}}, \\ \dot{\tilde{p}}_\mu &= -e \frac{\partial_\mu A_\nu (\tilde{p}^\nu - eA^\nu)}{\sqrt{\hat{p}^2 + m^2}}. \end{aligned}$$

Как видим, при взаимодействии с внешним полем импульс больше не сохраняется, и траектория частицы в общем случае отличается от экстремали. $\square$

Точечные частицы под действием гравитационного поля двигаются в пространстве-времени вдоль экстремалей. В разделе 16.2 было показано, что через данную точку в данном направлении проходит одна и только одна экстремаль. Это значит, что при постановке задачи Коши для экстремали достаточно задать точку $x \in \mathbb{M}$ и вектор $X \in \mathbb{T}_x(\mathbb{M})$. Анализ настоящего раздела показывает, что эта информация является избыточной. Действительно, длина касательного вектора к экстремали постоянна вдоль экстремали. Это значит, что для однозначного восстановления экстремали, проходящей через данную точку, достаточно задать не сам вектор, а его направление, которое определяется $n - 1$ параметром. Кроме того, остается еще произвол в выборе параметризации. Во временной калибровке в качестве параметра выбирается

наблюдаемое время $q^0 = \pm t$. Следовательно, для задания траектории частицы достаточно задать пространственные координаты q^μ и пространственные компоненты импульсов p_μ в начальный момент времени q^0 . Поэтому точечная частица на псевдоримановом многообразии имеет $n - 1$ степень свободы.

Пример 22.8.3. В четырехмерном пространстве-времени Минковского $\mathbb{R}^{1,3}$ точечная массивная частица имеет три степени свободы. $\square$

22.8.2 Калибровка светового конуса

Для многих приложений, например, в суперсимметричных моделях, удобно использовать калибровку светового конуса. Эта калибровка упрощает многие формулы, если частица движется в пространстве Минковского $\mathbb{R}^{1,n-1}$, что мы и предположим.

Чтобы определить калибровку светового конуса, вместо двух первых координат частицы q^0 и q^1 введем новые *конусные переменные*

$$q^\pm := \frac{1}{\sqrt{2}}(q^0 \pm q^1),$$

оставив остальные координаты без изменения. Поскольку известно выражение новых координат через старые, то совершим каноническое преобразование в фазовом пространстве с производящей функцией

$$S_3 = -\frac{1}{\sqrt{2}}(p_0 + p_1)q^+ - \frac{1}{\sqrt{2}}(p_0 - p_1)q^- - p_2q^2 - \dots - p_{n-1}q^{n-1},$$

зависящей от новых координат и старых импульсов (см. раздел 21.12). Отсюда следуют выражения для новых импульсов, сопряженных $q^\pm$:

$$p_\pm = -\frac{\partial S_3}{\partial q^\pm} = \frac{1}{\sqrt{2}}(p_0 \pm p_1).$$

Остальные импульсы при этом не меняются.

В новых канонических переменных первичная связь (22.73) принимает вид

$$\tilde{G} = 2p_+p_- + p^A p_A - m^2 = 0, \quad (22.99)$$

где суммирование ведется только по $n - 2$ значениям индексов: $A = 2, \dots, n - 1$. Поскольку квадратичная форма $p^A p_A - m^2$ отрицательно определена, то $p_+p_- > 0$. Полюсы гиперboloида, соответствующего поверхности связей, определяются неравенствами $(p_+ > 0, p_- > 0)$ и $(p_+ < 0, p_- < 0)$. И в любом случае $p_+p_- \neq 0$. Связь (22.99) просто решается

$$p_+ = \frac{1}{2p_-}(-p^A p_A + m^2), \quad (22.100)$$

или

$$p_- = \frac{1}{2p_+}(-p^A p_A + m^2).$$

Зафиксируем калибровку *светового конуса*

$$F_{LC} := q^+ - at = 0, \quad a = \pm 1. \quad (22.101)$$

Легко видеть, что

$$[\tilde{G}, F_{\text{LC}}] = -2p_-.$$

Поскольку на поверхности связей $p_- \neq 0$, то условие (22.101) определяет каноническую калибровку.

Так как связи $\tilde{G}$ и G пропорциональны, то обобщенный гамильтониан можно записать в виде

$$H_{\text{е}} = \mu \tilde{G} + \pi F_{\text{LC}},$$

где μ и π – множители Лагранжа. Из условий сохранения связи и калибровочного условия во времени,

$$\begin{aligned} \dot{\tilde{G}} &= \mu[\tilde{G}, \tilde{G}] + \pi[\tilde{G}, F_{\text{LC}}] = \pi[\tilde{G}, F_{\text{LC}}] \approx 0, \\ \dot{F}_{\text{LC}} &= \frac{\partial F_{\text{LC}}}{\partial t} + \mu[F_{\text{LC}}, \tilde{G}] + \pi[F_{\text{LC}}, F_{\text{LC}}] = -a + \mu[F_{\text{LC}}, \tilde{G}] \approx 0, \end{aligned}$$

находим множители Лагранжа:

$$\mu = \frac{a}{2p_-}, \quad \pi = 0.$$

В калибровке светового конуса (22.101) физическими переменными являются $\{q^-, q^A, p_-, p_A\}$, $A = 2, \dots, n-1$. После подстановки решения уравнения связи (22.100) и калибровочного условия (22.101) в действие,

$$S_{\text{eff}} = \int dt (p_+ \dot{q}^+ + p_- \dot{q}^- + p_A \dot{q}^A) = \int dt (p_- \dot{q}^- + p_A \dot{q}^A - H_{\text{eff}}),$$

получаем выражение для эффективного гамильтониана

$$H_{\text{eff}} = a \frac{p^A p_A - m^2}{2p_-}. \quad (22.102)$$

Как и раньше, нетривиальный эффективный гамильтониан для физических степеней свободы возникает из кинетического слагаемого $p_+ \dot{q}^+$ для нефизической степени свободы.

Уравнения движения для физических степеней свободы имеют вид

$$\begin{aligned} \dot{q}^- &= -a \frac{p^A p_A - m^2}{2p_-^2}, & \dot{q}^A &= a \frac{p^A}{p_-}, \\ \dot{p}_- &= 0, & \dot{p}_A &= 0. \end{aligned}$$

Мы видим, что, как и во временной калибровке, уравнения движения в калибровке светового конуса сводятся к сохранению обобщенных импульсов свободной точечной частицы.

Замечание. В калибровке светового конуса параметр эволюции $t \in \mathbb{R}$ целиком лежит на световом конусе в пространстве Минковского, который является характеристикой волнового уравнения (см. раздел 16.6). $\square$

Эффективный гамильтониан (22.102) в калибровке светового конуса мало чем отличается от эффективного гамильтониана (22.95) во временной калибровке. Тем не менее оба гамильтониана описывают одну и ту же массивную частицу. Ее траектории в конфигурационном пространстве – это экстремали. Выбор той или иной калибровки является существенным для анализа уравнений движения и квантования. Часто калибровка светового конуса упрощает вычисления, особенно в квантовой теории поля. Во многом это связано с возможностью решения связи (22.99) в классе рациональных функций.

22.9 Граничные слагаемые в калибровочных моделях

В настоящем разделе мы рассмотрим простой пример вариационной задачи, который проанализируем с различных точек зрения. Этот пример позволяет продемонстрировать тонкости вариационной задачи в теории поля, важную роль граничных слагаемых в действии и связь вариационной задачи на условную стационарную точку с фиксированием калибровки в калибровочных моделях.

22.9.1 Лагранжева формулировка

Обозначим декартовы координаты на евклидовой плоскости $\mathbb{R}^2$ через $x^0 = \tau$, $x^1 = \sigma$. Будем называть координату τ временем (параметр эволюции), а σ – пространством, хотя мы не предполагаем наличие на $\mathbb{R}^2$ какой либо метрики. Пусть в конечном прямоугольнике задано действие

$$S = \int_{\tau_1}^{\tau_2} d\tau \int_{\sigma_1}^{\sigma_2} d\sigma (p\dot{q} + p^*\dot{q}^*), \quad (22.103)$$

зависящее от четырех полей $q, p, q^*, p^* \in \mathcal{C}^1(\mathbb{R}^2)$. В действии точка обозначает дифференцирование по τ . В дальнейшем пределы интегрирования, для краткости, будем опускать. Рассмотрим задачу на условный экстремум для действия (22.103). Пусть на поля наложены две связи:

$$G := -\partial_1 q + H^*(q^*, p^*) = 0, \quad (22.104)$$

$$F := p = 0. \quad (22.105)$$

Будем считать, что функция $H^*(q^*, p^*) \geq 0$ достаточно гладко зависит только от полей и не зависит от их производных. Предположим также, что поля q и q^* имеют определенные граничные условия при $\tau = \tau_{1,2}$. Этого достаточно для того, чтобы избежать граничных вкладов в вариацию действия (22.103), возникающих при интегрировании по частям. Тогда для действия (22.103) определена задача на условный экстремум (см. раздел 20.1.4).

Решим эту задачу прямым методом и методом множителей Лагранжа. В первом случае исключим из действия переменные q и p с помощью уравнений связей. Поскольку на поверхности связей $p = 0$, то, независимо от вида функции $q(\tau, \sigma)$, первое слагаемое в (22.103) обращается в нуль, и мы получаем эффективное действие

$$S|_{G=0, F=0} = \int d\tau d\sigma p^* \dot{q}^*, \quad (22.106)$$

в котором переменные q^* и p^* уже рассматриваются, как независимые переменные. Поскольку вариация δq^* равна нулю на границе $\tau = \tau_{1,2}$, то из вариационного принципа следуют только уравнения Эйлера–Лагранжа, которые просто интегрируются:

$$\frac{\delta S}{\delta p^*} = \dot{q}^* = 0, \quad \Rightarrow \quad q^* = q^*(\sigma), \quad (22.107)$$

$$\frac{\delta S}{\delta q^*} = -\dot{p}^* = 0, \quad \Rightarrow \quad p^* = p^*(\sigma). \quad (22.108)$$

Мы видим, что решением уравнений Эйлера–Лагранжа являются произвольные функции от σ . Вид произвольной функции q^* находится из граничных условий, которые

должны быть заданы одинаковыми при $\tau = \tau_{1,2}$. Затем можно определить q из уравнения связи (22.104). Таким образом, задача на условный экстремум имеет решение, хотя и не для очень широкого класса граничных условий.

Поскольку с помощью уравнений связей мы исключили переменные q и p , то будем называть их нефизическими, а переменные q^* и p^* – физическими.

Теперь воспользуемся методом неопределенных множителей Лагранжа. Построим расширенное действие

$$S_E = \int d\tau d\sigma (p\dot{q} + p^*\dot{q}^* - \lambda G - \mu F), \quad (22.109)$$

где $\lambda, \mu \in \mathcal{C}^1(\mathbb{R}^2)$ – множители Лагранжа. Для этого действия нефизическими полями являются q, p, λ и μ . Полная система уравнений Эйлера–Лагранжа имеет вид

$$\frac{\delta S_E}{\delta p} = \dot{q} - \mu = 0, \quad (22.110)$$

$$\frac{\delta S_E}{\delta q} = -\dot{p} - \partial_1 \lambda = 0, \quad (22.111)$$

$$\frac{\delta S_E}{\delta p^*} = \dot{q}^* - \lambda \frac{\partial H^*}{\partial p^*} = 0, \quad (22.112)$$

$$\frac{\delta S_E}{\delta q^*} = -\dot{p}^* - \lambda \frac{\partial H^*}{\partial q^*} = 0, \quad (22.113)$$

$$\frac{\delta S_E}{\delta \lambda} = -G = 0, \quad (22.114)$$

$$\frac{\delta S_E}{\delta \mu} = -F = 0. \quad (22.115)$$

При вариации слагаемого $-\lambda \partial_1 q$ по q возникает также граничное условие на множитель Лагранжа

$$\lambda|_{\sigma_{1,2}} = 0, \quad (22.116)$$

поскольку значения переменной q фиксированы только на пространственноподобной границе $\tau = \tau_{1,2}$ и, следовательно, вариации δq на времениподобной границе $\sigma = \sigma_{1,2}$ произвольны.

Перейдем к анализу уравнений Эйлера–Лагранжа. Решение последней связи (22.115) тривиально. Решение связи (22.114) имеет вид

$$q = \int_{\sigma_1}^{\sigma} d\sigma' H^* + q_0(\tau), \quad (22.117)$$

где q_0 – произвольная функция τ . Отсюда следует, что задание граничных условий $q|_{\sigma_2}$, будет противоречить уравнению связи (22.114), т.к. значение поля q при $\sigma = \sigma_2$ определяется значениями физических полей q^* и p^* во внутренних точках области. Дифференцируя решение (22.117) по τ и используя уравнения (22.112) и (22.113), получим $\dot{q} = \dot{q}_0$. Затем решаем уравнения (22.110) и (22.111) относительно множителей Лагранжа:

$$\mu = \dot{q}_0, \quad \lambda = \lambda_0(\tau), \quad (22.118)$$

где $\lambda_0(\tau)$ – произвольная функция. Таким образом, мы решили уравнения движения для нефизических переменных q, p и множителей Лагранжа λ, μ , и это решение зависит от двух произвольных функций $q_0(\tau)$ и $\lambda_0(\tau)$. Для физических переменных q^* и p^* остаются уравнения (22.112), (22.113), которые имеют вид обычной гамильтоновой

системы. Эта система уравнений действительно воспроизводит уравнения Эйлера–Лагранжа на условный экстремум (22.107), (22.108) при $\lambda_0 = 0$. Заметим, что только это значение согласуется с граничным условием (22.116).

Таким образом, мы решили вариационную задачу для действия (22.103) с заданными граничными условиями для полей q и q^* на границе $\tau = \tau_{1,2}$ прямым способом и методом множителей Лагранжа. Как и следовало ожидать, результат одинаков, а класс решений очень беден.

Однако для расширенного действия (22.109) можно поставить более содержательную вариационную задачу. Предположим, что нефизическая переменная q задана на всей границе $q|_{\sigma_{1,2}}$ и $q|_{\tau_{1,2}}$, а физическое поле q^* – только на пространственноподобной границе $q^*|_{\tau_{1,2}}$. В этом случае вариации δq равны нулю на границе и граничного условия на множитель Лагранжа (22.116) не возникнет. Тогда, при $\lambda_0 \neq 0$, вместо τ можно ввести новую переменную t , определяемую дифференциальным уравнением

$$\frac{dt}{d\tau} = \lambda_0(\tau).$$

Это уравнение определяет координату t с точностью до сдвига на постоянную величину, что несущественно. Тогда уравнения для физических полей (22.112), (22.113) примут вид

$$\begin{aligned} \frac{dq^*}{dt} &= \frac{\partial H^*}{\partial p^*}, \\ \frac{dp^*}{dt} &= -\frac{\partial H^*}{\partial q^*}. \end{aligned} \quad (22.119)$$

Таким образом, мы получили систему гамильтоновых уравнений движения для физических полей, динамика которых определяется гамильтонианом $H^*(q^*, p^*)$. Это показывает, что модель, основанная на действии S_E с множителями Лагранжа допускает постановку более широкого класса вариационных задач, чем исходная задача на условную стационарную точку, и является более содержательной.

Здесь выявляется специфика полевых моделей, поскольку важно, что поля зависят не только от времени τ , но и от пространственной координаты σ . Действительно, если бы связь имела вид $-q + H^* = 0$, то уравнение (22.111) приняло бы вид $-\dot{p} + \lambda_0 = 0$. Откуда следовало бы единственное решение $\lambda_0 = 0$ при $p = 0$. (Мы употребили термины время и пространство, исходя из аналогии с теорией относительности, несмотря на то, что на плоскости $\mathbb{R}^2$ никакой метрики не задано.)

Заметим, что подстановка решения связей в расширенное действие S_E снова приводит к тривиальному действию (22.106), которое не воспроизводит уравнения Эйлера–Лагранжа (22.119). Это показывает, что подстановки решения части уравнений Эйлера–Лагранжа в действие и в оставшиеся уравнения в общем случае не эквивалентны. Это связано с тем, что в общем случае мы не можем накладывать граничные условия на нефизические поля произвольным образом и рассматривать исчезающие на границе вариации. Действительно, вариация δq определяется физическими полями и их вариациями во внутренних точках области:

$$\delta q = \int_{\sigma_1}^{\sigma} d\sigma' \left(\frac{\partial H^*}{\partial q^*} \delta q^* + \frac{\partial H^*}{\partial p^*} \delta p^* \right), \quad (22.120)$$

и в общем случае не равна нулю на границе $\sigma = \sigma_2$. То есть задание граничного условия $q|_{\sigma_2}$ противоречит уравнению связи (22.114).

Изменим постановку вариационной задачи таким образом, чтобы уравнения Эйлера–Лагранжа остались прежними, а исключение нефизических полей в действии приводило бы к новому действию, воспроизводящему уравнения (22.112), (22.113). Введем новое действие

$$S_{\text{ph}} := S_{\text{E}} - \int_{\tau_1}^{\tau_2} d\tau (\lambda q) \Big|_{\sigma=\sigma_1}^{\sigma_2}, \quad (22.121)$$

которое отличается от расширенного действия с множителями Лагранжа граничным слагаемым. Как и раньше, мы считаем, что на границе $\tau_{1,2}$ заданы значения переменных q и q^* . Добавление граничного члена не меняет уравнений Эйлера–Лагранжа, но меняет граничные условия. Этот граничный член подобран таким образом, чтобы компенсировать граничный вклад в вариацию действия, обусловленный вариацией δq . Его необходимо добавить, если мы не хотим получить граничное условие $\lambda|_{\sigma_2} = 0$ при произвольной вариации δq на границе. Теперь нетрудно проверить, что на решениях уравнений (22.111) и (22.114) действие принимает вид

$$S_{\text{ph}} \Big|_{G=0, F=0} = \int_{\tau_1}^{\tau_2} d\tau \int_{\sigma_1}^{\sigma_2} d\sigma p^* \dot{q}^* - \int_{\tau_1}^{\tau_2} d\tau \lambda(\tau, \sigma_2) q(\sigma_2) = \int_{\tau_1}^{\tau_2} d\tau \int_{\sigma_1}^{\sigma_2} d\sigma (p^* \dot{q}^* - \lambda_0 H^*), \quad (22.122)$$

где мы положили $\lambda_0(\tau) := \lambda(\tau, \sigma_2)$. Это действие воспроизводит гамильтоновы уравнения движения для физических полей. При этом мы отбросили слагаемое с $q_0(\tau)$, которое не влияет на уравнения Эйлера–Лагранжа, т.к. в этом действии функция $\lambda_0(\tau)$ рассматривается как заданная и не варьируется. Таким образом, добавление граничного слагаемого не меняет уравнений Эйлера–Лагранжа, позволяет избежать граничных условий на множители Лагранжа и, что самое важное, разрешает постановку решений уравнений Эйлера–Лагранжа непосредственно в действие. При этом возникает нетривиальный гамильтониан для физических полей.

В предыдущем построении была выделена роль точки σ_2 в определении функции $\lambda_0(\tau)$. Это не существенно и связано с выбором начальной точки в решении уравнения связи (22.117). Замена $\sigma_1 \mapsto \sigma_2$ в нижнем пределе этого интеграла приведет к переопределению $\lambda_0(\tau) = \lambda(\tau, \sigma_1)$.

В принципе, можно было бы ограничиться случаем $\lambda = 0$ и не добавлять граничный член. Однако исходное действие в калибровочных моделях и в моделях, инвариантных относительно общих преобразований координат, в канонической формулировке имеют вид расширенного действия S_{E} , содержащего множители Лагранжа. При этом многие решения, важные с физической точки зрения, соответствуют $\lambda \neq 0$. Например, решение Шварцшильда соответствует нетривиальному множителю Лагранжа, роль которого играет функция хода N .

22.9.2 Гамильтонова формулировка

Обозначения в предыдущем разделе были выбраны не случайно. По сути дела модель (22.109) уже записана в гамильтоновой форме, при этом переменные p и p^* являются импульсами, сопряженными координатам q и q^* . Нашей исходной точкой будет полное действие

$$S_{\text{T}} = \int d\tau d\sigma (p \dot{q} + p^* \dot{q}^* - \lambda G), \quad (22.123)$$

которое получится из действия (22.109), если положить $\mu = 0$. Координаты τ и σ будем считать временной и пространственной, соответственно. В рассматриваемом

случае гамильтониан системы задан единственной связью

$$H = \int d\sigma \lambda G, \quad G := -\partial_1 q + H^*(q^*, p^*). \quad (22.124)$$

Будем считать, что интегрирование в (22.123) проводится по всей плоскости τ, σ . При этом все возникающие интегралы предполагаются сходящимися.

Таким образом, модель описывается двумя парами канонически сопряженных переменных q, p и q^*, p^* , на которые наложена одна связь $G = 0$. Уравнения движения имеют прежний вид (22.110)–(22.114) (при $\mu = 0$), где точка обозначает дифференцирование по времени. Нетрудно проверить, что связь G является связью первого рода:

$$[G, G'] = 0,$$

где штрих обозначает, что соответствующие полевые переменные рассматриваются в точке σ' : $G' := G(\tau, \sigma')$. Поэтому на поверхности связей

$$\dot{G} = [G, H] \approx 0,$$

и никаких дополнительных связей в модели не возникает. Наличие связи первого рода означает, что действие (22.123) калибровочно инвариантно. Генератором калибровочных преобразований для канонических переменных является функционал

$$T = \int d\sigma \epsilon G,$$

где $\epsilon(\tau, \sigma)$ – малый параметр локальных преобразований. Бесконечно малые преобразования имеют вид

$$\begin{aligned} \delta q &= [q, T] = 0, \\ \delta p &= [p, T] = -\partial_1 \epsilon, \\ \delta q^* &= [q^*, T] = \epsilon \frac{\partial H^*}{\partial p^*}, \\ \delta p^* &= [p^*, T] = -\epsilon \frac{\partial H^*}{\partial q^*}. \end{aligned} \quad (22.125)$$

Если дополнить эти преобразования преобразованием множителя Лагранжа

$$\delta \lambda = \dot{\epsilon},$$

то, как нетрудно убедиться с помощью прямой подстановки, действие (22.123) является калибровочно инвариантным (см. раздел 22.3). Отметим, что эволюция во времени канонических переменных (22.110)–(22.113) является в рассматриваемом случае калибровочным преобразованием с параметром $\epsilon = \lambda d\tau$.

Согласно второй теореме Нетер калибровочная инвариантность приводит к зависимости уравнений движения:

$$\partial_1 \left(\frac{\delta S_T}{\delta p} \right) + \frac{\delta S_T}{\delta q^*} \frac{\partial H^*}{\partial p^*} - \frac{\delta S_T}{\delta p^*} \frac{\partial H^*}{\partial q^*} - \partial_\tau \frac{\delta S_T}{\delta \lambda} = 0.$$

С математической точки зрения наличие калибровочной инвариантности отражается в том, что решение уравнений движения зависит от произвольной функции

λ , которая ничем не фиксирована. Основным предположением в моделях с калибровочной симметрией является утверждение о том, что все физические наблюдаемые калибровочно инвариантны. В данном случае это означает, что наблюдаемые функции от канонических переменных не зависят от λ . Чтобы исключить произвол в решениях уравнений движения и исключить нефизические переменные необходимо наложить калибровочное условие. Согласно канонической процедуре фиксирования калибровки мы должны наложить одно калибровочное условие по числу связей первого рода. Выберем его в виде

$$F = p - p_0(\sigma) = 0, \quad (22.126)$$

где $p_0(\sigma)$ – произвольная заданная функция только от σ . Скобка Пуассона связи с калибровочным условием имеет вид

$$[G, F'] = \delta'(\sigma' - \sigma),$$

где δ' обозначает производную от δ -функции. Поскольку скобка Пуассона связи с калибровочным условием не обращается в нуль при выполнении связи, то вместе они не представляют собой систему связей первого рода. В то же время пара функций G, F не представляет собой также и систему связей второго рода, поскольку $\det [G, F'] = 0$, т.к. у δ' нетривиально ядро, состоящее из констант.

Канонически сопряженные переменные q, p являются нефизическими переменными, и могут быть исключены из рассмотрения. С этой целью решим их уравнения движения и связь, как это было сделано в предыдущем разделе. В результате получим, что модель описывает одну физическую пару канонически сопряженных полей q^*, p^* с эффективным гамильтонианом $\lambda_0 H^*(q^*, p^*)$. При этом подстановка связи и калибровочного условия в действие (22.123) приводит к неверному результату, который не воспроизводит уравнения движения для физических полей q^* и p^* . Причина этого и решение проблемы то же, что и в предыдущем разделе – к действию необходимо добавить граничный член (22.121). Заметим, что при выполнении уравнения связи граничный член в исходном действии превращается в интеграл по пространству от некоторой гамильтоновой плотности для физических переменных.

Исключение нефизических полей из уравнений движения не зависело от глобальной структуры пространства-времени, т.к. при этом решаются только уравнения движения, связи и калибровочные условия, которые локальны. Нетривиальный эффективный гамильтониан из “нулевого” исходного гамильтониана (22.124) для замкнутых многообразий можно получить следующим образом. Предположим, что пространство-время является прямым произведением $\mathbb{R} \times \mathbb{S}^1$, где первый сомножитель соответствует времени, а второй – пространству. Поскольку окружность $\mathbb{S}^1$ – компактное многообразие без края, то, казалось бы, никаких граничных вкладов не возникает, и можно свободно интегрировать по частям. Однако при внимательном рассмотрении оказывается, что при таком подходе можно потерять много решений уравнений движения, представляющих физический интерес. Опишем это подробнее.

Пусть пространственная координата $\sigma \in [0, 2\pi]$ параметризует окружность $\mathbb{S}^1$. Поскольку физические поля q^* и p^* ничем не ограничены, то их можно считать достаточно гладкими функциями на окружности. В то же время нефизические поля должны удовлетворять уравнению связи (22.104), которое в общем случае не имеет непрерывных решений (22.117) на окружности:

$$q(\tau, 0) = 0, \quad q(\tau, 2\pi) = \int_0^{2\pi} d\sigma H^* \neq 0, \quad \forall \tau \in \mathbb{R}.$$

Кроме того, вариация нефизического поля (22.120) в общем случае не может быть определена как непрерывная функция на окружности. Это значит, что при постановке вариационной задачи необходимо сделать разрез и добавить к действию (22.123) граничный член, который и приведет к нетривиальному эффективному действию для физических переменных. Поэтому, если пространство представляет собой окружность, то в физическом действии и эффективном гамильтониане достаточно просто изменить пределы интегрирования по σ .

Покажем, что нетривиальный эффективный гамильтониан возникает также в более общих калибровках, зависящих от времени явно. Пусть калибровочное условие имеет вид

$$F = p - p_0(\tau, \sigma) = 0,$$

где функция $p_0(\tau, \sigma)$ задана. В этом случае действие на поверхности связей получит дополнительный вклад за счет слагаемого $p\dot{q}$. Нетрудно проверить, что дополнительный вклад не меняет окончательного ответа. Действительно,

$$\Delta S \Big|_{F=0, G=0} = \int d\tau d\sigma p\dot{q} = \int d\tau d\sigma (-\dot{p}q),$$

поскольку интегрирование по частям по времени τ допустимо. Используя уравнение (22.111), определяющее λ , и интегрируя по частям, получим равенство

$$\Delta S \Big|_{F=0, G=0} = \int d\tau d\sigma (-\lambda \partial_1 q) + \int d\tau (\lambda q) \Big|_{\sigma=\sigma_1}^{\sigma_2}.$$

Теперь первое слагаемое воспроизводит эффективный гамильтониан λH^* , где $\lambda = \lambda(\tau, \sigma)$, а второе слагаемое сокращается с граничным членом, добавленным в физическое действие (22.121).

Таким образом, при подстановке связей и калибровочных условий в действие, необходимо проявлять осторожность. Рассмотренный пример показывает, что эффективный гамильтониан и действие для физических полей может полностью определяться граничным членом в исходном действии. Причина этого кроется в том, что связи могут не иметь решений, убывающих в бесконечности, и предположение о финитности вариации неправомерно. К сожалению, для определения явного вида граничных членов необходим глубокий анализ уравнений связей, что не всегда возможно, из-за их сложности.

22.10 О каноническом квантовании

В настоящей монографии рассматриваются классические модели математической физики, в основе которых лежит геометрия. Вместе с этим многие современные модели являются не классическими, а квантовыми, при этом они успешно применяются в физике микромира. Достаточно упомянуть нерелятивистскую квантовую механику, которая хорошо описывает атомные явления, или квантовую электродинамику, прекрасно согласующуюся с экспериментом. В настоящее время широко распространено мнение, что “правильная” модель квантовая, и она справедлива при всех значениях энергии и при всех расстояниях. Считается, что при высоких энергиях, которым соответствуют малые расстояния, существенны квантовые эффекты, а при низких энергиях и больших расстояниях квантовая модель переходит в некоторую классическую. Построение квантовой модели начинается, как правило, с выбора классической, и этот переход называется квантованием. В классических моделях наблюдаемым величинам ставятся в соответствие некоторые функции на многообразиях.

Например, мировая линия частицы. В квантовой механике состояние системы описывается вектором в гильбертовом пространстве, чья эволюция подчиняется уравнению Шредингера, а физическим наблюдаемым ставятся в соответствие некоторые операторы, действующие в этом гильбертовом пространстве.

В настоящее время используется несколько методов квантования. Иногда они приводят к эквивалентным моделям, а иногда – нет. Ниже мы скажем несколько слов о каноническом квантовании моделей математической физики, который считается наиболее последовательным.

При каноническом квантовании гамильтоновых систем без связей канонические переменные в фазовом пространстве $(q^i, p_i) \in \mathbb{N}$, $i = 1, \dots, N$, заменяются на самосопряженные операторы $\hat{q}^i, \hat{p}_i$, действующие в некотором гильбертовом пространстве $\mathbb{H}$, с каноническими одновременными коммутационными соотношениями:

$$i[q^i, p_j] = \hbar \delta_j^i, \quad i[\hat{q}^i, \hat{q}^j] = 0, \quad i[\hat{p}_i, \hat{p}_j] = 0,$$

где $\hbar$ – постоянная Планка, которая имеет размерность действия. Можно сказать, что скобки Пуассона в классической гамильтоновой механике заменяются на коммутаторы операторов гильбертова пространства. В квантовой теории поля гильбертово пространство, как правило, отождествляется с пространством Фока (см., например, [153]). При этом гамильтониан классической системы заменяется на эрмитов оператор, построенный из операторов $\hat{q}$ и $\hat{p}$ формальной заменой канонических переменных на операторы. Здесь возникает неоднозначность в упорядочении операторов, связанная с их некоммутативностью. Необходимо исследовать все возможности и выбрать наиболее подходящее упорядочение так, чтобы теория была самосогласованной. Каждой наблюдаемой величине A также ставится в соответствие некоторый эрмитов оператор с помощью формальной замены: $A(q, p) \mapsto \hat{A}(\hat{q}, \hat{p})$. Причем снова необходимо решить проблему упорядочения. После этого эволюция вектора состояния $\psi \in \mathbb{H}$ определяется уравнением Шредингера:

$$i \frac{\partial \psi}{\partial t} = \hat{H} \psi.$$

Этот переход от классических переменных к операторам называется каноническим квантованием. При этом необходимо следить, чтобы в пределе $\hbar \rightarrow 0$, возникала классическая теория.

В классических гамильтоновых моделях со связями ситуация сложнее. После исключения нефизических степеней свободы путем решения связей второго рода или связей вместе с калибровочными условиями для связей первого рода возникает фазовое пространство для физических степеней свободы без каких либо связей. Поэтому в физическом фазовом подпространстве $\mathbb{M} \subset \mathbb{N}$, $\dim \mathbb{M} = 2M$, мы постулируем канонические коммутационные соотношения:

$$i[\hat{q}^{*a}, \hat{p}_b^*] = \hbar \delta_b^a, \quad i[\hat{q}^{*a}, \hat{q}^{*b}] = 0, \quad i[\hat{p}_a^*, \hat{p}_b^*] = 0, \quad a, b = 1, \dots, M.$$

Так как явное исключение нефизических степеней свободы возможно только в редких случаях, то возникает вопрос: “Как квантовать калибровочные модели на всем фазовом пространстве $\mathbb{N} \supset \mathbb{M}$, поскольку здесь задано две пуассоновы структуры: каноническая и Дирака?” Общепринятый ответ дает скобка Дирака, которая при каноническом квантовании калибровочных систем со связями переходит в коммутатор соответствующих операторов:

$$[q^i, p_j]_D \mapsto i[\hat{q}^i, \hat{p}_j] = \hbar \delta_j^i,$$

а все остальные коммутаторы равны нулю.

Физические наблюдаемые в калибровочных моделях определяются функциями на физическом подпространстве $\mathbb{M} \subset \mathbb{N}$. Поэтому важно, чтобы наблюдаемые не зависели от выбора калибровочных условий (22.31). Можно доказать, что выбор различных калибровочных условий (22.31), удовлетворяющих условию (22.34) приводит к каноническому преобразованию координат q^*, p^* физического подпространства $\mathbb{M} \subset \mathbb{N}$.

В силу единственности решения задачи Коши, любая калибровочно инвариантная функция на поверхности связей $\mathbb{U}$ однозначно определяется своими начальными данными на $\mathbb{M} \subset \mathbb{U}$. Наоборот, любые данные Коши на $\mathbb{M}$ однозначно определяют некоторую калибровочно инвариантную функцию на $\mathbb{U}$. Ясно, что поверхность Коши в $\mathbb{M}$ может быть выбрана произвольно, лишь бы выполнялось условие (22.32). При этом эволюция во времени калибровочно инвариантной функции не будет зависеть от выбора поверхности Коши в $\mathbb{M}$, т.е. от выбора канонических калибровочных условий.

Подробное изложение канонического квантования систем со связями содержится в [149, 150].

Глава 23

Основы общей теории относительности

В настоящей главе мы приступим к изложению основ общей теории относительности, которая в настоящее время рассматривается в качестве основной модели гравитационных взаимодействий. После вступительного раздела, будут написаны основные уравнения и поставлена задача, которая решается в теории гравитации.

23.1 Пространство-время, метрика и гравитация

В основе общей теории относительности лежит ряд постулатов. Выделим среди них пять, на наш взгляд, основных.

1. Пространство-время $\mathbb{M}$, в котором мы живем, является четырехмерным многообразием.
2. Гравитационное взаимодействие между материальными телами описывается метрикой g лоренцевой сигнатуры, $\text{sign } g = (+ - - -)$, заданной на $\mathbb{M}$.
3. Метрика пространства-времени удовлетворяет уравнениям Эйнштейна.
4. Пробная точечная частица, собственным гравитационным полем которой в данной задаче можно пренебречь, под действием только гравитационного поля движется по экстремалиям (геодезическим) пространства-времени $(\mathbb{M}, g)$.
5. Пространство-время $(\mathbb{M}, g)$ должно быть максимально продолжено вдоль экстремалей (геодезических).

Первые два постулата являются “кинематическими”. Из них следует, что в общей теории относительности все законы природы формулируются на четырехмерном псевдоримановом многообразии (пространстве-времени) $(\mathbb{M}, g)$. Если выбрана некоторая система координат x^α , $\alpha = 0, 1, 2, 3$, то метрика имеет вид $g = dx^\alpha \otimes dx^\beta g_{\alpha\beta}$, $\text{sign } g_{\alpha\beta} = \text{diag } (+ - - -)$. В общем случае, если пространство-время топологически нетривиально, система координат может быть выбрана только локально.

Первые две аксиомы важны, поскольку позволяют описывать окружающий нас мир с помощью некоторого набора полей и формулировать законы природы в виде системы дифференциальных уравнений на $\mathbb{M}$. Этот подход оказался самым плодотворным в последние три столетия.

Третья и четвертая аксиома являются “динамическими”. В общей теории относительности постулируется, что метрика на $\mathbb{M}$ должна удовлетворять уравнениям Эйнштейна (23.2). Тем самым компоненты метрики пространства-времени удовлетворяют некоторой системе уравнений движения так же, как и все другие поля. Это – очень важное отличие общей теории относительности от специальной, где метрика Лоренца $\eta_{\alpha\beta} := \text{diag}(+ - - -)$ в пространстве Минковского $\mathbb{R}^{1,3}$ была постулирована.

Как мы увидим в дальнейшем, в правой части уравнений Эйнштейна стоит тензор энергии-импульса полей материи. Выбор полей материи зависит от рассматриваемой модели. Это может быть, например, сплошная среда, точечные массивные частицы, электромагнитное поле или что то еще. Возможны также произвольные комбинации полей материи.

Сама по себе система уравнений Эйнштейна неполна. Если мы выбрали какой либо набор полей материи, то уравнения Эйнштейна необходимо дополнить уравнениями движения полей материи. Вид дополнительных уравнений зависит от рассматриваемой задачи.

Четвертый постулат говорит о следующем. Допустим, что мы выбрали некоторый набор полей материи, записали и решили полную систему уравнений для метрики и полей материи. В результате мы получим псевдориманово многообразие $(\mathbb{M}, g)$, на котором заданы также поля материи. Теперь допустим, что к нашей системе добавлена точечная массивная частица, масса которой настолько мала, что она не влияет на решение уравнений Эйнштейна. То есть мы пренебрегаем собственным гравитационным полем частицы. Такую частицу назовем *пробной*. Тогда возникает вопрос, по какой траектории будет двигаться пробная частица под действием только гравитационных сил? Ответ на этот вопрос дает четвертый постулат: пробная частица будет двигаться по $\mathbb{M}$ вдоль экстремалей (геодезических), определяемых метрикой g . Мы также предполагаем, что безмассовые частицы (например, фотоны) также распространяются вдоль светоподобных (нулевых) экстремалей (геодезических).

На четвертой аксиоме основано экспериментальное подтверждение общей теории относительности. Два классических теста: смещение перигелия Меркурия и отклонение лучей света в поле тяготения основаны на анализе геодезических для решения Шварцшильда, о котором речь пойдет позже. Третий классический тест – красное смещение частоты электромагнитного излучения – это следствие второй аксиомы.

Обсудим пятую аксиому. В общей теории относительности не предполагается, что пространство-время снабжено какой либо линейной структурой, как это было в механике Ньютона и специальной теории относительности.

В общем случае глобальная структура (топология) пространства-времени $\mathbb{M}$ может быть нетривиальной и отличаться от пространства Минковского. Поскольку глобально структура $\mathbb{M}$ не фиксирована, то в моделях гравитации вводится новое требование. Пространство-время, по определению, должно быть максимально продолжено вдоль геодезических (экстремалей). Это значит, что любая геодезическая в пространстве-времени может быть продолжена либо до бесконечного значения канонического параметра в обе стороны, либо при конечном значении канонического параметра она попадет в сингулярную точку, где какой либо из геометрических инвариантов обращается в бесконечность. Поскольку канонический параметр вдоль экстремалей определен с точностью до линейных преобразований (см., главу 16), то данное требование инвариантно, т.е. не зависит от выбора системы координат.

Замечание. Требование максимального продолжения пространства-времени вдоль геодезических нельзя заменить на более жесткое требование геодезической полноты,

т.к. многие важные точные решения уравнений Эйнштейна не являются геодезически полными. Например, для решений, описывающих черные дыры, времениподобные геодезические линии достигают сингулярного края (черной дыры), на котором квадрат тензора кривизны обращается в бесконечность при конечном значении канонического параметра (собственного времени). $\square$

Приведенные аксиомы выделены, потому что лежат в основе любой модели, построенной в рамках общей теории относительности. Их недостаточно для построения конкретной модели гравитации, т.к. необходимо выбрать поля материи и дополнить уравнения Эйнштейна. При этом используются дополнительные аксиомы, которые мы не выделяем, поскольку их столько же, сколько и моделей.

Остановимся на некоторых общих свойствах и терминологии.

Все поля, кроме метрики, мы делим на поле излучения, описывающее безмассовое электромагнитное поле, и поля материи, описывающие массивные частицы. Многие современные модели математической физики содержат дополнительные безмассовые поля, например, глюоны, существование которых в природе пока экспериментально не подтверждено. Для определенности, все безмассовые поля мы также относим к полям излучения. Таким образом, все поля, за исключением метрики, мы делим на два класса: *излучение* (безмассовые поля или частицы) и *поля материи* (массивные поля или частицы).

Теперь скажем несколько слов о гравитационном взаимодействии. Для описания движения планет в солнечной системе с хорошей точностью используется механика Ньютона и закон всемирного тяготения. Мы говорим, что между планетами действуют гравитационные силы, которые определяют их движение. При этом движение происходит в плоском трехмерном евклидовом пространстве $\mathbb{R}^3$, а время играет роль параметра. Основное свойство гравитационного взаимодействия заключается в том, что движение пробной частицы, при заданных начальных условиях, не зависит от ее массы.

Пример 23.1.1. Ускорение свободного падения на Земле не зависит от массы падающего тела. Это утверждение в настоящее время экспериментально проверено с высокой степенью точности. $\square$

Независимость ускорения от массы частицы означает, что при одних и тех же начальных условиях траектории и мировые линии пробных частиц разной массы совпадают.

Замечание. В электродинамике траектория заряженной частицы зависит от ее заряда. Если электромагнитное поле задано, то траектория частицы однозначно определяется массой, зарядом и начальными условиями. $\square$

Рассмотрим движение пробной частицы в специальной теории относительности в инерциальной системе отсчета. По определению, если гравитационное поле отсутствует и на частицу не действуют никакие другие силы, то она движется равномерно и прямолинейно. Теперь рассмотрим движение той же частицы, но в неинерциальной системе отсчета, которая движется, например, с постоянным ускорением относительно инерциальной системы отсчета. В этой системе координат свободная частица движется с ускорением и наблюдатель в этой системе отсчета может сказать (если не наблюдает за другими телами), что его система инерциальна, а частица движется в постоянном и однородном гравитационном поле, которое и вызывает ускорение. При этом ускорение не зависит от массы частицы и определяется только неинерциальной

системой координат. Поэтому часто формулируют

Принцип эквивалентности. Движение свободной пробной частицы в неинерциальной системе отсчета такое же, как движение частицы в инерциальной системе отсчета, но находящейся под действием некоторого гравитационного поля. $\square$

Для одной частицы с этим можно согласиться, рассмотрев систему координат, в которой она покоится. Однако, если рассматривать две или более частиц, то никакой эквивалентности нет. А именно, обратное утверждение неверно. Если пробные частицы движутся в гравитационном поле, то в общем случае не существует такой системы координат, в которой они бы двигались равномерно и прямолинейно.

В инерциальной (декартовой) системе координат в пространстве Минковского метрика диагональна и имеет постоянные компоненты, $\eta_{\alpha\beta} := \text{diag}(+ - - -)$. В такой системе координат все экстремали и только они являются прямыми линиями. Поэтому можно сказать, что свободная пробная частица движется вдоль одной из экстремалей пространства Минковского. Если перейти в неинерциальную (криволинейную) систему координат, то в общем случае метрика перестанет быть диагональной, и ее компоненты станут зависеть от координат точки пространства-времени. В этой системе координат траектория свободной частицы уже не будет выглядеть прямолинейной, а движение – равномерным. Тем не менее траектория, конечно, будет оставаться экстремалью пространства Минковского, т.к. понятие экстремали инвариантно и не зависит от выбора системы координат.

В дифференциальной геометрии метрика является инвариантным объектом и определяется независимо от выбора системы координат. Метрика Лоренца не зависит от точки пространства Минковского. Однако ее компоненты могут быть непостоянными функциями от координат точки: это зависит от системы координат.

Таким образом, утверждение о том, что свободная пробная частица движется в пространстве Минковского вдоль одной из экстремалей инвариантно относительно выбора системы координат и лежит в основе перехода от механики Ньютона к общей теории относительности. Как уже было сказано, в общей теории относительности мы предполагаем, что пространство-время представляет собой четырехмерное многообразие M , на котором задана метрика лоренцевой сигнатуры. Мы постулируем, что любая пробная частица движется вдоль одной из экстремалей пространства-времени. Этот постулат согласуется с упомянутыми выше свойствами гравитационного взаимодействия: мировая линия пробной частицы не зависит от ее массы. При этом принцип эквивалентности является лишь наводящим соображением о том, что метрика с нетривиальными компонентами описывает гравитационное взаимодействие.

Тем самым метрика пространства-времени в общей теории относительности играет выделенную роль. Мы считаем, что метрика описывает гравитационные взаимодействия материальных тел и излучения. А именно, если частица движется в плоском пространстве-времени, которое изометрично пространству Минковского $\mathbb{R}^{1,3}$ с метрикой Лоренца, то на нее не действуют гравитационные силы. В этом случае частица в инерциальной системе координат движется равномерно и прямолинейно. Если гравитационное поле нетривиально, то частицы (массовые и безмассовые) движутся по экстремалиям в искривленном пространстве-времени, т.е. по многообразию M с метрикой g и связностью Леви–Чивиты Γ , для которой тензор кривизны отличен от нуля. В этом случае отсутствует понятие инерциальной системы отсчета, а экстремали отличаются от прямых линий.

Поскольку в пространстве-времени M задана метрика, то она однозначно определяет связность Леви–Чивиты или символы Кристоффеля. Это позволяет использовать аппарат ковариантного дифференцирования для построения инвариантов и за-

писи ковариантных уравнений движения. Введение связности Леви–Чивиты на многообразии M является постулатом общей теории относительности. То есть в теории тяготения Эйнштейна мы постулируем, что кручение и неметричность аффинной связности тождественно равны нулю.

В настоящее время теория тяготения Эйнштейна имеет много обобщений. Большой класс таких обобщений представляют собой модели, в которых на многообразии M помимо метрики задается также независимая аффинная связность Γ с нетривиальным кручением T и неметричностью Q . Эти обобщения естественны с геометрической точки зрения, т.к. метрика и аффинная связность являются совершенно независимыми геометрическими объектами. В общем случае, даже если ограничиться инвариантными лагранжианами, приводящими к уравнениям движения второго порядка, существует очень много возможностей для построения моделей, которые в настоящее время не исследованы в полной мере.

Считается, что общая теория относительности согласуется со всеми наблюдательными данными. Однако, поскольку мы не знаем экспериментальных следствий упомянутых выше геометрических обобщений теории тяготения, говорить о том, что они противоречат экспериментальным данным нельзя. Различные геометрические обобщения теории тяготения Эйнштейна представляют самостоятельный математический интерес и могут быть полезны при построении квантовой теории гравитации и единых моделей. В современной математической физике такие модели привлекают исследователей постоянно со времен создания общей теории относительности. Достаточно отметить, что геометрическими обобщениями общей теории относительности занимались А. Эйнштейн, Г. Вейль, Э. Шредингер и многие другие выдающиеся физики.

23.2 О постановке задач в теории гравитации

Отметим специфику задач, возникающих при рассмотрении произвольного функционала действия, инвариантного относительно общих преобразований координат и сохраняющего метрику. Уравнения Эйлера–Лагранжа записываются и решаются в произвольной, но фиксированной системе координат, т.е. локальны. Допустим, что мы поставили краевую задачу и нашли какое-то решение уравнений Эйлера–Лагранжа во всем $\mathbb{R}^n$ с заданными граничными условиями. Это решение может оказаться геодезически неполным. Поскольку исходное действие инвариантно относительно общих преобразований координат, то без потери общности можно отобразить все $\mathbb{R}^n$, и тем самым найденное решение, на ограниченную область, например, в шар конечного радиуса. В связи с этим возникает вопрос нельзя ли найденное решение продолжить, т.е. существует ли решение в большей области, сужение которого совпадает с уже найденным решением в шаре. Для ответа на этот вопрос необходимо инвариантное определение глобальности решения, которое дается через полноту геодезических и экстремалей.

В теории гравитации, основанной на аффинной геометрии, так же как и в общей теории относительности, понятие геодезической и экстремали является чрезвычайно важным, поскольку позволяет установить связь между локальными решениями уравнений движения и глобальными свойствами многообразий.

Определение. Назовем многообразие M с заданной аффинной геометрией, т.е. тройку (M, g, Γ) , *полным*, если любую геодезическую и экстремаль в M можно продолжить до бесконечного значения канонического параметра в обе стороны. На практике

часто локальное решение уравнений движения нельзя продолжить до полного решения, поскольку возможно появление сингулярностей. Назовем точку многообразия $\mathbb{M}$ *сингулярной*, если в этой точке по крайней мере одно скалярное поле, построенное из метрики и (или) аффинной связности, включая их производные обращается в бесконечность. $\square$

В этом определении важно, чтобы в сингулярной точке именно скалярная комбинация геометрических объектов обращалась в бесконечность, поскольку наряду с истинными сингулярностями могут существовать координатные сингулярности, связанные с неудачным выбором системы координат. Например, хорошо известна координатная особенность метрики на горизонте событий для решения Шварцшильда в координатах Шварцшильда, от которой можно избавиться, перейдя, например, к системе координат Эддингтона–Финкельштейна или Крускала–Секереша (см. главу 34). Простейшими функциями, определяющими положение сингулярностей являются скалярная кривизна и квадрат тензора кручения, имеющие одинаковую размерность. Примером истинной сингулярности может служить черная дыра в решении Шварцшильда, в которой квадрат тензора кривизны обращается в бесконечность.

В связи с возможным существованием сингулярностей оказывается полезным понятие максимально продолженного многообразия.

Определение. Многообразие с заданной аффинной геометрией назовем *максимально продолженным*, если любую геодезическую и экстремаль можно либо продолжить до бесконечного значения канонического параметра, либо они продолжаются до сингулярной точки при конечном значении канонического параметра. Соответствующую тройку $(\mathbb{M}, g, \Gamma)$ назовем *глобальным решением* в теории гравитации. $\square$

Слово “продолжение” в этом определении связано с тем, что в моделях математической физики, как правило, метрика и кручение сначала находятся только локально в какой-либо области, как решение уравнений движения в заданной системе координат, а затем, если необходимо, это решение гладко продолжается в соседние области до тех пор, пока дальнейшее продолжение становится невозможным либо вследствие сингулярностей, либо в связи с тем, что полученное решение станет полным.

Важно отметить, что определение максимально продолженного многообразия является инвариантным и не зависит от выбора локальной системы координат. Это следует из того, что канонический параметр определяется однозначно с точностью до линейного преобразования и не зависит от системы координат.

Пример 23.2.1. Рассмотрим евклидову плоскость $\mathbb{R}^2$ в полярных координатах r, φ . Будем считать, что полярный угол меняется в бесконечном интервале, $0 < \varphi < \infty$.

Определение. *Гиперболической спиралью* называется кривая на евклидовой плоскости, которая в полярных координатах задана уравнением

$$r = \frac{1}{\varphi}, \quad 0 < \varphi < \infty. \quad \square \quad (23.1)$$

Гиперболическая спираль соединяет бесконечно удаленную точку, $r = \infty$, и начало координат, $r = 0$, вокруг которого она наматывается бесконечное число раз. Если полярный угол меняется в интервале $\varphi \in (\varphi_1, \varphi_2) \subset (0, \infty)$, то ее длина определяется интегралом

$$l(\varphi_1, \varphi_2) = \int_{\varphi_1}^{\varphi_2} \sqrt{dr^2 + r^2 d\varphi^2} = \int_{\varphi_1}^{\varphi_2} d\varphi \frac{\sqrt{1 + \varphi^2}}{\varphi^2}.$$

Этот интеграл расходится на обоих концах гиперболической спирали, как при $\varphi \rightarrow 0$, так и при $\varphi \rightarrow \infty$. Расходимость интеграла при $\varphi \rightarrow \infty$ естественна, т.к. любая кривая, уходящая в бесконечность, имеет бесконечную длину. Начало координат является конечной точкой плоскости, т.к. любая экстремаль, проходящая через эту точку, проходит ее при конечном значении канонического параметра. В то же время не любая кривая, подходящая к этой точке, имеет конечную длину. Пример дает гиперболическая спираль, которая достаточно быстро наматывается вокруг начала координат. $\square$

Данный пример показывает, что в определении полноты многообразий требование полноты экстремалей, а не произвольных кривых является существенным.

С физической точки зрения требование полноты или максимального продолжения пространства-времени M является естественным. Действительно, если рассмотреть движение точечной частицы в сопутствующей системе отсчета, в которой время является каноническим параметром, то естественно предположить, что эволюция продолжается либо бесконечно долго, либо обрывается в сингулярной точке.

Отметим, что полное многообразие не может иметь края. При этом существуют две возможности: либо пространство-время M некомпактно, либо компактно и без края. В первом случае геометрические инварианты в бесконечности могут стремиться как к конечным, так и бесконечным значениям. В космологии принята следующая терминология. Если все пространственные сечения пространства-времени M некомпактны, то вселенная *бесконечна*. Если все пространственные сечения компактны и без края, то говорят, что вселенная *замкнута*. При наличии сингулярностей, соответствующих конечному значению канонического параметра, сингулярные точки образуют край пространства-времени, находящийся на конечном расстоянии (при конечных значениях канонического параметра).

Чтобы найти максимально продолженное решение в общей теории относительности, необходимо пройти несколько этапов: 1) решить уравнения Эйнштейна для метрики в некоторой области; 2) найти и проанализировать полноту всех геодезических; 3) если область, где найдено решение, оказалась неполной, то продолжить решение. Первые два этапа очень сложны, поскольку предполагают решение нелинейных систем дифференциальных уравнений. Последний этап также сложен. Его можно осуществить по крайней мере двумя способами. Либо перейти в новую систему координат, охватывающую большую область, либо найти решение в соседней области, а затем доказать гладкость склейки. Однако общего конструктивного метода продолжения решений в настоящее время не существует. В разделе 30 описан метод конформных блоков построения лоренцевых поверхностей в целом, если метрика допускает один вектор Киллинга.

В общей теории относительности существуют только отдельные примеры максимально продолженных многообразий. Например, расширение Крускала–Секереша решения Шварцшильда, которое будет обсуждаться в главе 34.

23.3 Действие Гильберта–Эйнштейна

В общей теории относительности постулируется, что пространство-время является псевдоримановым многообразием M , $\dim M = n$, с метрикой лоренцевой сигнатуры $g_{\alpha\beta}$. При этом считается, что метрика описывает гравитационные взаимодействия. Мы намеренно не ограничиваемся наиболее интересным случаем $n = 4$, потому что

модели гравитации в большем и меньшем числе измерений также важны для приложений.

Мы рассматриваем метрику пространства-времени в качестве одной из полевых переменных и постулируем для нее *уравнения Эйнштейна*:

$$\kappa \left(\tilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \tilde{R} \right) + g_{\alpha\beta} \frac{n-2}{2} \Lambda = -\frac{1}{2} T_{\alpha\beta}. \quad (23.2)$$

В левой части этой системы уравнений для метрики стоит *тензор Эйнштейна*

$$G_{\alpha\beta} := \tilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \tilde{R}, \quad (23.3)$$

умноженный на *гравитационную постоянную* κ , и *космологическая постоянная* Λ . В правой части уравнений Эйнштейна стоит *тензор энергии-импульса материи* $T_{\alpha\beta}$. Эти уравнения при $\Lambda = 0$ и $n = 4$ были впервые предложены А. Эйнштейном в статье [154].

В дальнейшем, для краткости, мы часто будем говорить “тензор энергии-импульса материи”, подразумевая под этим также возможное излучение.

Тензор энергии-импульса материи зависит от рассматриваемой модели, и в общем случае уравнения Эйнштейна необходимо дополнить уравнениями для полей материи. То есть сама по себе система уравнений Эйнштейна неполна.

Замечание. В уравнении (23.2) мы оставили знак тильды, чтобы подчеркнуть, что тензор кривизны строится только по метрике при нулевом кручении и неметричности. То есть метрика g на $\mathbb{M}$ определяет связность Леви–Чивиты (символы Кристоффеля), которые в свою очередь задают тензор кривизны. $\square$

Вклад космологической постоянной в уравнения Эйнштейна (23.2) можно перенести в правую часть

$$\kappa \left(\tilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \tilde{R} \right) = -\frac{1}{2} (T_{\alpha\beta} + T_{\Lambda\alpha\beta}),$$

где

$$T_{\Lambda\alpha\beta} := (n-2)\Lambda g_{\alpha\beta}$$

и рассматривать его как дополнение к тензору энергии-импульса материи $T_{\alpha\beta}$. Сравнивая это выражение с тензором энергии-импульса непрерывной среды (23.125), который будет рассмотрен позже, его можно интерпретировать, как вклад среды с постоянными давлением $\mathcal{P} = -(n-2)\Lambda$ и плотностью энергии противоположного знака $\mathcal{E} = -\mathcal{P} = (n-2)\Lambda$. Разность знаков давления и плотности энергии не позволяет интерпретировать космологическую постоянную, как распределение некоторой обычной материи. В космологических моделях вселенной ее часто связывают с наличием темной энергии.

Обсудим некоторые общие свойства уравнений Эйнштейна и введем терминологию.

Тензор Эйнштейна (23.3) инвариантен относительно вейлевского преобразования метрики с постоянным параметром

$$g_{\alpha\beta} \mapsto k g_{\alpha\beta}, \quad k = \text{const} \neq 0.$$

Эти преобразования меняют длины векторов, но сохраняют углы между ними.

Уравнения Эйнштейна при заданном тензоре энергии-импульса представляют собой систему из $n(n+1)/2$, где n – размерность пространства-времени, нелинейных уравнений в частных производных второго порядка для метрики. В частности, в четырехмерном пространстве-времени мы имеем десять уравнений. Уравнения Эйнштейна чрезвычайно сложны, и в настоящее время известны лишь отдельные классы решений, часть из которых будет обсуждаться в дальнейшем.

Уравнения Эйнштейна можно переписать в другом виде. След равенства (23.2) эквивалентен уравнению

$$\kappa \tilde{R} = n\Lambda + \frac{1}{n-2}T_M,$$

где $T_M := T_{M\alpha}{}^\alpha$ – след тензора энергии-импульса материи. Исключив скалярную кривизну из (23.2) с помощью этого равенства, получим эквивалентную систему уравнений

$$\kappa \tilde{R}_{\alpha\beta} - \Lambda g_{\alpha\beta} = -\frac{1}{2}\rho_{\alpha\beta}. \quad (23.4)$$

где

$$\rho_{\alpha\beta} := T_{M\alpha\beta} - \frac{1}{n-2}g_{\alpha\beta}T_M.$$

Пространство-время называется *пустым*, если тензор энергии-импульса материи всюду равен нулю. В этом случае уравнения Эйнштейна (23.4) принимают вид

$$\kappa \tilde{R}_{\alpha\beta} = \Lambda g_{\alpha\beta}. \quad (23.5)$$

Это – *вакуумные уравнения Эйнштейна* с космологической постоянной. Отсюда следует, что скалярная кривизна пустого пространства постоянна:

$$\tilde{R} = \frac{n\Lambda}{\kappa}.$$

Замечание. Коэффициент перед космологической постоянной в уравнениях Эйнштейна (23.2) подобран таким образом, чтобы вакуумные уравнения Эйнштейна имели вид (23.5) и не зависели от размерности пространства-времени. $\square$

При ненулевой космологической постоянной уравнения (23.5) означают, что тензор Риччи пропорционален метрике. Частным случаем таких пространств являются пространства постоянной кривизны. При нулевой космологической постоянной, $\Lambda = 0$, пустое пространство является *Риччи плоским*:

$$\tilde{R}_{\alpha\beta} = 0. \quad (23.6)$$

Следовательно, в этом случае скалярная кривизна также равна нулю, $\tilde{R} = 0$.

Пример 23.3.1. Для метрики Лоренца тензор кривизны равен нулю. Следовательно, пространство Минковского является пространством постоянной – нулевой – кривизны. В частности, оно является Риччи плоским. Ясно, что метрика Лоренца удовлетворяет вакуумным уравнениям Эйнштейна с нулевой космологической постоянной. $\square$

Пример 23.3.2. В дальнейшем мы увидим, что вакуумные уравнения Эйнштейна допускают решения в виде плоских волн. Для таких решений тензор Риччи равен нулю, но полный тензор кривизны отличен от нуля. $\square$

В двумерном пространстве-времени полный тензор кривизны однозначно восстанавливается по скалярной кривизне, а в трехмерном – по тензору Риччи и скалярной кривизне. Следовательно, полный тензор кривизны равен нулю, если выполнено условие (23.6). Это значит, что двумерное и трехмерное Риччи плоское пространство локально является пространством Минковского. То есть может быть либо пространством Минковского, либо цилиндром или тором. В четырех измерениях и выше равенства нулю тензора Риччи недостаточно для обращения в нуль полного тензора кривизны.

Пример 23.3.3. Решение Шварцшильда, описывающее черную и белую дыру, является Риччи плоским, но полный тензор кривизны отличен от нуля. А именно, отличен от нуля тензор Вейля (17.14). $\square$

Физическая интерпретация уравнений Эйнштейна при нулевой космологической постоянной следующая. В общей теории относительности постулируется, что метрика пространства-времени не является метрикой Лоренца, а находится, как решение уравнений Эйнштейна. Таким образом, пространство-время представляет собой псевдориманово многообразие с метрикой специального вида, удовлетворяющей уравнениям (23.2). Эти пространства называются *пространствами Эйнштейна*. Следующий постулат состоит в том, что пробные частицы под действием гравитационных сил двигаются по экстремалиям в пространстве Эйнштейна. При этом в правой части уравнений Эйнштейна подразумевается тензор энергии-импульса всей остальной материи. При этом мы говорим следующее. Пустое пространство при нулевой космологической постоянной и отсутствии гравитационных волн является пространством Минковского, и точечные частицы двигаются по прямым линиям. Это соответствует отсутствию сил тяготения. При наличии полей материи в уравнениях Эйнштейна появляется нетривиальная правая часть, что приводит к тому, что пространство-время становится нетривиальным псевдоримановым многообразием. В этом пространстве-времени экстремали уже не являются прямыми линиями, что интерпретируется, как наличие сил тяготения. Мы говорим, что пробная частица движется в поле тяготения, созданном остальной материей. При этом закон всемирного тяготения является следствием уравнений Эйнштейна в определенном приближении, которое рассмотрено в разделе 23.10.

Существующие наблюдательные данные свидетельствуют о том, что в отсутствие сил тяготения пространство-время в масштабах солнечной системы близко к пространству Минковского. Это значит, что если космологическая постоянная существует, то является в определенном смысле малой величиной. Отметим, что равенство или неравенство космологической постоянной нулю имеет принципиальное значение. Действительно, наличие даже малой космологической постоянной приводит к тому, что метрика Лоренца уже не будет удовлетворять вакуумным уравнениям Эйнштейна.

Теперь обсудим принцип наименьшего действия для уравнений Эйнштейна. Левую часть уравнений (23.2) можно получить из *действия Гильберта–Эйнштейна* [155, 156]:

$$S_{\text{HE}} = \int_{\mathbb{M}} dx \sqrt{|g|} \left(\kappa \tilde{R} - (n-2)\Lambda \right), \quad (23.7)$$

где интегрирование ведется по всему пространству-времени $\mathbb{M}$. Конечно, мы предполагаем, что интеграл сходится. Это действие было впервые предложено Д. Гильбертом в 1915 году в четырехмерном пространстве-времени. Он предложил действие в

более общем виде, включающем также электромагнитное поле [155]. Несколько позже А. Эйнштейн тоже рассмотрел это действие для вывода уравнений общей теории относительности в такой системе координат, где $\det g_{\alpha\beta} = 1$ [156].

В следующем разделе мы покажем, что вариационная производная действия Гильберта–Эйнштейна по метрике имеет вид

$$\frac{\delta S_{\text{HE}}}{\delta g_{\alpha\beta}} = -\sqrt{|g|}\kappa \left(\tilde{R}^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}\tilde{R} \right) - \sqrt{|g|}\frac{n-2}{2}\Lambda g^{\alpha\beta}. \quad (23.8)$$

При доказательстве этого равенства были отброшены все граничные вклады, возникающие при интегрировании по частям.

При наличии полей материи чаще удобнее варьировать по обратной метрике, что приводит к изменению знака вариационной производной:

$$\frac{\delta S_{\text{HE}}}{\delta g^{\alpha\beta}} = \sqrt{|g|}\kappa \left(\tilde{R}_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\tilde{R} \right) + \sqrt{|g|}\frac{n-2}{2}\Lambda g_{\alpha\beta}. \quad (23.9)$$

Полное действие для гравитационного поля и полей материи имеет вид суммы

$$S = S_{\text{HE}} + S_{\text{M}}, \quad (23.10)$$

где S_{M} – действие для полей материи. Обычно действие для полей материи в теории гравитации получают путем *минимальной подстановки*: выбирают лоренц-инвариантное действие в пространстве Минковского, заменяют лоренцеву метрику на псевдориманову $\eta_{\alpha\beta} \mapsto g_{\alpha\beta}(x)$, обычные производные – на ковариантные $\partial_\alpha \mapsto \tilde{\nabla}_\alpha$, и умножают лагранжиан на определитель репера $\sqrt{|g|}$, чтобы получить инвариантную меру интегрирования. В результате получим действие для полей материи, инвариантное относительно общих преобразований координат. Сравнивая правую часть уравнений Эйнштейна (23.2) с вариационной производной (23.9), получаем выражение для тензора энергии-импульса материи

$$T_{\text{M}\alpha\beta} := \frac{2}{\sqrt{|g|}} \frac{\delta S_{\text{M}}}{\delta g^{\alpha\beta}}. \quad (23.11)$$

Эту вариационную производную часто принимают за определение *тензора энергии-импульса* полей материи в общей теории относительности. При таком определении тензор энергии-импульса всегда симметричен. В ряде случаев, например, для скалярного и калибровочного полей, это определение совпадает с ковариантным обобщением канонического тензора энергии-импульса, т.е. получается из выражения (20.43) путем минимальной подстановки. Однако в общем случае это не так, потому что действие для полей материи (например, спинорных полей) не всегда может быть выражено через метрику.

Размерности постоянных и полей

В теории поля важную роль играет анализ размерностей. Он помогает контролировать проведение вычислений путем сравнения размерностей различных слагаемых и в некоторых случаях делать общие выводы. В квантовой теории поля, например, размерность констант связи позволяет судить о перенормируемости моделей. Здесь принят упрощенный вариант подсчета размерностей, где все измеряется в единицах длины. Однако, для сравнения общей теории относительности с теорией гравитации

Ньютона нам этого будет недостаточно. Поэтому мы опишем оба подхода к определению размерностей.

В квантовой теории поля, по определению, действие и метрика являются безразмерными величинами, а координаты имеют размерность длины (скорость света мы принимаем за единицу):

$$[S_{\text{нб}}] := [g_{\alpha\beta}] := 1, \quad [x^\alpha] := l.$$

При этом мы ограничиваем класс допустимых преобразований координат. Например, декартовы координаты в $\mathbb{R}^3$ имеют размерность длины по определению. Однако в сферической системе координат углы являются безразмерными. Тем самым мы допускаем только такие преобразования систем координат, которые не меняют их размерностей. Размерность компонент тензора кривизны легко вычислить:

$$[\tilde{R}_{\alpha\beta\gamma\delta}] = [\tilde{R}_{\alpha\beta}] = [\tilde{R}] = l^{-2},$$

т.к. они содержат две производные. Отсюда следует, что гравитационная и космологическая постоянные являются размерными величинами:

$$[\kappa] = l^{2-n}, \quad [\Lambda] = l^{-n}.$$

Для контроля вычислений такого описания размерностей, как правило, достаточно.

Теперь опишем размерности величин, как это принято в системе СГС, которые нам понадобятся в дальнейшем. При этом мы считаем, что размерность пространства-времени равна четырем, $n = 4$. Исходными являются размерности массы (грамм), расстояния (сантиметр) и времени (секунда):

$$[m] := \text{г}, \quad [x^\mu] := \text{см}, \quad [t] := \text{сек},$$

где индекс $\mu = 1, 2, 3$ пробегает только пространственные значения. Координата $x^0 := ct$, где c – скорость света, также измеряется в сантиметрах. По определению, компоненты метрики безразмерны:

$$[g_{\alpha\beta}] = [g^{\alpha\beta}] := 1.$$

Поскольку кривизна содержит две производные по координатам, то

$$[\tilde{R}_{\alpha\beta\gamma\delta}] = [\tilde{R}_{\alpha\beta}] = [\tilde{R}] = \frac{1}{\text{см}^2}.$$

Действие, как это принято в механике, имеет ту же размерность, что и произведение импульса на скорость $p dq$ или энергии на время $E dt$:

$$[S] := \frac{\text{г} \cdot \text{см}^2}{\text{сек}}.$$

Учитывая, что в действии $[dt d^3x] = \text{сек} \cdot \text{см}^3$, определяем размерность гравитационной постоянной:

$$[\kappa] = \frac{\text{г} \cdot \text{см}}{\text{сек}^2}. \quad (23.12)$$

В разделе 23.10 мы сравним эту гравитационную постоянную с той, которая входит во всемирный закон тяготения. Наконец, размерность космологической постоянной равна

$$[\Lambda] = \frac{\text{г}}{\text{см} \cdot \text{сек}^2}.$$

В дальнейшем для упрощения формул мы часто будем полагать $c = 1$ и $\kappa = 1$. Там, где это необходимо, степени скорости света легко восстановить, исходя из соображений размерности.

23.4 Вариация действия Гильберта–Эйнштейна

Докажем равенство (23.8) в более общем виде, который полезен при рассмотрении моделей, основанных на геометрии Римана–Картана или аффинной геометрии. А именно, рассмотрим инвариантное действие

$$S = \int_{\mathbb{M}} dx L = \int_{\mathbb{M}} dx \sqrt{|g|} \varphi R, \quad (23.13)$$

зависящее от скалярного поля $\varphi(x) \in \mathcal{C}^2(\mathbb{M})$ и скалярной кривизны $R(g, \Gamma)$, построенной по метрике $g_{\alpha\beta}$ и аффинной связности общего вида $\Gamma_{\alpha\beta}^\gamma$. Мы предполагаем, что компоненты метрики и связности являются достаточно гладкими функциями, и интеграл (23.13) сходится. Кроме этого предположим, что всеми граничными слагаемыми, возникающими при интегрировании по частям, можно пренебречь.

Подстановка в действие (23.13) римановой кривизны $\tilde{R}$, зависящей только от метрики, приводит к чрезвычайно трудоемкой вариационной задаче. Это связано с тем, что при дифференцировании по частям необходимо дифференцировать также и скалярное поле. Поскольку скалярная кривизна $\tilde{R}$ содержит вторые производные от метрики, то интегрировать по частям необходимо два раза, и это приводит к большому числу слагаемых. Значительное упрощение вносят последовательные действия. Сначала варьируем по метрике $g_{\alpha\beta}$ и связности $\Gamma_{\alpha\beta}^\gamma$, рассматривая их, как независимые переменные, а затем подставляем вариацию связности, выраженную через вариацию метрики, тензора кручения и неметричности. В общей теории относительности такой подход называется *формализмом первого порядка*.

Начнем с нескольких вспомогательных формул, необходимых в дальнейшем. Варьируя определение обратной метрики,

$$g^{\alpha\beta} g_{\beta\gamma} = \delta_\gamma^\alpha,$$

получаем тождество

$$\delta g^{\alpha\beta} g_{\beta\gamma} + g^{\alpha\beta} \delta g_{\beta\gamma} = 0.$$

Отсюда следует связь между вариацией самой метрики и ее обратной:

$$\delta g^{\alpha\beta} = -g^{\alpha\gamma} g^{\beta\delta} \delta g_{\gamma\delta}. \quad (23.14)$$

Из теории матриц известно, что для произвольной квадратной обратимой матрицы $A = (A_{\alpha\beta})$ справедливо тождество

$$\delta \det A = \det A A^{-1\alpha\beta} \delta A_{\alpha\beta}.$$

Отсюда следует, что вариация определителя метрики $g := \det g_{\alpha\beta}$ равна

$$\delta g = g g^{\alpha\beta} \delta g_{\alpha\beta} = -g g_{\alpha\beta} \delta g^{\alpha\beta}. \quad (23.15)$$

Эту вариацию мы записали в двух видах, т.к. в приложениях часто бывает удобнее варьировать действие не по самой метрике, а по ее обратной. Наличие квадратного корня в мере объема $\sqrt{|g|} := \sqrt{|\det g_{\alpha\beta}|}$ приводит к появлению множителя $1/2$. Поэтому для ее вариации справедливы равенства

$$\delta \sqrt{|g|} = \frac{1}{2} \sqrt{|g|} g^{\alpha\beta} \delta g_{\alpha\beta} = -\frac{1}{2} \sqrt{|g|} \delta g^{\alpha\beta} g_{\alpha\beta}. \quad (23.16)$$

Приступим к вариации действия (23.13). Вариационная производная по скалярному полю φ очевидна

$$S_{,\varphi} := \frac{\delta S}{\delta \varphi} = \sqrt{|g|} R. \quad (23.17)$$

Метрика входит в действие (23.13) дважды: в форму объема и в определение скалярной кривизны (6.90), причем без производных. Поэтому нетрудно проверить, что

$$S^{,\alpha\beta} := \frac{\delta S}{\delta g_{\alpha\beta}} = -\sqrt{|g|} \varphi \left(R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R \right). \quad (23.18)$$

Вариация действия по аффинной связности $\Gamma_{\alpha\beta}^\gamma$, все компоненты которой рассматриваются, как независимые переменные, более трудоемка. Это связано с тем, что приходится интегрировать по частям, т.к. тензор кривизны (6.74) зависит от производных аффинной связности. Прямые вычисления приводят к следующему выражению для вариации лагранжиана после интегрирования по частям

$$\begin{aligned} \delta L = & -\sqrt{|g|} \partial_\alpha \varphi g^{\alpha\gamma} \delta \Gamma_{\beta\gamma}^\beta + \sqrt{|g|} \partial_\gamma \varphi g^{\alpha\beta} \delta \Gamma_{\alpha\beta}^\gamma - \sqrt{|g|} \varphi \left(Q_\alpha^{\alpha\gamma} - \frac{1}{2} Q^\gamma \right) \delta \Gamma_{\beta\gamma}^\beta + \\ & + \sqrt{|g|} \varphi \left[- \left(T_{\delta\gamma}^\delta + \frac{1}{2} Q_\gamma \right) g^{\alpha\beta} - T_\gamma^{\beta\alpha} - T_\gamma^{\alpha\beta} + Q_\gamma^{\alpha\beta} \right] \delta \Gamma_{\alpha\beta}^\gamma, \end{aligned} \quad (23.19)$$

где $T_{\alpha\beta}^\gamma$ и $Q_{\alpha\beta\gamma}$ – тензоры кручения и неметричности. Для облегчения вычислений следует помнить, что выражение, стоящее перед вариацией связности, должно быть тензорным полем. Это поможет правильно сгруппировать слагаемые. Заметим также, что выражение, стоящее в квадратных скобках, симметрично по индексам α и β .

Теперь вычислим очень важную для приложений вариационную производную действия (23.13) по метрике $g_{\alpha\beta}$ в римановой геометрии. Обозначим соответствующее действие, зависящее только от скалярного поля и метрики, через

$$\tilde{S} = \int_{\mathbb{M}} dx \tilde{L} = \int_{\mathbb{M}} dx \sqrt{|g|} \varphi \tilde{R}. \quad (23.20)$$

Поскольку кручение и неметричность в римановой геометрии равны нулю, то из (23.18) и (23.19) следует, что вариация подынтегрального выражения в действии по метрике равна

$$\delta \tilde{L} = -\sqrt{|g|} \varphi \left(\tilde{R}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} \tilde{R} \right) \delta g_{\alpha\beta} - \sqrt{|g|} \partial_\gamma \varphi g^{\gamma\beta} \delta \tilde{\Gamma}_{\alpha\beta}^\alpha + \sqrt{|g|} \partial_\gamma \varphi g^{\alpha\beta} \delta \tilde{\Gamma}_{\alpha\beta}^\gamma.$$

Выразив вариацию связности $\delta \tilde{\Gamma}_{\alpha\beta}^\gamma$ через вариацию метрики, проинтегрировав по частям и приведя подобные члены, получим окончательное выражение для вариационной производной

$$\tilde{S}^{,\alpha\beta} := \frac{\delta \tilde{S}}{\delta g_{\alpha\beta}} = -\sqrt{|g|} \varphi \left(\tilde{R}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} \tilde{R} \right) + \sqrt{|g|} (\tilde{\square} \varphi g^{\alpha\beta} - \tilde{\nabla}^\alpha \tilde{\nabla}^\beta \varphi). \quad (23.21)$$

Напомним, что в римановой геометрии ввиду симметрии символов Кристоффеля по нижним индексам вторая ковариантная производная от скалярного поля симметрична: $\tilde{\nabla}^\alpha \tilde{\nabla}^\beta \varphi = \tilde{\nabla}^\beta \tilde{\nabla}^\alpha \varphi$ (сравните с равенством (6.95)).

Если скалярное поле равно единице, $\varphi = 1$, то действие (23.20) совпадает с действием Гильберта–Эйнштейна (23.7) без космологической постоянной, и мы получаем выражение для вариационной производной (23.8).

Формула для вариационной производной (23.21) со скалярным полем важна, например, при вычислении скобок Пуассона в канонической формулировке общей теории относительности.

Вариация действия (23.13) была проведена в пространстве произвольной размерности. Отметим, что в двумерном случае тензор Эйнштейна (23.3) тождественно равен нулю, и первое слагаемое в (23.21) отсутствует (см. раздел 28.1).

23.5 Зависимость уравнений Эйнштейна

В настоящем разделе мы считаем, что кручение и неметричность равны нулю, а связностью является связность Леви–Чивиты, построенная по заданной метрике.

Важным обстоятельством в общей теории относительности является линейная зависимость уравнений Эйнштейна (23.2). Предположим, что эти уравнения получены вариацией по метрике действия (23.10), которое инвариантно относительно общих преобразований координат. Если действие инвариантно относительно локальных преобразований, то согласно второй теореме Нетер между уравнениями движения существует линейная зависимость (20.63). Рассмотрим эту зависимость в случае общих преобразований координат. Для простоты предположим, что действие полей материи зависит только от некоторого конечного набора скалярных полей $\varphi^a(x)$, $a = 1, \dots, N$. При бесконечно малых преобразованиях координат $x^\alpha \mapsto x^\alpha + \epsilon^\alpha$ с параметром $\epsilon^\alpha(x)$ метрика и скалярные поля преобразуются по правилам (2.104) и (2.98):

$$\begin{aligned}\delta g_{\alpha\beta} &= -\tilde{\nabla}_\alpha \epsilon_\beta - \tilde{\nabla}_\beta \epsilon_\alpha, \\ \delta \varphi^a &= -\epsilon^\alpha \partial_\alpha \varphi^a,\end{aligned}$$

где $\epsilon_\alpha := g_{\alpha\beta} \epsilon^\beta$. Следовательно, инвариантность действия записывается в виде

$$\delta S = \int dx \left(\frac{\delta S}{\delta g_{\alpha\beta}} (-2\tilde{\nabla}_\beta \epsilon_\alpha) + \frac{\delta S}{\delta \varphi^a} (-\epsilon^\alpha \partial_\alpha \varphi^a) \right) = 0.$$

После интегрирования по частям первого слагаемого получаем искомую зависимость уравнений движения

$$2\sqrt{|g|}\tilde{\nabla}_\beta \left(\frac{\delta(S_{\text{HE}} + S_{\text{M}})}{\sqrt{|g|}\delta g_{\alpha\beta}} \right) - \partial_\alpha \varphi^a \frac{\delta S_{\text{M}}}{\delta \varphi^a} = 0,$$

т.к. действие Гильберта–Эйнштейна не зависит от полей материи. Это – тождества, которые выполняются независимо от того удовлетворяют поля уравнениям движения или нет. Поскольку каждое слагаемое в действии инвариантно само по себе, то выполняются два тождества:

$$\tilde{\nabla}_\beta G^{\beta\alpha} = 0, \quad (23.22)$$

$$\sqrt{|g|}\tilde{\nabla}_\beta T_{\text{M}\alpha}^\beta - \frac{\delta S_{\text{M}}}{\delta \varphi^a} \partial_\alpha \varphi^a = 0, \quad (23.23)$$

где мы воспользовались определением тензора энергии-импульса материи (23.11) в общей теории относительности. Первое из этих уравнений представляет собой свернутые тождества Бианки (6.121), а второе – ковариантный “закон сохранения” тензора энергии-импульса материи. Действительно, если выполнены уравнения для полей

материи,

$$\frac{\delta S_M}{\delta \varphi^a} = 0,$$

то дивергенция тензора энергии-импульса материи обращается в нуль

$$\tilde{\nabla}_\beta T_M^\beta{}_\alpha = 0. \quad (23.24)$$

Нетрудно видеть, что аналогичные выкладки можно проделать для любого набора полей материи. При этом второе слагаемое в (23.23) может усложниться, но оно всегда будет пропорционально уравнениям движения для полей материи. Единственное условие – это инвариантность действия. Таким образом, получаем следующее

Предложение 23.5.1. *Если действие полей материи инвариантно относительно общих преобразований координат и поля материи удовлетворяют своим уравнениям Эйлера–Лагранжа, то ковариантная дивергенция тензора энергии-импульса (23.24) равна нулю.*

На формулу (23.24) можно взглянуть с другой точки зрения. Допустим, что нам заданы уравнения Эйнштейна (23.2), а про инвариантное действие, приводящее к этим уравнениям, ничего не известно. Уравнения Эйнштейна – это система дифференциальных уравнений на метрику, и у них есть условия интегрируемости. Чтобы их получить возьмем ковариантную производную от обеих частей уравнений Эйнштейна. Дивергенция тензора Эйнштейна равна нулю (23.22) как следствие тождеств Бианки (6.121). Дивергенция метрики тоже равна нулю, т.к. связность Леви–Чивиты является метрической. Следовательно, ковариантный “закон сохранения” тензора энергии-импульса материи (23.24) является условием интегрируемости системы дифференциальных уравнений Эйнштейна для метрики (23.2). Это важно учитывать в тех случаях, когда тензор энергии-импульса материи получен не из принципа наименьшего действия, а из каких либо других соображений.

Пример 23.5.1. Если в качестве материи рассматривать жидкость или газ, для которой уравнения движения не следуют из принципа наименьшего действия, то условие (23.24) является независимым уравнением (см. далее раздел 23.12). $\square$

23.6 Действие для полей материи в обобщенных моделях гравитации

В настоящем разделе мы покажем, что при минимальной подстановке ковариантное обобщение канонического тензора энергии-импульса материи является источником в уравнении для репера e_α^a , а ковариантное обобщение спинового момента полей материи – источником в уравнении для лоренцевой связности $\omega_{\alpha a}^b$.

В различных обобщенных моделях гравитации мы обычно предполагаем, что инвариантное действие состоит из двух слагаемых

$$S = S_G + S_M, \quad (23.25)$$

где S_G – гравитационная часть действия и S_M – действие для полей материи, в которое для простоты мы включили также калибровочные поля (электромагнитное поле и поле Янга–Миллса). В общей теории относительности $S_G = S_{\text{не}}$ – это действие Гильберта–Эйнштейна (23.7), равное интегралу от скалярной кривизны с возможным

добавлением космологической постоянной. Это действие зависит только от метрики или репера. В более общих моделях гравитационная часть действия может включать также инварианты более высокого порядка по кривизне, кручению, тензору неметричности и их ковариантных производных. В таких случаях мы рассматриваем в качестве независимых переменных, по которым производится варьирование, переменные Картана: репер e_α^a и линейную $\mathbb{GL}(n, \mathbb{R})$ связность $\omega_{\alpha a}^b$ (см. раздел 5.4).

Если действие для полей материи может быть выражено через метрику и аффинную связность, то в качестве независимых переменных можно рассматривать также метрику, кручение и неметричность. Однако это не всегда имеет место. Например, для спинорного поля в геометрии Римана–Картана лоренцева связность не может быть выражена через метрику и аффинную связность. В этом случае введение репера необходимо. Таким образом, в общем случае использование переменных Картана предпочтительнее.

В настоящем разделе мы обсудим общие свойства уравнений движения для полей материи, не используя конкретный вид гравитационной части действия S_G .

Остановимся более подробно на действии S_M для полей материи. К полям материи в настоящем разделе мы относим скалярные, спинорные поля, электромагнитное поле, поля Янга–Миллса и все другие поля, кроме репера и линейной связности. Обозначим всю совокупность полей материи через $\varphi = \{\varphi^A\}$, $A = 1, 2, \dots$. Пусть действие для полей материи в плоском пространстве-времени Минковского $\mathbb{R}^{1,n-1}$ с координатами x^a , $a = 0, 1, \dots, n-1$, имеет вид

$$S_M = \int_{\mathbb{R}^{1,n-1}} dx L_M(\eta, \hat{\eta}, \varphi, \partial\varphi),$$

где лагранжиан полей материи L_M зависит от метрики Минковского $\eta = \{\eta_{ab}\}$, инвариантной метрики в пространстве-мишени для полей материи $\hat{\eta} = \{\eta_{AB}\}$ (во многих случаях это просто символ Кронекера, $\eta_{AB} = \delta_{AB}$), полей материи $\varphi = \{\varphi^A\}$ и их частных производных первого порядка $\partial\varphi = \{\partial_a\varphi^A\}$.

В моделях гравитации мы предполагаем, что пространство-время $\mathbb{M}$ является многообразием той же размерности, что и исходное пространство Минковского $\mathbb{R}^{1,n-1}$. Обозначим локальные координаты на $\mathbb{M}$ через x^α , $\alpha = 0, 1, \dots, n-1$. В простейшем случае инвариантное действие для полей материи получается из действия в пространстве Минковского с помощью *минимальной подстановки*:

$$\begin{aligned} \mathbb{R}^{1,n-1} &\rightarrow \mathbb{M}, & \dim \mathbb{M} &= n, \\ L_M &\mapsto \sqrt{|g|} L_M, \\ \eta_{ab} &\mapsto g_{\alpha\beta} = e_\alpha^a e_\beta^b g_{ab}, \\ \eta_{AB} &\mapsto g_{AB}(x), \\ \varphi^A &\mapsto \varphi^A, \\ \partial_a \varphi^A &\mapsto \nabla_\alpha \varphi^A, & \nabla_\alpha \varphi^A &:= \partial_\alpha \varphi^A + \omega_{\alpha B}^A \varphi^B, \end{aligned}$$

где $\omega_{\alpha A}^B$ – компоненты линейной связности в том представлении, в котором преобразуются поля материи.

В случае общей линейной группы преобразований $\mathbb{GL}(n, \mathbb{R})$ инвариантной метрики в пространстве-мишени не существует (если поля материи преобразуются по точному представлению $\mathbb{GL}(n, \mathbb{R})$), и мы вынуждены рассматривать нетривиальную метрику в пространстве-мишени $g_{AB}(x)$ как дополнительное поле. Метрика $g_{ab}(x)$ – это произвольная симметричная невырожденная матрица лоренцевой сигнатуры, с помощью которой определяется репер.

Пример 23.6.1. Рассмотрим набор массивных скалярных полей одинаковой массы, лагранжиан которых в пространстве Минковского имеет вид (см. раздел 25.1)

$$L_M = \frac{1}{2} \eta^{ab} \partial_a \varphi^A \partial_b \varphi^B \eta_{AB} - \frac{1}{2} m^2 \varphi^A \varphi^B \eta_{AB}. \quad (23.26)$$

Мы считаем, что скалярные поля преобразуются по некоторому представлению группы Лоренца (индекс A), и η_{AB} – инвариантная метрика. Данный лагранжиан инвариантен относительно глобальных преобразований Лоренца, действующих на координаты x^a и поля φ^A . При этом координаты и поля можно преобразовывать либо одновременно, либо по отдельности. После минимальной подстановки он примет вид

$$L_M = \sqrt{|g|} \left(\frac{1}{2} g^{\alpha\beta} \nabla_\alpha \varphi^A \nabla_\beta \varphi^B g_{AB} - \frac{1}{2} m^2 \varphi^A \varphi^B g_{AB} \right),$$

где ковариантная производная определяется линейной связностью (см. раздел 5.2)

$$\nabla_\alpha \varphi^A := \partial_\alpha \varphi^A + \omega_{\alpha B}^A \varphi^B, \quad \omega_{\alpha B}^A := \omega_\alpha^{ab} L_{abB}^A, \quad (23.27)$$

g_{AB} – метрика в пространстве-мишени и $g^{\alpha\beta} = e^\alpha_a e^\beta_b g^{ab}$. В приведенной формуле L_{abB}^A – представление генераторов линейной группы для выбранного набора скалярных полей. Соответствующее действие инвариантно относительно общих преобразований координат и локальных $\mathbb{GL}(n, \mathbb{R})$ вращений в пространстве-мишени:

$$\begin{aligned} e_\alpha^b &\mapsto e_\alpha^b S_b^{-1a}, \\ g_{ab} &\mapsto S_a^c S_b^d g_{cd}, \\ \omega_{\alpha a}^b &\mapsto S_a^c \omega_{\alpha c}^d S_d^{-1b} + \partial_\alpha S_a^c S_c^{-1b}, \\ \varphi^A &\mapsto \varphi^B S_B^{-1A}, \\ g_{AB} &\mapsto S_A^c S_B^d g_{cd}, \end{aligned}$$

где $S_a^b(x) \in \mathbb{GL}(n, \mathbb{R})$ – матрица локальных вращений и S_A^B – ее представление, которое соответствует выбранному набору скалярных полей. Если ограничить группу $\mathbb{GL}(n, \mathbb{R})$ до группы Лоренца $\mathbb{O}(1, n-1)$, то в качестве метрики $g_{ab}(x)$ следует выбрать инвариантную метрику Лоренца η_{ab} , которая уже не будет зависеть от точки многообразия. Ей соответствует некоторая инвариантная метрика η_{AB} в пространстве-мишени. $\square$

Ограничимся моделями, основанными на геометрии Римана–Картана. Тогда группа Ли общих линейных преобразований $\mathbb{GL}(n, \mathbb{R})$ сужается до группы Лоренца, и мы полагаем $g_{ab} = \eta_{ab}$ и $g_{AB} = \eta_{AB}$.

Из общего выражения для действия (23.25) следуют уравнения Эйлера–Лагранжа для геометрических переменных e_α^a , $\omega_{\alpha a}^b$ и полей материи φ^A :

$$\frac{\delta S}{\delta e_\alpha^a} = 0, \quad (23.28)$$

$$\frac{\delta S}{\delta \omega_{\alpha a}^b} = 0, \quad (23.29)$$

$$\frac{\delta S}{\delta \varphi^A} = \frac{\delta S_M}{\delta \varphi^A} = 0. \quad (23.30)$$

Поскольку при минимальной подстановке репер входит в действие полей материи только в качестве общего множителя $\sqrt{|g|} := \det e_\alpha^a$ и в ковариантную производную

с латинским индексом $\nabla_a \varphi^A := e^\alpha_a \nabla_\alpha \varphi^A$, то первое уравнение можно переписать в виде

$$\frac{\delta S_G}{\delta e^\alpha_a} = \sqrt{|g|} \frac{\partial L_M}{\partial (\nabla_\alpha \varphi^A)} e^\beta_a \nabla_\beta \varphi^A - \sqrt{|g|} e^\alpha_a L_M = \sqrt{|g|} e^\beta_a T_{M\beta}{}^\alpha, \quad (23.31)$$

где

$$T_{M\beta}{}^\alpha := -\frac{1}{\sqrt{|g|}} \frac{\delta S_M}{\delta e^\alpha_a} e^\beta_a = \nabla_\beta \varphi^A \frac{\partial L_M}{\partial (\nabla_\alpha \varphi^A)} - \delta^\alpha_\beta L_M.$$

Сравнивая последнее выражение с каноническим тензором энергии-импульса (20.43), полученным из теоремы Нетер, мы видим, что в правой части гравитационных уравнений (23.31), полученных после варьирования по реперу, стоит его ковариантное обобщение. Поэтому говорят, что тензор энергии-импульса материи является источником для репера.

Линейная связность $\omega_\alpha{}^{ab}$ входит в действие полей материи только через ковариантную производную. Поэтому уравнение (23.29) можно переписать в виде

$$\frac{\delta S_G}{\delta \omega_\alpha{}^{ab}} = -\sqrt{|g|} \frac{\partial L_M}{\partial (\nabla_\alpha \varphi^A)} L_{ab\mathbf{B}}{}^A \varphi^{\mathbf{B}} = \sqrt{|g|} S_{Mab}{}^\alpha,$$

где

$$S_{Mab}{}^\alpha := -\frac{\partial L_M}{\partial (\nabla_\alpha \varphi^A)} L_{ab\mathbf{B}}{}^A \varphi^{\mathbf{B}}. \quad (23.32)$$

В геометрии Римана–Картана общая линейная группа $\mathbb{GL}(n, \mathbb{R})$ сужается до группы Лоренца $\mathbb{O}(1, n-1)$, линейная связность $\omega_\alpha{}^{ab}$ становится лоренцевой связностью, $L_{ab\mathbf{B}}{}^A$ – генераторы группы Лоренца, $g_{ab} = \eta_{ab}$, $g_{\mathbf{A}\mathbf{B}}$ – инвариантная метрика в пространстве-мишени, которая всегда существует, т.к. группа Лоренца проста. В этом случае выражение (23.32) является ковариантным обобщением спинowego момента полей материи (20.53). Поэтому мы говорим, что спиновый момент полей материи является источником для лоренцевой связности.

Таким образом, выбор переменных Картана для моделей гравитации позволил дать физическую интерпретацию источников гравитационного поля: тензор энергии-импульса материи является источником для репера, а спиновый момент – источником для лоренцевой связности. При выборе в качестве независимых переменных в геометрии Римана–Картана метрики и кручения, в правой части соответствующих уравнений будут стоять выражения, которые не имеют столь простой и привлекательной интерпретации.

23.7 Скалярно-тензорные модели

В качестве одного из возможных обобщений эйнштейновской теории гравитации, основанного на римановой геометрии, рассматривают скалярно-тензорные модели, в которых гравитационное взаимодействие описывается метрикой $g_{\alpha\beta}$ и скалярным полем ϕ . Лагранжиан скалярно-тензорных моделей обычно записывают в виде

$$L = \kappa \sqrt{|g|} \phi R + \sqrt{|g|} \frac{\omega(\phi)}{\phi} \partial \phi^2 - \sqrt{|g|} V(\phi) + L_M, \quad (23.33)$$

где $R(g)$ – скалярная кривизна (в настоящем разделе знаки тильды опущены),

$$\partial \phi^2 := g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi$$

– квадрат градиента скалярного поля, κ – гравитационная постоянная, $\omega(\phi)$ и $V(\phi)$ – некоторые достаточно гладкие функции от скалярного поля, характеризующие неминимальность взаимодействия и потенциал самодействия. Все остальные поля включены в лагранжиан полей материи L_m .

С точки зрения общей теории относительности можно сказать следующее. Лагранжиан (23.33) описывает гравитационное взаимодействие полей материи, причем гравитационная “постоянная” $\kappa\phi(x)$ зависит от точки пространства-времени, т.е. является скалярным полем, которое удовлетворяет своим уравнениям движения. При этом связь этого скалярного поля с метрикой неминимальна и включается некоторое самодействие, описываемое потенциалом V .

При $n = 2$ лагранжиан вида (23.33) описывает двумерную дилатонную гравитацию общего вида, а скалярное поле ϕ называется полем дилатона. Этот класс моделей характеризуется двумя произвольными функциями $U(\phi) := 2\omega(\phi)/\phi$ и $V(\phi)$ и приводит к интегрируемым уравнениям движения. Эти модели будут рассмотрены в главе 32.

Впервые действие вида (23.33) было рассмотрено М. Фирцем в 1956 году [157]. Скалярно-тензорные модели гравитации привлекли значительное внимание после работ П. Йордана [158] и С. Бранса и Р. Дике [159]. Основная идея этих исследований восходит к работе П. Дирака, который предположил, что гравитационная постоянная может меняться со временем [160]. М. Фирц пошел дальше, выдвинув гипотезу о том, что гравитационная постоянная описывается независимым скалярным полем, удовлетворяющим некоторому нелинейному уравнению движения.

Получим уравнения движения для скалярно-тензорных моделей (23.33). Используя вид вариационной производной (23.21), нетрудно получить вариационные производные действия (23.33) по обратной метрике и скалярному полю

$$\begin{aligned} \frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta g^{\alpha\beta}} : \quad & \kappa\phi \left(R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R \right) - \kappa(\Box\phi g_{\alpha\beta} - \nabla_\alpha \nabla_\beta \phi) + \\ & + \frac{\omega}{\phi} \left(\partial_\alpha \phi \partial_\beta \phi - \frac{1}{2}g_{\alpha\beta} \partial\phi^2 \right) + \frac{1}{2}g_{\alpha\beta}V + \frac{1}{2}T_{m\alpha\beta} = 0. \end{aligned} \quad (23.34)$$

$$\frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta \phi} : \quad \kappa R - 2\frac{\omega}{\phi}\Box\phi - \frac{\omega'}{\phi}\partial\phi^2 + \frac{\omega}{\phi^2}\partial\phi^2 - V' = 0, \quad (23.35)$$

где $T_{m\alpha\beta}$ – тензор энергии-импульса материи (23.11), а штрих обозначает дифференцирование функции по аргументу. След уравнения (23.34) имеет вид

$$\phi\left(\frac{n}{2} - 1\right) \left(\kappa R + \frac{\omega}{\phi^2} \partial\phi^2 \right) + \kappa(n-1)\Box\phi - \frac{n}{2}V - T_m = 0,$$

где $T_m := T_{m\alpha}{}^\alpha$ – след тензора энергии-импульса материи. Это уравнение при $n \neq 2$ позволяет исключить сумму

$$\kappa R + \frac{\omega}{\phi^2} \partial\phi^2$$

из уравнения для скалярного поля (23.35). В результате получим эквивалентное уравнение

$$2 \left(\frac{n-1}{n-2} \kappa + \omega \right) \Box\phi + \frac{1}{n-2} (nV + 2T_m) - \omega' \partial\phi^2 - \phi V' = 0. \quad (23.36)$$

Таким образом систему уравнений движения (23.34), (23.35) можно переписать в эквивалентном виде, заменив уравнение для скалярного поля (23.35) на уравнение

(23.36). В четырехмерном пространстве-времени уравнения движения обычно записывают в виде

$$\kappa\phi\left(R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R\right) = -\frac{1}{2}T_{\text{м}\alpha\beta} + \kappa(\Box\phi g_{\alpha\beta} - \nabla_\alpha\nabla_\beta\phi) - \\ - \frac{\omega}{\phi}\left(\partial_\alpha\phi\partial_\beta\phi - \frac{1}{2}g_{\alpha\beta}\partial\phi^2\right) - \frac{1}{2}g_{\alpha\beta}V, \quad (23.37)$$

$$\Box\phi = \frac{1}{2\omega + 3\kappa}(2V + T_{\text{м}} - \omega'\partial\phi^2 - \phi V'). \quad (23.38)$$

С математической точки зрения уравнения движения скалярно-тензорных моделей значительно сложнее уравнений общей теории относительности и изучены недостаточно полно. Поэтому практически ничего нельзя сказать о виде функций $\omega(\phi)$ и $V(\phi)$, которые приводят к удовлетворительным результатам с теоретической точки зрения и не противоречат существующим наблюдательным данным.

23.8 Полиномиальная форма действия Гильберта–Эйнштейна

Рассмотрим пространство-время $\mathbb{M}$ произвольной размерности $\dim \mathbb{M} = n \geq 3$. Обозначим локальные координаты пространства-времени через x^α , $\alpha = 0, 1, \dots, n-1$. Уравнения движения для метрики $g_{\alpha\beta}(x)$ в общей теории относительности без полей материи следуют из вариационного принципа для действия Гильберта–Эйнштейна (23.7).

Действие Гильберта–Эйнштейна неполиномиально по компонентам метрики по двум причинам. Во-первых, оно содержит неполиномиальный элемент объема $\sqrt{|g|}$. Во-вторых, выражение для скалярной кривизны содержит обратную метрику $g^{\alpha\beta}$, компоненты которой также неполиномиальны по $g_{\alpha\beta}$. Поэтому действие Гильберта–Эйнштейна в теории возмущений представляет собой очень сложный бесконечный ряд, что является существенной технической трудностью при анализе уравнений движения и квантовании. По этим же причинам действие Гильберта–Эйнштейна неполиномиально по компонентам обратной метрики.

Покажем, что конфигурационное пространство общей теории относительности можно расширить, включив определитель метрики в качестве дополнительной независимой переменной таким образом, что действие примет полиномиальный вид. Эквивалентность полиномиального действия исходному действию достигается за счет наложения в расширенном конфигурационном пространстве связи, которая также полиномиальна по полям. Изложение настоящего раздела следует статьям [161, 162].

Координатами конфигурационного пространства $\mathcal{M}$ общей теории относительности являются компоненты метрики $g_{\alpha\beta}(x)$. Размерность этого пространства равна

$$\dim \mathcal{M} = \frac{n(n+1)}{2} \times \infty^{(n-1)},$$

где символический множитель $\infty^{(n-1)}$ соответствует точкам пространства $((n-1)$ -мерным пространственным подмногообразиям) в пространстве-времени $\mathbb{M}$.

Рассмотрим другое конфигурационное пространство $\mathcal{N}$ с координатами $\varrho(x)$, $k_{\alpha\beta}(x)$. При этом будем считать, что $\varrho > 0$ и матрица $k_{\alpha\beta}$ симметрична и невырождена в каждой точке пространства-времени:

$$k_{\alpha\beta} = k_{\beta\alpha}, \quad \det k_{\alpha\beta} \neq 0.$$

Мы предполагаем также, что матрицы $g_{\alpha\beta}$ и $k_{\alpha\beta}$ имеют одинаковую лоренцеву сигнатуру

$$\text{sign } g_{\alpha\beta} = \text{sign } k_{\alpha\beta} = (+ \underbrace{- \dots -}_{n-1}).$$

Размерность нового конфигурационного пространства равна

$$\dim \mathcal{N} = \left(\frac{n(n+1)}{2} + 1 \right) \times \infty^{(n-1)}.$$

Выделим в $\mathcal{N}$ подпространство $\mathcal{M}'$ с помощью дополнительного условия

$$\det k_{\alpha\beta} = \begin{cases} 1, & \det g_{\alpha\beta} > 0, \quad \text{нечетные } n, \\ -1, & \det g_{\alpha\beta} < 0, \quad \text{четные } n. \end{cases} \quad (23.39)$$

Тогда между точками подпространства $\mathcal{M}' \subset \mathcal{N}$ и исходного конфигурационного пространства $\mathcal{M}$ можно установить взаимно однозначное соответствие

$$g_{\alpha\beta} := \varrho^2 k_{\alpha\beta}. \quad (23.40)$$

Обратное преобразование имеет вид

$$\varrho = |\det g_{\alpha\beta}|^{1/2n}, \quad k_{\alpha\beta} = \varrho^{-2} g_{\alpha\beta}. \quad (23.41)$$

Поэтому мы отождествим $\mathcal{M} = \mathcal{M}'$. Представление для обратной метрики $g^{\alpha\beta} = \varrho^{-2} k^{\alpha\beta}$, где $k^{\alpha\beta} k_{\beta\gamma} = \delta_\gamma^\alpha$ следует из уравнения (23.40).

Сделаем два важных замечания. Во-первых, компоненты обратной матрицы $k^{\alpha\beta}$ полиномиальны по $k_{\alpha\beta}$. В общем случае из условия (23.39) следует, что они являются полиномами степени $n-1$ по компонентам $k_{\alpha\beta}$. Во-вторых, компоненты $k_{\alpha\beta}$ являются компонентами не тензора, а тензорной плотности второго ранга. Действительно, потребуем, чтобы равенство (23.40) было выполнено в произвольной системе координат. Поскольку определитель метрики g представляет собой скалярную плотность степени $\deg g = -2$, то матрица $k_{\alpha\beta}$ является симметричной тензорной плотностью второго ранга и степени $\deg k_{\alpha\beta} = 2/n$, а поле ϱ – скалярной плотностью степени $\deg \varrho = -1/n$. То есть при преобразовании координат $x^\alpha \rightarrow x^{\alpha'}(x)$ новые переменные преобразуются по правилу

$$k_{\alpha'\beta'} = \partial_\alpha x^\alpha \partial_{\beta'} x^\beta k_{\alpha\beta} J^{2/n}, \quad \varrho' = \varrho J^{-1/n},$$

где $J := \det \partial_\alpha x^{\alpha'}$ – якобиан преобразования координат. Это значит, что дополнительное условие (23.39) инвариантно относительно преобразования координат. В связи с этим будем называть поле $k_{\alpha\beta}(x)$ *плотностью метрики*.

Явная формула для компонент обратной плотности метрики имеет вид

$$k^{\alpha\beta} = \frac{1}{(n-1)!} \hat{\varepsilon}^{\alpha\gamma_2 \dots \gamma_n} \hat{\varepsilon}^{\beta\delta_2 \dots \delta_n} k_{\gamma_2 \delta_2} \dots k_{\gamma_n \delta_n},$$

где $\hat{\varepsilon}^{\alpha_1 \dots \alpha_n}$ – полностью антисимметричная тензорная плотность степени -1 с единичными компонентами $|\hat{\varepsilon}^{\alpha_1 \dots \alpha_n}| = 1$. Данная формула показывает, что компоненты $k^{\alpha\beta}$ – это полиномы степени $n-1$ от компонент $k_{\alpha\beta}$ и наоборот.

Перепишем действие Гильберта–Эйнштейна в новых переменных $\varrho, k_{\alpha\beta}$. Для вычислений нам понадобятся следующие тождества, которые следуют после дифференцирования равенства (23.39) два раза

$$\begin{aligned} k^{\gamma\delta} \partial_\alpha k_{\gamma\delta} &= 0, \\ k^{\gamma\delta} \partial_{\alpha\beta}^2 k_{\gamma\delta} - k^{\gamma\delta} k^{\epsilon\zeta} \partial_\alpha k_{\gamma\epsilon} \partial_\beta k_{\delta\zeta} &= 0. \end{aligned}$$

Прямые вычисления приводят к равенствам

$$\begin{aligned} \partial_\alpha g_{\beta\gamma} &= \varrho^2 (\partial_\alpha k_{\beta\gamma} + 2\partial_\alpha \ln \varrho k_{\beta\gamma}), \\ \partial_{\alpha\beta}^2 g_{\gamma\delta} &= \varrho^2 (\partial_{\alpha\beta}^2 k_{\gamma\delta} + 2\partial_{\alpha\beta}^2 \ln \varrho k_{\gamma\delta} + \\ &\quad + 2\partial_\alpha \ln \varrho \partial_\beta k_{\gamma\delta} + 2\partial_\beta \ln \varrho \partial_\alpha k_{\gamma\delta} + 4\partial_\alpha \ln \varrho \partial_\beta \ln \varrho k_{\gamma\delta}), \\ g^{\beta\delta} \partial_\alpha g_{\delta\gamma} &= k^{\beta\delta} \partial_\alpha k_{\delta\gamma} + 2\partial_\alpha \ln \varrho \delta_\gamma^\beta, \\ g^{\epsilon\zeta} \partial_\alpha g_{\gamma\epsilon} \partial_\beta g_{\delta\zeta} &= \varrho^2 (k^{\epsilon\zeta} \partial_\alpha k_{\beta\epsilon} \partial_\beta k_{\delta\zeta} + 2\partial_\alpha \ln \varrho \partial_\beta k_{\gamma\delta} + 2\partial_\beta \ln \varrho \partial_\alpha k_{\gamma\delta} + \\ &\quad + 4\partial_\alpha \ln \varrho \partial_\beta \ln \varrho k_{\gamma\delta}), \\ g^{\gamma\delta} g^{\epsilon\zeta} \partial_\alpha g_{\gamma\epsilon} \partial_\beta g_{\delta\zeta} &= k^{\gamma\delta} k^{\epsilon\zeta} \partial_\alpha k_{\gamma\epsilon} \partial_\beta k_{\delta\zeta} + 4n \partial_\alpha \ln \varrho \partial_\beta \ln \varrho, \\ g^{\gamma\delta} \partial_{\alpha\beta}^2 g_{\gamma\delta} &= k^{\gamma\delta} \partial_{\alpha\beta}^2 k_{\gamma\delta} + 2n \partial_{\alpha\beta}^2 \ln \varrho + 4n \partial_\alpha \ln \varrho \partial_\beta \ln \varrho, \\ \partial_{\alpha\beta}^2 g_{\gamma\delta} - \frac{1}{2} g^{\epsilon\zeta} \partial_\alpha g_{\gamma\epsilon} \partial_\beta g_{\delta\zeta} - \frac{1}{2} g^{\epsilon\zeta} \partial_\beta g_{\gamma\epsilon} \partial_\alpha g_{\delta\zeta} &= \\ &= \varrho^2 \left(\partial_{\alpha\beta}^2 k_{\gamma\delta} - \frac{1}{2} k^{\epsilon\zeta} \partial_\alpha k_{\gamma\epsilon} \partial_\beta k_{\delta\zeta} - \frac{1}{2} k^{\epsilon\zeta} \partial_\beta k_{\gamma\epsilon} \partial_\alpha k_{\delta\zeta} + 2\partial_{\alpha\beta}^2 \ln \varrho k_{\gamma\delta} \right). \end{aligned} \quad (23.42)$$

Теперь нетрудно найти выражение для скалярной кривизны

$$R = \varrho^{-4} [\varrho^2 R^{(k)} + 2(n-1)\varrho \partial_\alpha (k^{\alpha\beta} \partial_\beta \varrho) + (n-1)(n-4)\partial \varrho^2], \quad (23.43)$$

где

$$\partial \varrho^2 := k^{\alpha\beta} \partial_\alpha \varrho \partial_\beta \varrho,$$

и “скалярная” кривизна $R^{(k)}$ для плотности метрики $k_{\alpha\beta}$ принимает удивительно простой вид

$$R^{(k)} = \partial_{\alpha\beta} k^{\alpha\beta} + \frac{1}{2} k^{\alpha\beta} \partial_\alpha k^{\gamma\delta} \partial_\gamma k_{\beta\delta} - \frac{1}{4} k^{\alpha\beta} \partial_\alpha k^{\gamma\delta} \partial_\beta k_{\gamma\delta}. \quad (23.44)$$

Отметим, что это выражение полиномиально и по плотности метрики $k_{\alpha\beta}$, и по ее обратной $k^{\alpha\beta}$.

Заметим, что для заданной плотности метрики $k_{\alpha\beta}$ мы можем формально построить символы Кристоффеля, тензоры кривизны и Риччи, а также скалярную кривизну. Соответствующие “символы Кристоффеля” не определяют связность на $\mathbb{M}$, а “кривизна” не является тензором, потому что новая переменная $k_{\alpha\beta}$ является не тензором, а тензорной плотностью. Например, скаляром в выражении (23.43) является не просто $R^{(k)}$, а вся правая часть выражения. Вместе с этим, группа общих преобразований координат содержит подгруппу, состоящую из преобразований с единичным якобианом. Относительно этой подгруппы “символы Кристоффеля” $\Gamma_{\alpha\beta}^{(k)\gamma}$ для $k_{\alpha\beta}$ преобразуются, как компоненты связности, а “кривизна” $R_{\alpha\beta\gamma}^{(k)\delta}$ является тензором.

Вторые производные $\partial_{\alpha\beta}^2 k^{\alpha\beta}$ и $\partial_\alpha (k^{\alpha\beta} \partial_\beta \varrho)$, как и в общей теории относительности, можно исключить из действия Гильберта–Эйнштейна (23.7), вычтя граничный член

$$\partial_\alpha (\varrho^{n-2} \partial_\beta k^{\alpha\beta} + 2(n-1)\varrho^{n-3} k^{\alpha\beta} \partial_\beta \varrho).$$

Здесь, для краткости, мы положили $\kappa = 0$. В результате действие примет вид

$$S_{\text{HE}} \stackrel{\text{div}}{=} \int dx L_{\text{HE}}, \quad (23.45)$$

где

$$L_{\text{HE}} = \varrho^{n-4} \left[\frac{1}{2} \varrho^2 k^{\alpha\beta} \partial_\alpha k^{\gamma\delta} \partial_\gamma k_{\beta\delta} - \frac{1}{4} k^{\alpha\beta} \partial_\alpha k^{\gamma\delta} \partial_\beta k_{\gamma\delta} - \right. \\ \left. - (n-2) \varrho \partial_\alpha k^{\alpha\beta} \partial_\beta \varrho - (n-1)(n-2) \partial \varrho^2 - (n-2) \Lambda \varrho^4 \right]. \quad (23.46)$$

При размерности пространства-времени $n \geq 4$ этот лагранжиан полиномиален по полям $\varrho, k_{\alpha\beta}$ со степенями n и $n+1$, соответственно. Общая степень полинома равна $2n-1$. По построению, с точностью до граничных слагаемых это действие инвариантно относительно общих преобразований координат. Необходимо только помнить, что поля ϱ и $k_{\alpha\beta}$ являются не тензорами, а тензорными плотностями.

Лагранжиан (23.46) имеет вид лагранжиана дилатонной гравитации, где роль дилатона играет определитель метрики. От обычных моделей он отличается тем, что содержит перекрестный член с производными $\partial_\alpha k^{\alpha\beta} \partial_\beta \varrho$. Кроме того, он содержит меньшее число независимых полей, т.к. на плотность метрики наложено условие (23.39).

Дополнительное условие на плотность метрики (23.39) можно учесть, добавив к лагранжиану связь,

$$L_{\text{HE}} \mapsto L_{\text{HE}} + \lambda(\det k_{\alpha\beta} \pm 1),$$

где $\lambda(x)$ – множитель Лагранжа.

Введение множителей Лагранжа необязательно. Из условия (23.39) следует условие на вариации плотности метрики

$$\delta \det k_{\alpha\beta} = \pm k^{\alpha\beta} \delta k_{\alpha\beta} = \mp \delta k^{\alpha\beta} k_{\alpha\beta} = 0. \quad (23.47)$$

Поэтому действие (23.45) можно варьировать по $k_{\alpha\beta}$ или $k^{\alpha\beta}$, рассматривая все компоненты как независимые, а затем взять бесследовую часть получившихся уравнений.

Уравнения Эйлера–Лагранжа для действия (23.45) эквивалентны вакуумным уравнениям Эйнштейна (23.5). Вариация действия по ϱ приводит к уравнению

$$2(n-1) \varrho \square \varrho + (n-1)(n-4) \partial \varrho^2 + R^{(k)} \varrho^2 = n \Lambda \varrho^4, \quad (23.48)$$

которое эквивалентно следу уравнений Эйнштейна (23.5). Вариация действия (23.45) по плотности метрики дает $n(n+1)/2 - 1$ уравнений:

$$\varrho^2 \left(R_{\alpha\beta}^{(k)} - \frac{1}{n} k_{\alpha\beta} R^{(k)} \right) + \\ + (n-2) (\varrho \nabla_\alpha \nabla_\beta \varrho - \nabla_\alpha \varrho \nabla_\beta \varrho) - \frac{n-2}{n} k_{\alpha\beta} (\varrho \square \varrho - \partial \varrho^2) = 0, \quad (23.49)$$

где

$$\begin{aligned} \nabla_\alpha \varrho &= \partial_\alpha \varrho, \\ \nabla_\alpha \nabla_\beta \varrho &= \partial_{\alpha\beta}^2 \varrho - \Gamma_{\alpha\beta}^{(k)\gamma} \partial_\gamma \varrho, \\ \square \varrho &= \partial_\alpha (k^{\alpha\beta} \partial_\beta \varrho), \end{aligned} \quad (23.50)$$

и символы Кристоффеля $\Gamma_{\alpha\beta}^{(k)\gamma}$ построены по плотности метрики $k_{\alpha\beta}$. Эти уравнения эквивалентны бесследовой части уравнений Эйнштейна, потому что при варьировании необходимо учесть условие на вариации компонент плотности метрики (23.47). Заметим, что “ковариантные” производные тензоров и тензорных плотностей при условии (23.39) совпадают, т.к. $\Gamma_{\alpha\beta}^{(k)\beta} = 0$. “Тензор Риччи” в (23.49) имеет вид

$$R_{\alpha\beta}^{(k)} = \frac{1}{2}k^{\gamma\delta}(\partial_{\gamma\delta}^2 k_{\alpha\beta} - \partial_{\alpha\gamma}^2 k_{\beta\delta} - \partial_{\beta\gamma}^2 k_{\alpha\delta}) - \frac{1}{2}\partial_{\gamma}k^{\gamma\delta}(\partial_{\alpha}k_{\beta\delta} + \partial_{\beta}k_{\alpha\delta} - \partial_{\delta}k_{\alpha\beta}) - \\ - \frac{1}{4}\partial_{\alpha}k_{\gamma\delta}\partial_{\beta}k^{\gamma\delta} - \frac{1}{2}k^{\gamma\delta}k^{\epsilon\zeta}(\partial_{\gamma}k_{\alpha\epsilon}\partial_{\delta}k_{\beta\zeta} - \partial_{\gamma}k_{\alpha\zeta}\partial_{\delta}k_{\beta\epsilon}).$$

Уравнения (23.48) и (23.49) представляют собой, соответственно, след и бесследовую часть уравнений Эйнштейна (23.5). Для любого решения уравнений (23.48) и (23.49) существует единственная метрика (23.40), которая удовлетворяет уравнениям Эйнштейна (23.5). Наоборот, для любого решения уравнений Эйнштейна (23.5) можно построить единственные плотность метрики и поле ϱ (23.41), которые будут удовлетворять уравнениям (23.48) и (23.49).

Таким образом, при $n \geq 4$, действие (23.45), уравнения Эйлера–Лагранжа (23.48), (23.49) и связь (23.39) полиномиальны по полям $\varrho, k_{\alpha\beta}$. Это существенное упрощение теории достигнуто за счет расширения конфигурационного пространства путем введения дополнительной полевой переменной ϱ . Если связь (23.39) явно решить относительно одной из компонент плотности метрики $k_{\alpha\beta}$ и подставить найденное решение в действие (23.45), то полиномиальность теории будет утеряна. Отметим, что введение нового поля и связи не является чем-то, из ряда вон выходящим: исходная метрика $g_{\alpha\beta}$ и так содержит нефизические степени свободы, которые исключаются из теории путем решения калибровочных условий и связей, содержащихся в общей теории относительности. Мы лишь увеличили число переменных и связей, оставив физические степени свободы прежними.

На гамильтоновом языке проведенная процедура означает следующее. Фазовое пространство, соответствующее переменным $\varrho, k_{\alpha\beta}$, также расширится, при этом возникнет дополнительная связь на импульсы (равенство нулю следа импульсов, соответствующих $k_{\alpha\beta}$), которая вместе со связью (23.39) образует пару связей второго рода. Однако полное фазовое пространство уже будет не симплектическим, а только пуассоновым, т.к. матрица скобок Пуассона координат этого пространства будет вырождена. Подробнее этот вопрос рассмотрен далее в разделе 24.7.

Добавление скалярного $\varphi(x)$ и электромагнитного $A_{\alpha}(x)$ полей сохраняет действие и уравнения Эйлера–Лагранжа полиномиальными. При минимальной подстановке получаем следующие лагранжианы:

$$L_s = \frac{1}{2}\sqrt{g}[g^{\alpha\beta}\partial_{\alpha}\varphi\partial_{\beta}\varphi - V(\varphi)] = \frac{1}{2}\varrho^{n-2}[k^{\alpha\beta}\partial_{\alpha}\varphi\partial_{\beta}\varphi - \varrho^2V(\varphi)], \\ L_{\text{ем}} = -\frac{1}{4}\sqrt{g}g^{\alpha\beta}g^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta} = -\frac{1}{4}\varrho^{n-4}k^{\alpha\beta}k^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta},$$

где $V(\varphi)$ – потенциал скалярного поля, включая массовый член, и $F_{\alpha\beta} := \partial_{\alpha}A_{\beta} - \partial_{\beta}A_{\alpha}$ – напряженность электромагнитного поля.

В четырехмерном пространстве-времени уравнение (23.48) упрощается

$$\square\varrho + \frac{1}{6}R^{(k)}\varrho = \frac{2}{3}\Lambda\varrho^3. \quad (23.51)$$

При этом мы разделили его на ϱ , что возможно в силу предположения $\varrho > 0$.

Отбросим на время условие на плотность метрики (23.51) и будем считать $k_{\alpha\beta}$ метрикой, а ϱ – скалярным полем. Тогда уравнение (23.51) ковариантно относительно конформных преобразований

$$\bar{k}_{\alpha\beta} = \Omega^2 k_{\alpha\beta}, \quad \bar{\varrho} = \Omega^{-1} \varrho, \quad (23.52)$$

где $\Omega(x) > 0$ – дважды дифференцируемая функция. Оно рассматривалось в [163, 164, 165] при $\Lambda = 0$ и в [166] при $\Lambda \neq 0$. Рассматриваемый подход существенно отличается, поскольку на плотность метрики $k_{\alpha\beta}$ наложено дополнительно условие (23.39), которое явно нарушает конформную инвариантность. Однако появление множителя $1/6$ в уравнении (23.51) не случайно, т.к. параметризация (23.40) по виду совпадает с конформным преобразованием метрики (23.52).

Унимодулярная гравитация

При создании общей теории относительности А. Эйнштейн для упрощения вычислений часто использовал систему координат, в которой форма объема $\sqrt{|g|}$ является постоянной. Не ограничивая общности, можно считать, что эта постоянная равна единице:

$$|g| := |\det g_{\alpha\beta}| = 1. \quad (23.53)$$

Дифференцируя это равенство по координатам, получим уравнения:

$$\partial_\alpha \sqrt{|g|} = 0 \quad \Leftrightarrow \quad \partial_\alpha g = 0.$$

Используя формулу для дифференцирования определителя метрики (37.10), полученные условия можно также записать в другом эквивалентном виде

$$g^{\beta\gamma} \partial_\alpha g_{\beta\gamma} = 0 \quad \Leftrightarrow \quad \tilde{\Gamma}_\alpha := \tilde{\Gamma}_{\beta\alpha}{}^\beta = 0.$$

Дивергенция произвольного векторного поля в унимодулярной гравитации имеет тот же вид, что и в плоском пространстве в декартовых координатах:

$$\nabla_\alpha X^\alpha = \partial_\alpha X^\alpha.$$

В общековариантных теориях, которой является общая теория относительности, существует произвол в выборе системы координат, которым можно воспользоваться. Нетрудно доказать, что в окрестности произвольной точки пространства-времени существует такая система координат, что определитель метрики равен по модулю единице (23.53). Такие системы координат определены с точностью до преобразований координат с единичным якобианом. Условие $|\det g_{\alpha\beta}| = 1$ существенно упрощает многие формулы, в частности, выражение для скалярной кривизны (23.44) становится полиномиальным.

Модели гравитации для метрики с единичным определителем рассматривались в физике неоднократно. Этим обстоятельством пользовался Эйнштейн при создании общей теории относительности [156]. В настоящее время такие модели часто называют *унимодулярной гравитацией*. Если рассматривается только гравитация без взаимодействия с другими полями, то унимодулярная гравитация есть просто общая теория относительности в специальной фиксированной системе координат, которая упрощает вид многих формул. Поэтому она эквивалентна общей теории относительности. Однако при введении взаимодействия с полями материи возникают дополнительные возможности, которые мы обсуждать не будем.

23.9 Точечные частицы в теории гравитации

Пусть задано пространство-время, т.е. многообразие $\mathbb{M}$ с метрикой g лоренцевой сигнатуры. Будем считать, что кручение и неметричность на $\mathbb{M}$ равны нулю, и, для простоты, не будем помечать это обстоятельство знаком тильды. Размерность пространства-времени пока не фиксируем, $\dim \mathbb{M} = n \geq 2$. Точечная частица движется в пространстве-времени $\mathbb{M}$ по некоторой дифференцируемой времениподобной кривой $\{q^\alpha(\tau)\} \in \mathbb{M}$, где $\tau \in \mathbb{R}$ – произвольный параметр вдоль этой кривой. Напомним, что кривая называется времениподобной, если вектор скорости кривой, $u^\alpha := \dot{q}^\alpha := dq^\alpha/d\tau$, времениподобен, $u^2 := u^\alpha u^\beta g_{\alpha\beta} > 0$. Мы считаем, что частица движется в будущее, т.е. $u^0 > 0$. Форма кривой определяется рассматриваемой задачей и силами, которые действуют на частицу. В общем случае параметр вдоль кривой произволен, и его выбирают из соображений удобства. Наиболее часто в качестве параметра вдоль траектории частицы выбирают ее длину s . Это всегда возможно, т.к. обыкновенное дифференциальное уравнение $ds = \sqrt{\dot{q}^\alpha \dot{q}^\beta g_{\alpha\beta}} d\tau$ разрешимо относительно $s = s(\tau)$. В дальнейшем точка, как правило, будет обозначать дифференцирование по параметру s .

Определение. Времениподобная дифференцируемая кривая $\{q^\alpha(\tau)\} \in \mathbb{M}$, вдоль которой движется точечная частица, называется *траекторией* или *мировой линией* частицы. Если параметр вдоль траектории частицы совпадает с ее длиной, $\tau = s$, то он называется *собственным временем*. Векторное поле

$$u^\alpha := \frac{dq^\alpha}{ds}, \quad (23.54)$$

определенное на траектории частицы, называется *собственной скоростью* частицы. Ковариантная производная от скорости частицы вдоль ее траектории

$$w^\alpha := \frac{dq^\beta}{ds} \nabla_\beta u^\alpha = u^\beta \nabla_\beta u^\alpha \quad (23.55)$$

называется *ускорением* частицы. Частица называется *свободной*, если ее ускорение равно нулю. Производная вдоль траектории частицы

$$v^\alpha := \frac{dq^\alpha}{dq^0} = \frac{\dot{q}^\alpha}{\dot{q}^0} \quad (23.56)$$

называется *наблюдаемой скоростью* частицы в системе координат x^α . $\square$

Собственная скорость и ускорение частицы являются n -мерными векторами, определенным вдоль траектории частицы. Собственное время – это то время, которое показывают часы наблюдателя, движущегося вместе с частицей. Когда наблюдатель движется вместе с частицей, то он может измерить свою скорость относительно системы координат x^α , это и будут компоненты собственной скорости. Наблюдаемая скорость, как следует из определения, не является векторным полем и зависит от выбора системы координат. Это та скорость, которую измеряет внешний наблюдатель в выбранной системе отсчета.

Равенство нулю ускорения частицы

$$u^\beta \nabla_\beta u^\alpha = 0,$$

определяет экстремали (16.25). Это значит, что свободные частицы в теории тяготения движутся вдоль экстремалей пространства-времени. Если на частицу действуют негравитационные силы, например, электромагнитные, то в уравнении движения $mw^\alpha = f^\alpha$, где $m = \text{const} > 0$ – масса частицы, появится внешняя сила с компонентами f^α . В этом случае ее траектория будет отличаться от экстремали.

Предложение 23.9.1. *Если параметр вдоль мировой линии частицы канонический, то квадрат собственной скорости равен единице,*

$$u^2 := u^\alpha u^\beta g_{\alpha\beta} = 1. \quad (23.57)$$

При этом ускорение всегда ортогонально скорости

$$u^\alpha w_\alpha = 0.$$

Доказательство. Первое утверждение следует из определения:

$$u^2 = \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} g_{\alpha\beta} = \frac{ds^2}{ds^2} = 1.$$

Продифференцируем это равенство вдоль траектории

$$u^\beta \nabla_\beta (u^2) = 2u^\beta u^\alpha \nabla_\beta u_\alpha = 2u^\alpha w_\alpha = 0,$$

где мы воспользовались тем, что ковариантная производная от метрики для связности Леви–Чивиты равна нулю. Отсюда вытекает второе утверждение предложения. $\square$

Наблюдаемая скорость является нековариантным объектом. Из определения следует соотношение между компонентами собственной и наблюдаемой скоростями

$$v^\alpha = \frac{u^\alpha}{\dot{q}^0}. \quad (23.58)$$

Более подробно,

$$v^0 = 1, \quad v^\mu = \frac{dq^\mu}{dq^0}, \quad \mu = 1, \dots, n-1.$$

Возведем равенство $v^\alpha \dot{q}^0 = u^\alpha$ в квадрат и учтем, что $u^2 = 1$. Тогда получим, что компоненты собственной скорости частицы можно записать в виде

$$u^0 := \dot{q}^0 = \frac{1}{\sqrt{v^2}}, \quad u^\mu := \dot{q}^\mu = \frac{v^\mu}{\sqrt{v^2}}, \quad (23.59)$$

где $v^2 = v^\alpha v^\beta g_{\alpha\beta}$ – квадрат наблюдаемой скорости. Из равенства (23.58) также следует, что $v^2 > 0$. Понятие наблюдаемой скорости частицы будет использовано в дальнейшем при определении ультрарелятивистского предела для точечной частицы.

Замечание. Определения даны для точечной частицы, находящейся под действием произвольных сил, не только гравитационных. Мы также предполагаем, что, если наблюдатель находится в определенной точке пространства в определенный момент времени, то ему известны значения всех координат x^α , соответствующих данной точке пространства-времени. Мы также предполагаем, что ему известны координаты всех близлежащих точек, что необходимо для вычисления производных. Эти упрощающие предположения сделаны для того, чтобы обойти важный и интересный, но сложный вопрос измерений. $\square$

В общем случае, когда на частицу действуют произвольные силы, она может двигаться по любой времениподобной кривой. В моделях гравитации мы предполагаем, что точечная частица, на которую действуют только гравитационные силы, описывается инвариантным действием

$$S_m = -mc \int_p^q ds = -mc \int_{\tau_1}^{\tau_2} d\tau \sqrt{\dot{q}^\alpha \dot{q}^\beta g_{\alpha\beta}}, \quad (23.60)$$

где $m = \text{const} > 0$ – масса частицы, c – скорость света и интегрирование проводится вдоль времениподобной кривой $q(\tau)$, $\tau \in \mathbb{R}$, соединяющей точки $p := q(\tau_1)$ и $q := q(\tau_2)$. Действие (23.60) отличается от длины мировой линии частицы (16.20) постоянным множителем $-mc$ и, если метрика задана, варьируется только по траектории частицы $\delta q^\alpha(\tau)$. В дальнейшем положим $c = 1$.

Допустим, что точки $p, q \in \mathbb{M}$ можно соединить времениподобной экстремалью и притом только одной. Ясно, что эти точки всегда можно соединить также ломаными светоподобными кривыми γ . Длина этих кривых равна нулю, потому что каждый сегмент ломаных светоподобен. Если некоторая времениподобная кривая аппроксимирует одну из этих кривых γ , то ее длина будет близка к нулю. Остальные времениподобные кривые, соединяющие точки p и q будут иметь некоторую положительную длину. Поэтому, если точки p и q соединяются только одной экстремалью, то это будет времениподобная кривая наибольшей длины. Благодаря знаку минус перед действием (23.60), экстремаль соответствует минимуму функционала (23.60) среди всех времениподобных кривых, соединяющих точки p и q . Отсюда следует, что точечные частицы движутся вдоль экстремалей пространства-времени, если на них действуют только внешние гравитационные силы.

Если метрика пространства-времени считается заданной, то это означает, что мы не учитываем гравитационное поле, создаваемое самой частицей. Во многих случаях такая постановка задачи вполне допустима, например, если масса частицы мала. Такие частицы называются *пробными*.

Замечание. Мы предполагаем, что точечная частица описывается действием (23.60) независимо от того, задана ли на $\mathbb{M}$ аффинная геометрия общего вида с кручением и неметричностью или задана только метрика. Тогда в общем случае точечная частица будет двигаться по экстремали, а не по геодезической. Если кручение и неметричность пространства-времени равны нулю, то экстремали совпадают с геодезическими. Поэтому в общей теории относительности можно также сказать, что точечные частицы под действием гравитационных сил движутся по геодезическим. Если, помимо гравитационных сил, присутствуют также другие, например, электромагнитные взаимодействия, то вид действия (23.60) изменится. Поэтому при наличии сил негравитационного происхождения траектории частиц, вообще говоря, отличаются от экстремалей. $\square$

Выше было введено понятие собственного времени для произвольной времениподобной линии. Для экстремали собственное время по существу совпадает с каноническим параметром. Действительно, для любой экстремали квадрат вектора скорости сохраняется, $u^2 = \text{const}$, (16.29). Поскольку канонический параметр определен с точностью до умножения на отличную от нуля постоянную, то его всегда можно выбрать таким образом, чтобы квадрат вектора скорости частицы был равен единице (23.57). Другими словами, в качестве канонического параметра выбирается длина экстремали, $\tau = s$.

Предположим, что в пространстве-времени находится N частиц с массами m_i , $i = 1, \dots, N$, которые взаимодействуют между собой только посредством гравитационных сил. В общей теории относительности суммарное действие гравитационного поля и совокупности точечных частиц равно сумме действия Гильберта–Эйнштейна (23.7) и действий для каждой частицы,

$$S = S_{\text{HE}} + \sum_{i=1}^N S_i = \kappa \int_{\mathbb{M}} dx \sqrt{|g|} R - \sum_i m_i \int_{\tau_{i1}}^{\tau_{i2}} d\tau_i \sqrt{\dot{q}_i^\alpha \dot{q}_i^\beta g_{\alpha\beta}}, \quad (23.61)$$

где мы, для простоты, опустили космологическую постоянную и знак тильды у скалярной кривизны. Во втором слагаемом метрика рассматривается как сложная функция $g_{\alpha\beta}(\tau_i) = g_{\alpha\beta}(q(\tau_i))$, и параметры τ_i могут быть выбраны произвольно для каждой частицы. Первый интеграл берется по всему пространству-времени, а последующие – в пределах $\tau_{i1,2}$ (возможно, бесконечных), которые соответствуют пересечению мировых линий частиц с краем пространства-времени $\partial\mathbb{M}$, если таковой имеется. Для простоты, предположим, что мировые линии частиц нигде не пересекаются, т.е. частицы не сталкиваются между собой.

В настоящем разделе нас не будут интересовать граничные эффекты. Поэтому пределы интегрирования, для простоты, мы в дальнейшем опустим.

Действие (23.61) инвариантно относительно общих преобразований координат и независимой перепараметризации параметров τ_i вдоль каждой траектории. В дальнейшем мы предполагаем, что вдоль каждой траектории параметр совпадает с собственным временем. Более того, поскольку в действие входит сумма интегралов вдоль траекторий, то индекс i у параметров τ_i можно опустить,

$$S = \int dx \sqrt{|g|} \kappa R - \int d\tau \sum_i m_i \sqrt{\dot{q}_i^\alpha \dot{q}_i^\beta g_{\alpha\beta}}. \quad (23.62)$$

В общем случае пределы интегрирования для различных частиц могут отличаться. Однако, поскольку нас не интересуют граничные эффекты, мы этого указывать не будем.

В разделе 16.2 действие для экстремалей было проварьировано в предположении, что вдоль нее выбран канонический параметр. Сейчас мы получим уравнения без этого предположения, т.к. в ряде случаев удобнее выбирать параметр вдоль экстремалей, исходя из других соображений. Рассмотрим одну точечную частицу с мировой линией $q^\alpha(\tau)$ в произвольной параметризации. Простые вычисления приводят к следующим уравнениям движения

$$S_{m,\alpha} := \frac{\delta S_m}{\delta q^\alpha} = m \frac{g_{\alpha\beta}}{u^2} \left(\ddot{q}^\beta + \Gamma_{\gamma\delta}^\beta \dot{q}^\gamma \dot{q}^\delta - \frac{\dot{q}^\beta}{2u^2} \frac{d(u^2)}{d\tau} \right) = 0, \quad (23.63)$$

где

$$u^2 := \dot{q}^\alpha \dot{q}^\beta g_{\alpha\beta}.$$

Поскольку исходное действие инвариантно относительно произвольной замены параметра вдоль мировой линии, то согласно второй теореме Нетер между уравнениями движения существует линейная зависимость. Чтобы ее найти, рассмотрим бесконечно малое изменение параметра

$$\tau \mapsto \tau + \epsilon(\tau).$$

Соответствующая вариация формы функций $q^\alpha(\tau)$ (см. раздел 2.13) имеет вид

$$\delta q^\alpha = -\epsilon \dot{q}^\alpha.$$

Следовательно, вариация действия равна

$$\delta S_m = - \int d\tau S_{m,\alpha} \epsilon \dot{q}^\alpha.$$

Поскольку функция $\epsilon(\tau)$ произвольна, то из инвариантности действия, $\delta S_m = 0$, вытекает зависимость уравнений движения:

$$S_{m,\alpha} \dot{q}^\alpha = 0. \quad (23.64)$$

В этом тождестве можно убедиться прямой проверкой.

При произвольной параметризации мировой линии частицы квадрат вектора скорости не является постоянным, $u^2 \neq \text{const}$. Если выбран канонический параметр вдоль экстремали, то $d(u^2)/d\tau = 0$, и последнее слагаемое в уравнении (23.63) обращается в нуль.

Действие для совокупности частиц (23.62) инвариантно относительно независимой перепараметризации каждой мировой линии. Поэтому вдоль каждой кривой можно выбрать свой канонический параметр. В дальнейшем мы это предположим.

Для вывода полной системы уравнений движения, действие (23.62) необходимо проварьировать по метрике $g_{\alpha\beta}(x)$ и траекториям частиц $q_i^\alpha(\tau)$. Вариация действия для частиц по компонентам метрики $\delta g_{\alpha\beta}(x)$ не определена, т.к. оно записано только вдоль траекторий. Поэтому мы преобразуем интегралы вдоль траекторий в интегралы по всему пространству-времени. Для этого вставим в подынтегральное выражение единицу,

$$1 = \int dx \delta(x - q_i) := \int dx \delta(x^0 - q_i^0) \delta(x^1 - q_i^1) \dots \delta(x^{n-1} - q_i^{n-1}),$$

и изменим порядок интегрирования, предположив, что это возможно. Тогда действие примет вид

$$S = \int dx \left[\kappa \sqrt{|g|} R - \int d\tau \sum_i m_i \sqrt{\dot{q}_i^\alpha \dot{q}_i^\beta g_{\alpha\beta}} \delta(x - q_i) \right]. \quad (23.65)$$

Теперь метрику во втором слагаемом можно рассматривать, как функцию от точки пространства-времени, $g_{\alpha\beta} = g_{\alpha\beta}(x)$. Вариационные производные этого действия по траекториям частиц и метрике при канонической параметризации равны

$$\frac{\delta S}{\delta q_i^\alpha} = m_i (\ddot{q}_i^\beta + \Gamma_{\gamma\delta}{}^\beta \dot{q}_i^\gamma \dot{q}_i^\delta) g_{\beta\alpha}, \quad (23.66)$$

$$\frac{\delta S}{\delta g_{\alpha\beta}} = -\sqrt{|g|} \kappa \left(R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R \right) - \int d\tau \sum_i \frac{m_i}{2\sqrt{\dot{q}_i^\gamma \dot{q}_i^\delta g_{\gamma\delta}}} \dot{q}_i^\alpha \dot{q}_i^\beta \delta(x - q_i). \quad (23.67)$$

Вариационная производная по траекториям частиц (23.65) была получена ранее в произвольной параметризации (16.22). Вариационная производная действия Гильберта–Эйнштейна также была получена ранее в разделе (23.4). Для любого решения уравнения (23.66) канонический параметр τ можно выбрать так, что $\sqrt{\dot{q}_i^\gamma \dot{q}_i^\delta g_{\gamma\delta}} = 1$ для каждой частицы. Поэтому, не ограничивая общности, знаменатель во втором слагаемом (23.67) можно упростить, отбросив квадратный корень. Таким образом, связанная система уравнений движения гравитационного поля и системы точечных частиц примет вид

$$\kappa \sqrt{|g|} \left(R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R \right) = -\frac{1}{2} \sqrt{|g|} T_m^{\alpha\beta}, \quad (23.68)$$

$$g_{\alpha\beta} (\ddot{q}_i^\beta + \Gamma_{\gamma\delta}{}^\beta \dot{q}_i^\gamma \dot{q}_i^\delta) = 0, \quad (23.69)$$

где

$$T_m^{\alpha\beta} := \frac{1}{\sqrt{|g|}} \int d\tau \sum_i m_i \dot{q}_i^\alpha \dot{q}_i^\beta \delta(x - q_i) \quad (23.70)$$

– тензор энергии-импульса точечных частиц и $\dot{q}_{i\alpha} := \dot{q}_i^\beta g_{\beta\alpha}$. Интегрирование по каноническим параметру τ в тензоре энергии-импульса можно снять, используя одну δ -функцию, а именно $\delta(x^0 - q_i^0(\tau))$. Поскольку $\dot{q}_i^0 > 0$ (все частицы движутся в будущее), то для тензора энергии-импульса точечных частиц получаем следующее выражение

$$T_m^{\alpha\beta} = \frac{1}{\sqrt{|g|}} \sum_i \frac{m_i \dot{q}_i^\alpha \dot{q}_i^\beta}{\dot{q}_i^0} \delta(x - q_i), \quad (23.71)$$

где

$$\delta(x - q_i) := \delta(x^1 - q_i^1) \dots \delta(x^{n-1} - q_i^{n-1})$$

– пространственная δ -функция и параметр τ является неявной функцией x^0 , заданной уравнением $x^0 = q^0(\tau)$.

Замечание. Появление множителя $1/\sqrt{|g|}$ в выражении для тензора энергии-импульса точечных частиц не случайно. Напомним, что δ -функция является не функцией на многообразии, а скалярной плотностью степени -1 , как и элемент объема $\sqrt{|g|}$. Если параметризация мировой линии противоположна, $\dot{q}_i^0 < 0$, что соответствует античастице, то в знаменателе тензора энергии-импульса (23.71) необходимо поставить знак модуля: $|\dot{q}_i^0|$, в соответствии с правилом замены переменных интегрирования. $\square$

Таким образом, для точечных частиц, на которые действуют только гравитационные силы, мы имеем связанную систему уравнений (23.68), (23.69). Каждая частица движется по экстремали пространства-времени в соответствии с уравнением (23.69), где метрика определяется уравнениями Эйнштейна (23.68). В свою очередь, метрика зависит от распределения частиц, т.к. в правой части уравнений Эйнштейна стоит нетривиальный тензор энергии-импульса.

Обсудим трудности, которые возникают при решении уравнений (23.68) и (23.69) в связи с наличием δ -функций.

Уравнение для экстремалей (траекторий точечных частиц) (23.69) хорошо определены, если компоненты метрики – дифференцируемые функции, $g_{\alpha\beta} \in C^1(\mathbb{M})$. Однако детальный анализ системы уравнений движения показывает, что это условие не выполняется, т.к. компоненты метрики имеют особенности на мировых линиях частиц. Именно по этой причине мы не опустили индексы в уравнениях Эйнштейна (23.68). В рассматриваемом случае уравнения Эйнштейна с контравариантными и ковариантными индексами не эквивалентны. По этой же причине мы не сократили элемент объема $\sqrt{|g|}$, т.к. он обращается в нуль на мировых линиях частиц. В уравнениях для экстремалей (23.69) мы также не произвели каких либо манипуляций с индексами.

Наличие δ -функций в полной системе уравнений приводит к серьезным математическим трудностям. Поскольку есть δ -функции, то решения системы уравнений Эйнштейна нужно понимать в обобщенном смысле после интегрирования с пробными функциями. Если в качестве пробных функций выбрать пространство гладких функций с финитными носителями $\mathcal{D}(\mathbb{R}^3)$ (см., например, [3]), то компоненты метрики должны лежать в сопряженном пространстве $\mathcal{D}'(\mathbb{R}^3)$. Однако уравнения Эйнштейна нелинейны, а умножение в $\mathcal{D}'(\mathbb{R}^3)$ в общем случае определить нельзя.

Насколько известно автору, решение этой проблемы в настоящее время отсутствует. Поэтому вычисления данного раздела следует рассматривать, как ориентир, с которым следует сравнивать более строгие выкладки.

В заключение раздела проведем еще одно обобщение. Действие (23.62) описывает совокупность точечных частиц, взаимодействующих с гравитационным полем. Если частицы находятся дополнительно под действием некоторых потенциальных сил негравитационного происхождения, то действие можно обобщить, вставив соответствующий потенциал $V(x) \in C^1(\mathbb{M})$:

$$S = \int dx \sqrt{|g|} \kappa R - \int d\tau \sum_I m_I V \sqrt{\dot{q}_I^\alpha \dot{q}_I^\beta g_{\alpha\beta}}, \quad (23.72)$$

где потенциал рассматривается на мировых линиях, $V(\tau) := V(g_I(\tau))$. При этом мы не нарушаем инвариантность действия относительно независимого преобразования параметров мировых линий частиц. Если вдоль каждой мировой линии выбрать канонический параметр, то уравнения движения для действия (23.72) примут вид

$$\kappa \sqrt{|g|} \left(R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R \right) = -\frac{1}{2} \sqrt{|g|} T_m^{\alpha\beta} V, \quad (23.73)$$

$$V g_{\alpha\beta} (\ddot{q}_I^\beta + \Gamma_{\gamma\delta}^\beta \dot{q}_I^\gamma \dot{q}_I^\delta) + \dot{q}_I^\gamma \partial_\gamma V g_{\alpha\beta} \dot{q}_I^\beta - \partial_\alpha V = 0. \quad (23.74)$$

При $V \neq 0$ второе уравнение можно разделить на V и переписать в следующем виде

$$g_{\alpha\beta} (\ddot{q}_I^\beta + \Gamma_{\gamma\delta}^\beta \dot{q}_I^\gamma \dot{q}_I^\delta) = \Pi^\tau_\alpha{}^\beta \partial_\beta \ln|V|, \quad (23.75)$$

где

$$\Pi^\tau_\alpha{}^\beta := \delta_\alpha^\beta - \dot{q}_{I\alpha} \dot{q}_I^\beta$$

– проекционный оператор на направление, перпендикулярное к мировой линии I -той частицы. То есть потенциальные силы всегда действуют в направлении, перпендикулярном траекториям.

23.9.1 Нерелятивистский предел для точечной частицы

В настоящем разделе будет показана связь между уравнениями движения для точечных частиц (23.69) и хорошо знакомыми уравнениями движения частиц под действием гравитационного поля в механике Ньютона.

Пусть метрика пространства Минковского $\mathbb{R}^{1,n-1}$ имеет вид

$$\eta_{\alpha\beta} = \text{diag}(1, \underbrace{-1, \dots, -1}_{n-1}). \quad (23.76)$$

Для простоты рассмотрим движение одной частицы. В пространстве-времени $\mathbb{M}$ с нетривиальной метрикой $g_{\alpha\beta}(x)$ функции $\{q^\alpha(\tau)\}$ задают траекторию точечной частицы. Пусть $\tau = t$ – канонический параметр (собственное время). Предположим, что координата x^0 на $\mathbb{M}$ выбрана таким образом, что на траектории частицы $c\tau = ct = x^0$. Здесь мы ввели явно скорость света c для того, чтобы в дальнейшем строить разложение по малому параметру $\mathbf{u}^2/c^2 \ll 1$, где $\mathbf{u}$ – пространственная часть собственной скорости частицы. Условимся нумеровать, как обычно, пространственные координаты буквами из середины греческого алфавита:

$$\{x^\alpha\} = \{x^0, x^\mu\}, \quad \mu = 1, \dots, n-1.$$

Поскольку исходное действие (23.60) инвариантно относительно общих преобразований координат, то у нас имеется возможность дополнительно фиксировать $n - 1$ компонент метрики. Положим $g_{0\mu} = 0$. Тогда метрика примет блочно диагональный вид

$$g_{\alpha\beta} = \begin{pmatrix} g_{00} & 0 \\ 0 & g_{\mu\nu} \end{pmatrix} \quad (23.77)$$

Другими словами, система координат выбрана таким образом, чтобы времениподобный вектор ∂_0 был ортогонален всем касательным векторам к пространственным сечениям $x^0 = \text{const}$.

Введем два параметра разложения. Во-первых, нерелятивистский предел соответствует скоростям, малым по сравнению со скоростью света,

$$\frac{\mathbf{u}^2}{c^2} \sim \epsilon \ll 1, \quad \mathbf{u}^2 := -\eta_{\mu\nu} \dot{q}^\mu \dot{q}^\nu = \delta_{\mu\nu} u^\mu u^\nu \geq 0.$$

Во-вторых, слабому гравитационному полю соответствует метрика, которая мало отличается от метрики Минковского:

$$g_{00} = 1 + h_{00}, \quad g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad h_{00}, h_{\mu\nu} \sim \epsilon \ll 1.$$

Если гравитационное поле мало, $g_{\alpha\beta} \rightarrow \eta_{\alpha\beta}$, то в нерелятивистском пределе $\mathbf{u}^2 \rightarrow 0$ временная компонента собственной скорости стремится к скорости света, $\dot{q}^0 \rightarrow c$, т.к. выполнено соотношение (23.57). Если гравитационное поле не очень велико, $0 < g_{00} < \text{const}$, то в нерелятивистском пределе $g_{00}(\dot{q}^0)^2 \rightarrow c^2$ и производная $\dot{q}^0$ ограничена снизу некоторой положительной постоянной. Поскольку пространственные компоненты собственной u^μ и наблюдаемой v^μ скорости частицы отличаются на строго положительный множитель $\dot{q}^0$, то пределы $\mathbf{u}^2 \rightarrow 0$ и $\mathbf{v}^2 \rightarrow 0$, где $\mathbf{v}^2 := -\eta_{\mu\nu} v^\mu v^\nu \rightarrow 0$, эквивалентны.

Нерелятивистской частице в слабом гравитационном поле соответствует интервал

$$ds^2 = (c^2 + c^2 h_{00} + \eta_{\mu\nu} u^\mu u^\nu + h_{\mu\nu} u^\mu u^\nu) dt^2.$$

В силу сделанных предположений о малости гравитационного поля и скоростей последним слагаемым в этом представлении можно пренебречь. Тогда в нерелятивистском пределе с учетом первой поправки интервал для точечной частицы примет вид

$$ds^2 \approx \left(c^2 + \frac{2U}{m} - \mathbf{u}^2 \right) dt^2, \quad (23.78)$$

где введено обозначение

$$h_{00} = \frac{2U}{mc^2}. \quad (23.79)$$

Подставим приближенное выражение для интервала (23.78) в действие для точечной частицы (23.60), умноженное на скорость света, и разложим по степеням ϵ . Тогда в первом порядке по ϵ получим приближенное выражение

$$S_m \approx -mc \int d\tau \sqrt{c^2 + \frac{2U}{m} - \mathbf{u}^2} \approx \int d\tau \left(-mc^2 + \frac{m\mathbf{u}^2}{2} - U \right). \quad (23.80)$$

С точностью до энергии покоя точечной частицы с обратным знаком $-mc^2$ подынтегральное выражение совпадает с хорошо известным выражением для лагранжиана точечной частицы в нерелятивистской механике (21.17). Тем самым мы показали, что в нерелятивистском пределе поправку к временной компоненте метрики

(23.79), умноженной на mc^2 , следует интерпретировать, как потенциальную энергию $U := mc^2 h_{00}/2$ точечной частицы, находящейся во внешнем гравитационном поле.

Отметим, что разумный нерелятивистский предел обуславливает выбор общего знака минус в исходном действии для точечной частицы (23.60).

Интервалу (23.78) соответствует метрика

$$g_{\alpha\beta} = \text{diag}(1 + 2U/(mc^2), -1, \dots, -1). \quad (23.81)$$

Предположим, что потенциал U не зависит от времени t , т.е. движение частицы происходит в статическом гравитационном поле. Символы Кристоффеля в первом порядке малости имеют только три нетривиальные компоненты:

$$\Gamma_{00}^\mu = -\frac{\eta^{\mu\nu}\partial_\nu U}{mc^2}, \quad \Gamma_{0\mu}^0 = \Gamma_{\mu 0}^0 = \frac{\partial_\mu U}{mc^2}.$$

Соответствующие уравнения для экстремалей (16.21) принимают вид

$$\ddot{q}^0 = -2\frac{\partial_\mu U}{mc^2}\dot{q}^\mu\dot{q}^0, \quad (23.82)$$

$$\ddot{q}^\mu = -\frac{\delta^{\mu\nu}\partial_\nu U}{mc^2}\dot{q}^0\dot{q}^0. \quad (23.83)$$

Поскольку в выбранной системе координат $\ddot{q}^0 = 0$ и $\dot{U} = \dot{q}^\mu\partial_\mu U = 0$ (поле статично), то первое уравнение удовлетворяется автоматически. Второе уравнение совпадает с *вторым законом Ньютона* для движения точечной частицы в статическом потенциальном поле

$$m\ddot{q}^\mu = -\delta^{\mu\nu}\partial_\nu U, \quad (23.84)$$

где с рассматриваемой точностью сделана замена $\dot{q}^0 = c$. Таким образом, потерянное уравнение выполняется с точностью ϵ^2 . Заметим, что уравнение (23.82) не следует из действия (23.80), т.к. переменная q^0 в нем отсутствует.

Тот факт, что уравнение для q^0 удовлетворяется с рассматриваемой степенью точности не является удивительным. Действительно, среди n исходных уравнений для точечной частицы имеется одна линейная зависимость (23.64). Поэтому только $n - 1$ уравнений являются независимыми, которые в нерелятивистском пределе сводятся к уравнениям Ньютона (23.84).

В общей теории относительности ($n = 4$) метрика вдали от точечной массы M дается решением Шварцшильда (34.1)

$$ds^2 = \left(1 - \frac{2GM}{rc^2}\right) dt^2 - \frac{dr^2}{1 - \frac{2GM}{rc^2}} - r^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

где G – гравитационная постоянная. Соответствующее нерелятивистское выражение для потенциальной энергии пробной частицы имеет вид

$$U = -G\frac{mM}{r}, \quad (23.85)$$

что совпадает с *законом всемирного тяготения*. Таким образом, мы показали, что закон всемирного тяготения вытекает из общей теории относительности в нерелятивистском пределе.

В конце предыдущего раздела было рассмотрено действие для точечных частиц, (23.72), взаимодействующих не только с гравитационным полем, но и с другими потенциальными полями. Нерелятивистский предел в этом случае определяется так

же, как и раньше. Дополнительно мы требуем, чтобы потенциал V мало отличался от единицы:

$$V = 1 + \frac{W}{mc^2}, \quad \frac{W}{mc^2} \sim \epsilon \ll 1.$$

Тогда в нерелятивистском пределе приближенное выражение действия для точечной частицы (23.80) примет вид

$$S_m \approx \int d\tau \left(-mc^2 + \frac{m\mathbf{u}^2}{2} - U - W \right).$$

В этом случае движение пробной частицы определяется суммой гравитационного потенциала U , который возник из g_{00} компоненты метрики, и негравитационного потенциала W , который возник из множителя V .

23.9.2 Теория гравитации Ньютона

В настоящем разделе мы опишем гравитационное взаимодействие точечных частиц в механике Ньютона. При этом, по возможности, мы будем следовать общей схеме, принятой в общей теории относительности.

Пусть $\mathbb{M}$ – это тривиальное четырехмерное многообразие, $\mathbb{M} \approx \mathbb{R}^4$, с декартовой системой координат $x = \{x^\alpha\} = \{x^0, x^\mu\} = \{t, \mathbf{x}\}$, где $\alpha = 0, 1, 2, 3$ и $\mu = 1, 2, 3$. Как и раньше мы отождествим нулевую координату с временем, $x^0 = t$. Поставим нерелятивистскую задачу. Пусть в $\mathbb{M}$ движутся N точечных частиц по траекториям $q_i(t) = \{q_i^\mu(t)\}$, $i = 1, \dots, N$. В механике Ньютона время имеет абсолютный характер в том смысле, что оно одинаково для всех частиц и играет роль параметра вдоль траектории каждой частицы. Предположим, что между частицами действуют только гравитационные силы. Это значит, что каждая частица движется в гравитационном поле, которое создается другими частицами. В свою очередь каждая частица создает собственное гравитационное поле, которое влияет на движение других частиц. Обозначим суммарный потенциал гравитационного поля через $\varphi(x)$, который, по определению, является функцией (скалярным полем) на $\mathbb{R}^4$.

Действие для точечных частиц, взаимодействующих посредством гравитационного поля, является суммой действия для гравитационного поля

$$S_g = \frac{1}{4\pi G} \int_{\mathbb{R}^4} dx \frac{1}{2} \eta^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi, \quad (23.86)$$

где G – гравитационная постоянная и $\eta_{\mu\nu} = \text{diag}(-, -, -, -)$ – отрицательно определенная пространственная метрика, и действия для точечных частиц

$$S_m = - \sum_{i=1}^N \int_{-\infty}^{\infty} dt \frac{1}{2} m_i \dot{q}_i^\mu \dot{q}_i^\nu \eta_{\mu\nu} - \sum_{i=1}^N \int_{\mathbb{R}^4} dx m_i \varphi \delta(\mathbf{x} - \mathbf{q}_i), \quad (23.87)$$

где второе слагаемое описывает взаимодействие точечной частицы с гравитационным полем. Действие для гравитационного поля (23.86) отрицательно определено и равно потенциальной энергии гравитационного поля, взятой с обратным знаком. Действие для точечных частиц (23.87) как обычно представляет собой разность кинетической и потенциальной энергии. Отметим, что потенциальная энергия взаимодействия точечных частиц с гравитационным полем содержит только трехмерную дельта-функцию.

Варьирование суммарного действия, $S = S_g + S_m$, по гравитационному полю и траекториям частиц дает уравнения движения:

$$\frac{\delta S}{\delta \varphi} = \frac{1}{4\pi G} \Delta \varphi - \sum_I m_I \delta(\mathbf{x} - \mathbf{q}_I) = 0, \quad (23.88)$$

$$\frac{\delta S}{\delta q_I^\mu} = m_I (\ddot{q}_{I\mu} - \partial_\mu \varphi) = 0, \quad (23.89)$$

где $\Delta := \partial_1^2 + \partial_2^2 + \partial_3^2$ – лапласиан, $\ddot{q}_{I\mu} := \ddot{q}_I^\nu \eta_{\mu\nu}$ и градиент потенциала гравитационного поля $\partial_\mu \varphi$ во втором уравнении берется в той точке $q_I = \{t, q_I^\mu\} \in \mathbb{M}$, где в данный момент времени расположена частица.

Уравнение для гравитационного поля (23.88) представляет собой уравнение Пуассона

$$\Delta \varphi = 4\pi G \sum_I m_I \delta(\mathbf{x} - \mathbf{q}_I). \quad (23.90)$$

Мы рассматриваем решения этого уравнения в слабом смысле, т.е. равенство достигается только после свертки левой и правой части с основными функциями. Если ограничить класс рассматриваемых решений только теми решениями, которые равны нулю на бесконечности, то решение единственно и имеет вид (см., например, [3], глава III, §11, пункт 8).

$$\varphi(t, \mathbf{x}) = -G \sum_I \frac{m_I}{|\mathbf{x} - \mathbf{q}_I(t)|}, \quad (23.91)$$

где

$$|\mathbf{x} - \mathbf{q}_I| = \sqrt{(x^1 - q_I^1)^2 + (x^2 - q_I^2)^2 + (x^3 - q_I^3)^2} = \sqrt{-\eta_{\mu\nu}(x^\mu - q_I^\mu)(x^\nu - q_I^\nu)}$$

– расстояние от точки пространства $\mathbf{x} = \{x^\mu\} \in \mathbb{R}^3$ до I -той частицы в момент времени t .

Таким образом мы нашли потенциал гравитационного поля для произвольного движения частиц, которое описывается функциями $\mathbf{q}_I(t)$. Принято говорить, что гравитационное взаимодействие в механике Ньютона является дальнодействующим и распространяется с бесконечной скоростью. Это отражает тот факт, что решение (23.91) в произвольной точке пространства в произвольный момент времени однозначно определяется только массами частиц и их расположением в тот же момент времени. Можно сказать по другому: изменение положения частицы мгновенно приводит к изменению гравитационного поля во всем пространстве.

Теперь подставим решение для потенциала гравитационного поля (23.91) в уравнения движения точечных частиц (23.89),

$$\ddot{q}_{I\mu} = -G \frac{\partial}{\partial q_I^\mu} \sum_J \frac{m_J}{|\mathbf{q}_I - \mathbf{q}_J|},$$

чтобы полностью исключить потенциал гравитационного поля. Однако на этом этапе возникает серьезная трудность, т.к. правая часть уравнения расходится в точке $\mathbf{q}_I = \mathbf{q}_J$ и поэтому неопределена. Чтобы устранить эту трудность мы отбросим в сумме слагаемое с $I = J$. Физически это означает, что частица не взаимодействует с собственным гравитационным полем. Таким образом получаем систему, состоящую из $3N$ обыкновенных дифференциальных уравнений:

$$\ddot{q}_I^\mu = -G \sum_{J \neq I} m_J \frac{q_I^\mu - q_J^\mu}{|\mathbf{q}_I - \mathbf{q}_J|^3}. \quad (23.92)$$

В механике Ньютона уравнение (23.92) интерпретируется следующим образом. Если имеется всего две частицы с массами m_1 и m_j , то между ними возникает притяжение, обусловленное гравитационным взаимодействием. При этом сила $\mathbf{F}_{1j} = \{F_{1j}^\mu\}$, действующая на частицу m_1 со стороны частицы m_j , равна

$$F_{1j}^\mu = -Gm_1m_j \frac{q_1^\mu - q_j^\mu}{|\mathbf{q}_1 - \mathbf{q}_j|^3}. \quad (23.93)$$

Это – закон всемирного тяготения Ньютона.

Замечание. В настоящее время закон всемирного тяготения Ньютона подтвержден экспериментально с высокой степенью точности в лабораторных условиях и в небесной механике. С его помощью в первом приближении рассчитывают движение планет в солнечной системе и звезд в галактиках. Численное значение гравитационной постоянной в системе СГС равно

$$G = (6,673 \pm 0,003) \cdot 10^{-8} \frac{\text{см}^3}{\text{г} \cdot \text{сек}^2}.$$

Общая теория относительности в первом приближении приводит к результатам, которые совпадают с результатами, полученными в рамках механики Ньютона. Кроме этого общая теория относительности приводит к поправкам, которые называются *постньютоновскими*. $\square$

Систему уравнений движения для точечных частиц (23.92) можно получить из эффективного действия

$$S_{\text{eff}} = \int dt \left(- \sum_{i=1}^N \frac{1}{2} \dot{q}_i^\mu \dot{q}_i^\nu \eta_{\mu\nu} + \frac{1}{2} G \sum_{i=1}^N \sum_{j \neq i}^N \frac{m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|} \right), \quad (23.94)$$

которое варьируется только по траекториям частиц. Это действие получается из исходного действия для точечных частиц (23.87) подстановкой в него общего решения уравнения Эйлера–Лагранжа для потенциала (23.91). Эта процедура подстановки решения части уравнений Эйлера–Лагранжа в исходное действие с целью исключения некоторых динамических переменных была описана в разделе эффективное действие (20.4) в общем виде. Множитель $1/2$ во втором слагаемом в действии (23.94) возникает из-за двойной суммы, где каждое слагаемое встречается дважды.

Таким образом, движение точечных частиц, между которыми действуют только гравитационные силы, сводится к системе обыкновенных дифференциальных уравнений (23.92). Это полное описание, и в таком виде оно обычно встречается в курсах классической механики. Как видим, введение гравитационного поля $\varphi(x)$ в механике Ньютона совсем не обязательно. Мы проделали более длинный путь с тем, чтобы показать аналогию с общей теорией относительности. К сожалению, решить уравнения Эйнштейна (23.68) для компонент метрики при произвольном движении частиц, как в случае механики Ньютона (23.91), не удастся. Поэтому эффективное действие для точечных частиц в общей теории относительности в настоящее время неизвестно. Более того, его просто не существует. Дело в том, что при постановке задачи Коши в общей теории относительности необходимо задать не только начальные данные для точечных частиц, но и для части компонент метрики (гравитационные волны). Поэтому полного описания гравитационного взаимодействия частиц на языке эффективного действия не может существовать.

23.9.3 Свойства тензора энергии-импульса точечных частиц

В настоящем разделе мы обсудим некоторые свойства тензора энергии-импульса точечных частиц: ковариантное сохранение тензора энергии-импульса, аналогию с тензором энергии-импульса сплошной среды, неотрицательность следа тензора энергии-импульса и ультрарелятивистский предел. Для краткости мы не будем писать знак тильды над геометрическими объектами, построенными при нулевом кручении и неметричности.

Ковариантное сохранение тензора энергии-импульса

Ниже мы проведем формальные выкладки, предполагая, что компоненты метрики являются достаточно гладкими функциями на мировых линиях частиц. Как уже отмечалось, это предположение неверно: компоненты метрики расходятся на мировых линиях. Тем не менее мы проведем эти выкладки по нескольким причинам. Во-первых, рассматриваемый вопрос очень важен. Во-вторых, приведенные выкладки часто можно встретить в научной литературе и, в-третьих, данный пример покажет насколько легко получить неверный результат в случае работы с обобщенными функциями.

Поскольку тензор энергии-импульса (23.70) получен из вариации действия, инвариантного относительно общих преобразований координат, то, в силу второй теоремы Нетер (см. раздел 20.3), между уравнениями движения существует линейная зависимость,

$$2\sqrt{|g|}\nabla_\gamma \left(\frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta g_{\beta\gamma}} \right) g_{\alpha\beta} + \sum_i \frac{\delta S}{\delta q_i^\alpha} = 0.$$

На языке уравнений движения (23.68), (23.69) это означает следующее. Возьмем ковариантную дивергенцию от уравнения (23.68). Поскольку тензор Эйнштейна удовлетворяет свернутым тождествам Бианки (20.67),

$$\nabla_\alpha G^{\alpha\beta} = 0,$$

то тензор энергии-импульса точечных частиц ковариантно сохраняется,

$$\nabla_\alpha T_m^{\alpha\beta} = 0, \quad (23.95)$$

для любого решения системы уравнений движения для точечных частиц (23.69). То есть условия интегрируемости уравнений Эйнштейна выполняются автоматически, если выполнены уравнения движения точечных частиц.

Явная проверка ковариантного сохранения тензора энергии-импульса точечных частиц (23.95) требует осторожности. Поэтому проведем соответствующие вычисления, для простоты, для одной частицы,

$$\begin{aligned} \nabla_\beta T_m^{\alpha\beta} &= \partial_\alpha T_m^{\alpha\beta} + \Gamma_\alpha T_m^{\alpha\beta} + \Gamma_{\alpha\gamma}{}^\beta T_m^{\alpha\gamma} = \\ &= \frac{m}{\sqrt{|g|}} \left[\partial_\alpha \int d\tau \dot{q}^\alpha \dot{q}^\beta \delta(x - q) + \Gamma_{\alpha\gamma}{}^\beta \int d\tau \dot{q}^\alpha \dot{q}^\gamma \delta(x - q) \right], \end{aligned}$$

где мы воспользовались определением тензора энергии-импульса (23.70) и тождеством (6.49)

$$\partial_\alpha \sqrt{|g|} = \sqrt{|g|} \Gamma_\alpha.$$

Во втором слагаемом символы Кристоффеля можно внести под знак интеграла и рассматривать их, как функции от q ввиду наличия δ -функции. Для первого слагаемого в слабом смысле справедливо равенство

$$\partial_\alpha \int d\tau \dot{q}^\alpha \dot{q}^\beta \delta(x - q) = \int d\tau \ddot{q}^\beta \delta(x - q). \quad (23.96)$$

Чтобы доказать это равенство, напомним определения [3].

Определение. Назовем функцию φ на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, *финитной*, если ее носитель (замыкание множества точек, в которых функция отлична от нуля) компактен. Пространство $\mathcal{D}(\mathbb{M})$, состоящее из всех финитных бесконечно дифференцируемых функций на $\mathbb{M}$ называется пространством *основных (пробных) функций*. *Обобщенной функцией* или *распределением* f на многообразии $\mathbb{M}$ называется линейный непрерывный функционал на пространстве основных функций. Равенство двух обобщенных функций $f = g$ называется *слабым*, если справедливо равенство

$$\int_{\mathbb{M}} dx \varphi f = \int_{\mathbb{M}} dx \varphi g,$$

для всех $\varphi \in \mathcal{D}(\mathbb{M})$. □

Мы не будем останавливаться на определении топологии в пространствах основных и обобщенных функций, отсылая читателя к [3].

Вернемся к равенству (23.96). Умножим левую часть на пробную функцию $\varphi \in \mathcal{D}(\mathbb{M})$ и проинтегрируем по $\mathbb{M}$:

$$\begin{aligned} \int_{\mathbb{M}} dx \varphi \partial_\alpha \left(\int d\tau \dot{q}^\alpha \dot{q}^\beta \delta(x - q) \right) &= - \int_{\mathbb{M}} dx \partial_\alpha \varphi \int d\tau \dot{q}^\alpha \dot{q}^\beta \delta(x - q) = \\ &= - \int_{\mathbb{M}} dx \int d\tau \frac{d\varphi}{d\tau} \dot{q}^\beta \delta(x - q) = - \int d\tau \frac{d\varphi}{d\tau} \dot{q}^\beta = \int d\tau \varphi \ddot{q}^\beta, \end{aligned}$$

где мы два раза проинтегрировали по частям, воспользовались равенством $d\varphi/d\tau = \dot{q}^\alpha \partial\varphi/\partial q^\alpha$ и взяли интеграл по $\mathbb{M}$, используя δ -функцию. Интеграл по $\mathbb{M}$ от правой части (23.96) с пробной функцией приводит, очевидно, к тому же результату. Таким образом, слабое равенство (23.96) доказано.

Теперь равенство (23.95) можно переписать в виде

$$\nabla_\alpha T_m^{\alpha\beta} = \frac{1}{\sqrt{|g|}} \int d\tau \sum_i m_i (\ddot{q}_i^\beta + \Gamma_{\alpha\gamma}^{\beta} \dot{q}_i^\alpha \dot{q}_i^\gamma) \delta(x - q), \quad (23.97)$$

где мы вернулись к общему случаю N частиц. Отсюда следует, что тензор энергии-импульса точечных частиц ковариантно сохраняется, если выполнены уравнения движения (23.69). Таким образом мы привели прямое доказательство

Предложение 23.9.2. *Тензор энергии-импульса точечных частиц ковариантно сохраняется (23.95) если выполнены уравнения движения для точечных частиц.*

Зададимся вопросом: “Верно ли обратное утверждение?” Для ответа умножим равенство (23.97) на пробную функцию и проинтегрируем по $\mathbb{M}$, используя δ -функцию. Если тензор энергии-импульса ковариантно сохраняется, то должно быть выполнено равенство

$$\int d\tau \varphi \frac{1}{\sqrt{|g|}} \sum_i m_i (\ddot{q}_i^\beta + \Gamma_{\alpha\gamma}^{\beta} \dot{q}_i^\alpha \dot{q}_i^\gamma) = 0.$$

Поскольку пробная функция φ произвольна, и ее носитель может быть отличен от нуля для одной произвольной частицы, то отсюда следуют уравнения движения для точечных частиц (23.69). При этом существенно, что равенство (23.97) понимается в слабом смысле.

Вывод о том, что уравнения геодезических следуют из ковариантного закона сохранения тензора энергии-импульса точечных частиц, что, в свою очередь, необходимым условием совместности уравнений Эйнштейна с источниками в виде точечных частиц, широко распространен (см., например, [167], глава I, §6). Однако этот вывод не верен. А именно, если найдено точное решение уравнений Эйнштейна с источниками в виде точечных частиц, то отсюда не следует, что уравнения геодезических будут автоматически удовлетворены. Это связано с расходимостями компонент метрики на мировых линиях частиц. Пример одной точечной частицы рассмотрен в разделе 34.7.

Аналогия с тензором энергии-импульса сплошной среды

Введем новую временную координату $\tau = \tau(x^0, x^\mu)$ в пространстве-времени $\mathbb{M}$ таким образом, чтобы вдоль каждой мировой линии частицы она совпадала с собственным временем $\tau(x^\alpha = q_1^\alpha) = \tau$. Это всегда можно сделать, причем не единственным образом, т.к. траектории всех частиц не пересекаются, времениподобны, а канонический параметр определен с точностью до сдвигов. Продолжим векторные поля скоростей $\dot{q}_1^\alpha$, определенных вдоль мировых линий частиц, на все пространство-время гладким образом. Тогда в новой системе координат τ, x^μ производные $\dot{q}_1^\alpha$ в выражении (23.70) можно заменить на частные производные $\dot{x}^\alpha = \partial x^\alpha / \partial \tau$ ввиду наличия δ -функций, и вынести за знак интегрирования

$$T_m^{\alpha\beta} = \frac{\dot{x}^\alpha \dot{x}^\beta}{\sqrt{|g|}} \int d\tau \sum_i m_i \delta(x - q_i) = \frac{\dot{x}^\alpha \dot{x}^\beta}{\sqrt{|g|}} \sum_i \frac{m_i}{\dot{q}_i^0} \delta(\mathbf{x} - \mathbf{q}_i), \quad (23.98)$$

где мы проинтегрировали одну δ -функцию: $\delta(x - q_i^0)$. Полученное выражение для тензора энергии-импульса точечных частиц имеет такой же вид, как и для сплошной среды (23.125). Для точечных частиц давление равно нулю, $\mathcal{P} = 0$, а плотность энергии принимает вид

$$\mathcal{E} = \frac{1}{\sqrt{|g|}} \sum_i \frac{m_i}{\dot{q}_i^0} \delta(\mathbf{x} - \mathbf{q}_i).$$

Поскольку, по построению, временная координата x^0 на траекториях частиц совпадает с собственным временем τ , то $\dot{q}_i^0 = 1$ и выражение для плотности энергии приобретает интуитивно ясную форму,

$$\mathcal{E} = \frac{1}{\sqrt{|g|}} \sum_i m_i \delta(\mathbf{x} - \mathbf{q}_i).$$

То есть энергия сосредоточена в точках расположения частиц, и каждая частица несет энергию, которая равна ее массе. Тензор энергии-импульса точечных частиц соответствует пылевидной материи, поскольку давление равно нулю.

След тензора энергии-импульса

Вернемся в произвольную систему координат. Из формулы (23.71) следует выражение для следа тензора энергии-импульса произвольного распределения точечных

частиц

$$T_m^{\alpha}{}_{\alpha} = \frac{1}{\sqrt{|g|}} \sum_I \frac{m_I}{\dot{q}_I^0} \delta(\mathbf{x} - \mathbf{q}_I), \quad (23.99)$$

где использовано равенство $\dot{q}_I^2 = 1$ для канонического параметра. Поскольку $m_I > 0$ и $\dot{q}_I^0 > 0$, то след тензора энергии-импульса положителен (при этом мы рассматриваем δ -функцию, как положительную).

Поскольку след тензора энергии-импульса положителен для *произвольного* распределения точечных частиц, то в моделях математической физики делается предположение о том, что след тензора энергии-импульса для любой обычной (наблюдаемой) материи всегда неотрицателен, $T^{\alpha}{}_{\alpha} \geq 0$. При этом равенство следа тензора энергии-импульса нулю, как будет показано ниже, достигается только для частиц, движущихся со скоростью света, или излучения.

След тензора энергии-импульса электромагнитного поля (25.143), который соответствует электромагнитному излучению, также равен нулю. Это согласуется с утверждением о том, что след тензора энергии-импульса произвольного распределения ультрарелятивистских частиц равен нулю. Напомним, что в квантовой электродинамике электромагнитное поле описывает безмассовые частицы – фотоны, которые распространяются со скоростью света.

Мы выделили рассмотрение следа тензора энергии-импульса точечных частиц в отдельный пункт именно в свете последнего замечания, т.к. положительность следа тензора энергии-импульса для произвольной материи ниоткуда больше не следует, и в то же время ведет к важным следствиям.

Ультрарелятивистский предел

Рассмотрим ультрарелятивистский предел для тензора энергии-импульса точечных частиц (23.98). Этот предел, прежде всего, требует определения, потому что, глядя на определение собственной скорости частицы (23.54), непонятно что и куда стремиться. Тензор энергии-импульса точечных частиц и его след можно выразить через наблюдаемую скорость

$$T_m^{\alpha\beta} = \frac{v^{\alpha}v^{\beta}}{\sqrt{|g|}v^2} \sum_I m_I \delta(\mathbf{x} - \mathbf{q}_I), \quad (23.100)$$

$$T_m^{\alpha}{}_{\alpha} = \frac{\sqrt{v^2}}{\sqrt{|g|}} \sum_I m_I \delta(\mathbf{x} - \mathbf{q}_I), \quad (23.101)$$

где мы воспользовались равенствами (23.58) и (23.59).

Определение. Предел, когда квадрат наблюдаемой скорости частицы (23.56) стремится к нулю,

$$v^2 = v^{\alpha}v^{\beta}g_{\alpha\beta} \rightarrow 0, \quad (23.102)$$

называется *ультрарелятивистским пределом* для точечной частицы. $\square$

Если метрика имеет блочно диагональный вид (23.77), то

$$v^2 = g_{00} - \mathbf{v}^2, \quad \text{и} \quad \{u^{\alpha}\} = \left\{ \frac{1}{\sqrt{g_{00} - \mathbf{v}^2}}, \frac{v^{\mu}}{\sqrt{g_{00} - \mathbf{v}^2}} \right\}$$

где $v^2 := -v^\mu v^\nu g_{\mu\nu} \geq 0$ – квадрат пространственной наблюдаемой скорости и собственная скорость частицы u^α определена формулой (23.54). В этом случае ультрарелятивистский предел соответствует пределу $v^2 \rightarrow g_{00}$. Для пространства Минковского это означает, что наблюдаемая скорость частицы стремится к скорости света.

В ультрарелятивистском пределе компоненты самого тензора энергии-импульса точечных частиц (23.100) не определены, однако след тензора энергии-импульса (23.101) стремится к нулю

$$\lim_{v^2 \rightarrow 0} T_m^\alpha{}_\alpha = 0.$$

Это утверждение справедливо для произвольного количества и распределения точечных частиц, находящихся только под действием гравитационных сил. Если присутствуют другие взаимодействия, то действие для точечных частиц (23.62) может измениться и, следовательно, изменится выражение для тензора энергии-импульса. В таком случае требуется дополнительное исследование.

23.10 Ньютонов предел

Для того, чтобы сказать, что общая теория относительности не противоречит экспериментальным данным, желательно показать, что теория тяготения Ньютона в каком то смысле (приближении) следует из уравнений Эйнштейна. Поскольку гравитация Ньютона находится в хорошем согласии с экспериментом, то в этом случае можно утверждать, что общая теория относительности описывает гравитационные взаимодействия по крайней мере не хуже, чем законы Ньютона. Такое приближение существует, и будет описано в настоящем разделе.

В разделе 23.9.1 было показано, что закон всемирного тяготения следует из уравнений Эйнштейна в частном случае, а именно, для решения Шварцшильда. Ниже мы рассмотрим общий случай.

Сначала сделаем общее замечание. Уравнения Эйнштейна существенно нелинейны, в то время как гравитация Ньютона линейна: гравитационные потенциалы различных массивных тел просто складываются. Поэтому естественно ожидать, что закон всемирного тяготения вытекает из уравнений Эйнштейна в линейном приближении.

Рассмотрим вакуумные уравнения Эйнштейна без космологической постоянной в четырехмерном пространстве-времени

$$\kappa \left(R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R \right) = -\frac{1}{2} T_{\alpha\beta}. \quad (23.103)$$

Будем считать, что пространство-время топологически тривиально, и существует глобальная система координат, в которой метрика, удовлетворяющая уравнениям (23.103), мало отличается от метрики Лоренца в пространстве Минковского $\mathbb{R}^{1,3}$:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + \epsilon h_{\alpha\beta}, \quad \epsilon \ll 1. \quad (23.104)$$

При этом мы считаем малыми также все частные производные: $\partial_\alpha g_{\beta\gamma} \sim \epsilon$. Символы Кристоффеля пропорциональны производным $\partial_\alpha g_{\beta\gamma}$, и поэтому их квадраты дают вклад в тензор кривизны порядка ϵ^2 . Тем самым вкладом квадратичных слагаемых по символам Кристоффеля в тензор кривизны (6.88) можно пренебречь по сравнению

со вторыми производными от компонент метрики, которые дают вклад порядка ϵ . Таким образом, в линейном приближении по ϵ тензор Риччи имеет вид

$$\frac{1}{\epsilon} R_{\alpha\beta} = \frac{1}{2} \eta^{\gamma\delta} (\partial_{\alpha\beta}^2 h_{\gamma\delta} - \partial_{\alpha\gamma}^2 h_{\beta\delta} - \partial_{\beta\gamma}^2 h_{\alpha\delta} + \partial_{\gamma\delta}^2 h_{\alpha\beta}).$$

В правой части равенства свертка проводится с помощью метрики Лоренца, т.к. выражение в скобках имеет первый порядок. Соответствующая скалярная кривизна имеет вид

$$\frac{1}{\epsilon} R = \eta^{\alpha\beta} \eta^{\gamma\delta} (\partial_{\alpha\beta}^2 h_{\gamma\delta} - \partial_{\alpha\gamma}^2 h_{\beta\delta}).$$

Введем новые переменные

$$\bar{h}_{\alpha\beta} := h_{\alpha\beta} - \frac{1}{2} \eta_{\alpha\beta} h, \quad (23.105)$$

где $h := \eta^{\alpha\beta} h_{\alpha\beta}$ – след возмущения метрики. Обратное преобразование имеет вид

$$h_{\alpha\beta} = \bar{h}_{\alpha\beta} - \frac{1}{2} \eta_{\alpha\beta} \bar{h}, \quad \bar{h} := \eta^{\alpha\beta} \bar{h}_{\alpha\beta}. \quad (23.106)$$

Тогда тензор Эйнштейна в линейном приближении равен следующему выражению

$$\frac{1}{\epsilon} \left(R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R \right) = \frac{1}{2} \eta^{\gamma\delta} (\partial_{\gamma\delta}^2 \bar{h}_{\alpha\beta} - \partial_{\alpha\gamma}^2 \bar{h}_{\beta\delta} - \partial_{\beta\gamma}^2 \bar{h}_{\alpha\delta} + \eta_{\alpha\beta} \eta^{\epsilon\zeta} \partial_{\gamma\epsilon}^2 \bar{h}_{\delta\zeta}).$$

Теперь воспользуемся инвариантностью действия Гильберта–Эйнштейна относительно общих преобразований координат. Рассмотрим бесконечно малые преобразования координат, которые генерируются некоторым векторным полем (2.96),

$$x^\alpha \mapsto x^\alpha + \epsilon u^\alpha(x). \quad (23.107)$$

При этом компоненты метрики получают приращение (2.104)

$$h_{\alpha\beta} \mapsto h_{\alpha\beta} + \partial_\alpha u_\beta + \partial_\beta u_\alpha. \quad (23.108)$$

Возьмем в качестве векторного поля u^α произвольное решение уравнения

$$\square u^\alpha = -\partial_\beta \bar{h}_\alpha{}^\beta,$$

где $\square := \eta^{\alpha\beta} \partial_\alpha \partial_\beta$ – оператор Даламбера. Тогда в новой системе координат возмущение компонент метрики будет удовлетворять уравнению

$$\partial_\beta \bar{h}_\alpha{}^\beta = 0. \quad (23.109)$$

Это есть ни что иное как условие гармоничности координат (16.65) в линейном приближении.

Замечание. Гармонические координаты в общей теории относительности являются аналогом лоренцевой калибровки в электродинамике (см. раздел 25.3). $\square$

С учетом условия гармоничности уравнения Эйнштейна (23.103) принимают вид

$$\epsilon \square \bar{h}_{\alpha\beta} = -\frac{1}{\kappa} T_{\alpha\beta}. \quad (23.110)$$

Замечание. Если поля материи отсутствуют, $T_{\alpha\beta} = 0$, то система уравнений (23.109), (23.110) совпадает с уравнениями для безмассового поля со спином 2 в плоском пространстве-времени Минковского [168]. Поэтому общую теорию относительности в целом можно рассматривать как теорию безмассового поля со спином 2 и с некоторым самодействием, которое соответствует отброшенным нелинейным членам. Следует однако заметить, что понятие массы и спина требует наличия метрики Лоренца, которая является фоновой метрикой для линейного приближения. В общем случае, без обращения к линейному приближению, утверждению о том, что метрика описывает безмассовое поле спина 2 придать точный смысл весьма затруднительно. $\square$

Рассмотрим в качестве источника в уравнениях Эйнштейна одну частицу массы M . Поскольку мы рассматриваем слабые гравитационные поля, то будем считать, что $M \sim \epsilon$. Этой частице соответствует тензор энергии-импульса (23.71)

$$T_{\alpha\beta} = \frac{1}{\sqrt{|g|}} M \frac{\dot{q}_\alpha \dot{q}_\beta}{\dot{q}^0} \delta(\mathbf{x} - \mathbf{q}).$$

В линейном приближении по ϵ можно сделать замену $\sqrt{|g|} \mapsto 1$.

Предположим, что частица покоится в начале координат, т.е. $\{\dot{q}^\alpha\} = (1, 0, 0, 0)$ и $\{q^\alpha\} = (\tau, 0, 0, 0)$. Предположим также, что компоненты метрики не зависят от времени (статическое решение). Тогда полная система уравнений Эйнштейна примет вид

$$\kappa \epsilon \Delta \bar{h}_{00} = M \delta(\mathbf{x}), \quad (23.111)$$

$$\Delta \bar{h}_{0\mu} = \Delta \bar{h}_{\mu\nu} = 0, \quad (23.112)$$

где $\Delta := \partial_1^2 + \partial_2^2 + \partial_3^2$ – лапласиан в трехмерном евклидовом пространстве. Если предположить, что компоненты возмущений метрики $\bar{h}_{\alpha\beta}$ стремятся к нулю на бесконечности, то уравнения (23.112) имеют единственное решение

$$\bar{h}_{0\mu} = 0, \quad \bar{h}_{\mu\nu} = 0.$$

Для сравнения уравнения (23.111) с законом всемирного тяготения, необходимо восстановить размерные постоянные. Во-первых, положим

$$\epsilon \bar{h}_{00} = \frac{4\varphi}{c^2},$$

где φ – потенциал гравитационного поля. Это следует из нерелятивистского предела для точечной частицы (23.79). Кроме того, в правую часть уравнения (23.111) надо вставить множитель c^2 : один множитель c следует из опущенного множителя в действии для точечной частицы (23.60), а второй – из равенства $\dot{q}^0 = c$. Если после этого положить

$$\kappa := \frac{c^4}{16\pi G}, \quad (23.113)$$

где G – гравитационная постоянная в законе тяготения Ньютона, то уравнение (23.111) совпадет с уравнением Пуассона для гравитационного поля (23.90). В этом случае решение уравнения Пуассона (23.111) примет вид

$$\varphi = -G \frac{M}{r},$$

где $r = |\mathbf{x}|$.

Ясно, что для найденных компонент метрики калибровочное условие (23.109) выполнено, и, следовательно, найдено самосогласованное решение задачи.

Поскольку след $\epsilon \bar{h} = 4\varphi/c^2$, то из формулы для обратного преобразования возмущения компонент метрики (23.106), получаем, что в ньютоновом приближении метрика имеет вид

$$g = \begin{pmatrix} 1 - G\frac{2M}{rc^2} & 0 & 0 & 0 \\ 0 & -1 - G\frac{2M}{rc^2} & 0 & 0 \\ 0 & 0 & -1 - G\frac{2M}{rc^2} & 0 \\ 0 & 0 & 0 & -1 - G\frac{2M}{rc^2} \end{pmatrix}. \quad (23.114)$$

Напомним, что метрика имеет такой вид в линейном приближении в гармонической системе координат.

Таким образом, мы показали, что теория тяготения Ньютона согласуется с общей теорией относительности. Она возникает в статическом случае для слабых гравитационных полей. При этом константа связи перед действием Гильберта–Эйнштейна имеет вид (23.113). По этой причине многие авторы записывают действие Гильберта–Эйнштейна именно с такой константой связи, часто полагая $c = 1$.

Замечание. Наш выбор выражения для тензора кривизны, тензора Риччи и скалярной кривизны через метрику и аффинную связность согласован с обычными обозначениями в теории калибровочных полей и приспособлен к анализу общей теории относительности. В то же время он отличается от стандартных определений, принятых в математической литературе. Например, скалярная кривизна сферы радиуса r , вложенной в трехмерное евклидово пространство, $\mathbb{S}_r^2 \hookrightarrow \mathbb{R}^3$, равна $-2/r^2$. В этом нет ничего страшного. Просто об этом следует помнить. $\square$

Несмотря на то, что общая теория относительности содержит в себе теорию тяготения Ньютона в качестве предельного случая, отметим принципиальное отличие. В механике Ньютона свободная частица движется по прямой линии. Если она находится в поле другой массивной частицы, то на нее действует сила гравитационного притяжения. Теперь она уже не является свободной и ее траектория отличается от прямой линии в соответствии с законом всемирного тяготения. В общей теории относительности ситуация совершенно иная. Массивная частица искривляет пространство-время в соответствии с уравнениями Эйнштейна. Пробная частица в гравитационном поле остается свободной и движется вдоль экстремали. Однако теперь с точки зрения внешнего наблюдателя экстремаль не является прямой линией, поскольку пространство-время перестает быть плоским из-за наличия массивной частицы.

23.11 Гравитационные волны

В механике Ньютона гравитационных волн нет, что вытекает из системы уравнений (23.88), (23.89). При этом изменение положения одного из массивных тел мгновенно приводит к изменению гравитационного поля во всем пространстве. Кроме этого, если массивные тела отсутствуют, то потенциал гравитационного поля равен нулю. В

общей теории относительности ситуация другая. Во-первых, гравитационные взаимодействия распространяются с конечной постоянной скоростью света c в локально инерциальной системе отсчета. Во-вторых, даже если материальные тела отсутствуют, уравнения Эйнштейна допускают нетривиальные решения в виде гравитационных волн. То есть гравитационное поле может быть отлично от нуля даже если материальные тела отсутствуют. В настоящем разделе мы изучим решения уравнений Эйнштейна, описывающие гравитационные волны.

Рассмотрим вакуумные уравнения Эйнштейна без космологической постоянной

$$R_{\alpha\beta} = 0. \quad (23.115)$$

Как и в предыдущем разделе будем считать, что пространство-время топологически тривиально, $M \approx \mathbb{R}^4$, и существует глобальная система координат, в которой метрика мало отличается от метрики Лоренца (23.104). В нулевом порядке по ϵ вакуумные уравнения Эйнштейна, очевидно, удовлетворяются, т.к. кривизна пространства Минковского равна нулю. Найдем решение уравнений (23.115) в первом порядке по ϵ .

Используя инвариантность действия Гильберта–Эйнштейна относительно общих преобразований координат, выберем систему отсчета таким образом, чтобы выполнялось калибровочное условие (23.109) (гармонические координаты). Тогда все компоненты метрики в первом порядке будут удовлетворять волновому уравнению

$$\square \bar{h}_{\alpha\beta} = 0 \quad \Leftrightarrow \quad \square h_{\alpha\beta} = 0,$$

где компоненты $\bar{h}_{\alpha\beta}$ определены уравнением (23.105) и $\square$ – оператор Даламбера в пространстве Минковского. Однако далеко не все компоненты метрики являются независимыми и описывают физические степени свободы, от которых нельзя избавиться путем выбора соответствующей системы координат.

Покажем это. Во-первых, выберем гармоническую систему координат. Тогда в линейном приближении выполнено уравнение (23.109), которое рассматривается как калибровочное условие. Это условие не фиксирует систему координат однозначно. Действительно, допустим, что в некоторой системе координат это условие выполнено. Совершим преобразование координат $x^\alpha \mapsto x^\alpha + \epsilon u^\alpha$, где все компоненты векторного поля u удовлетворяют волновому уравнению

$$\square u^\alpha = 0. \quad (23.116)$$

Нетрудно проверить, что в новой системе координат калибровочное условие (23.109) будет также выполнено. Следовательно, оставшуюся свободу в выборе системы координат можно использовать для того, чтобы зафиксировать дополнительные компоненты метрики.

Чтобы найти подходящие дополнительные калибровочные условия, необходимо решить уравнения (23.116) с некоторыми начальными условиями. Посмотрим как преобразуется след возмущений метрики $h := h_\alpha^\alpha$ и компоненты $h_{0\mu}$, $\mu = 1, 2, 3$, при бесконечно малых преобразованиях координат (23.107):

$$\begin{aligned} h &\mapsto h + 2\partial_\alpha u^\alpha, \\ h_{0\mu} &\mapsto h_{0\mu} + \partial_0 u_\mu + \partial_\mu u_0. \end{aligned}$$

Рассмотрим пространственное сечение $x^0 := t = \text{const}$ в качестве поверхности Коши для уравнения (23.116). На этой поверхности найдем какое либо решение системы

линейных уравнений

$$2(\dot{u}_0 + \partial_\mu u^\mu) = -h, \quad (23.117)$$

$$2(\Delta u_0 + \partial_\mu \dot{u}^\mu) = -\dot{h}, \quad (23.118)$$

$$\dot{u}_\mu + \partial_\mu u_0 = -h_{0\mu}, \quad (23.119)$$

$$\Delta u_\mu + \partial_\mu \dot{u}_0 = -\dot{h}_{0\mu}, \quad (23.120)$$

где точка обозначает дифференцирование по времени t . Из этой системы уравнений определяем компоненты u^α и их производные по времени $\dot{u}^\alpha$ на поверхности Коши. После этого решаем задачу Коши для уравнений (23.116) с найденными начальными условиями в обе стороны по времени и определяем векторное поле u во всем пространстве-времени. Поскольку выполнены уравнения (23.117)–(23.120), то на поверхности Коши справедливы равенства

$$\begin{aligned} h &= 0, & \dot{h} &= 0, \\ h_{0\mu} &= 0, & \dot{h}_{0\mu} &= 0. \end{aligned} \quad (23.121)$$

Это очевидно, если заметить, что уравнения (23.118) и (23.120) возникают после дифференцирования уравнений (23.117) и (23.119) по времени и использования уравнений движения. Поскольку компоненты h и $h_{0\mu}$ удовлетворяют волновому уравнению, то условия (23.121) выполнены также во всем пространстве-времени.

Определение. Система координат, в которой в линейном приближении выполнены условия

$$\partial_\beta h_\alpha{}^\beta = 0, \quad h = 0, \quad h_{0\mu} = 0 \quad (23.122)$$

называется *радиационной калибровкой*. $\square$

В электродинамике аналогичная калибровка называется кулоновской.

Предложение 23.11.1. *При отсутствии полей материи и космологической постоянной радиационная калибровка существует в линейном приближении к метрике Лоренца.*

Доказательство. Было приведено выше. $\square$

Несмотря на то, что в четырехмерном пространстве-времени общие преобразования координат параметризуются четырьмя произвольными функциями, мы сумели наложить восемь калибровочных условий (23.122).

Из первого условия (23.122) для $\alpha = 0$ и с учетом того, что $h_{0\mu} = 0$, получаем уравнение $\dot{h}_{00} = 0$, которое должно быть выполнено во всем пространстве-времени. Тогда уравнение движения для временной компоненты сводится к уравнению Лапласа $\Delta h_{00} = 0$. С учетом нулевых граничных условий на бесконечности получаем дополнительное условие на компоненты метрики: $h_{00} = 0$.

Рассмотрим плоскую волну, которая распространяется в направлении волнового вектора $k = \{k^\alpha\}$:

$$h_{\alpha\beta} = H_{\alpha\beta} e^{ik_\gamma x^\gamma}, \quad (23.123)$$

где $H_{\alpha\beta}$ – некоторая постоянная матрица и $k^2 := k^\alpha k_\alpha = 0$. Радиационная калибровка (23.122) для этого решения уравнений движения задается следующими восемью условиями:

$$k^\beta H_{\alpha\beta} = 0, \quad H_\alpha{}^\alpha = 0, \quad H_{0\mu} = 0.$$

Из первого и третьего условия вытекает равенство $k^0 H_{00}$. Для нетривиального решения $k^0 \neq 0$, и поэтому $H_{00} = 0$. Ввиду симметрии по индексам матрица $H_{\alpha\beta}$ имеет 10 независимых элементов. Радиационная калибровка накладывает на них 8 независимых условий. Отсюда вытекает, что в выбранной системе координат только 2 компоненты возмущенной метрики являются независимыми.

Допустим, что гравитационная волна распространяется вдоль оси x^1 , т.е. нормированный волновой вектор имеет только две отличные от нуля компоненты $k = (1, 1, 0, 0)$. Тогда матрица $H_{\alpha\beta}$ в радиационной калибровке имеет вид

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & A & B \\ 0 & 0 & B & -A \end{pmatrix},$$

где $H_{22} = -H_{33} := A$ и $H_{23} = H_{32} := B$ — два произвольных числа (амплитуды волн).

Введем новое понятие спиральности плоской волны. Для этого рассмотрим вращение пространства $\mathbb{R}^3 \subset \mathbb{R}^{1,3}$ на угол φ вокруг оси x^1 . Матрица вращения задана в явном виде формулой (1.71), и ясно, что такое вращение не меняет волнового вектора k . При преобразовании координат компоненты метрики преобразуются по тензорному закону

$$g_{\alpha\beta} \mapsto g'_{\alpha\beta} = S_{\alpha}^{\gamma} S_{\beta}^{\delta} g_{\gamma\delta}$$

С соответствующей матрицей вращений S . Простые вычисления приводят к равенствам для новых амплитуд:

$$\begin{aligned} A' &= \cos 2\varphi A + \sin 2\varphi B, \\ B' &= -\sin 2\varphi A + \cos 2\varphi B. \end{aligned}$$

Это означает, что при повороте системы координат на угол φ амплитуда волны поворачивается на удвоенный угол 2φ . В физике часто рассматривают комплексные амплитуды

$$H_{\pm} := H_{22} \mp i H_{23} = A \mp i B.$$

При вращении они преобразуются по правилу

$$H'_{\pm} = e^{\pm 2i\varphi} H_{\pm}.$$

Определение. Если амплитуда плоской поперечной волны при повороте на угол φ вокруг направления распространения волны поворачивается на угол $h\varphi$, то говорят, что волна имеет *спиральность* h . $\square$

Таким образом, плоские гравитационные волны описывают поперечные волны спиральности два.

Тензор Риччи и скалярная кривизна для данного решения вакуумных уравнений Эйнштейна в линейном приближении равны, конечно, нулю. Это следует из того, что мы решаем вакуумные уравнения Эйнштейна без космологической постоянной (23.6). Тем не менее полный тензор кривизны отличен от нуля. В линейном приближении тензор кривизны имеет вид (см. выражение (6.88))

$$\frac{1}{\epsilon} R_{\alpha\beta\gamma\delta} = \frac{1}{2} (\partial_{\alpha\gamma}^2 h_{\beta\delta} - \partial_{\alpha\delta}^2 h_{\beta\gamma} - \partial_{\beta\gamma}^2 h_{\alpha\delta} + \partial_{\beta\delta}^2 h_{\alpha\gamma}). \quad (23.124)$$

Простые вычисления показывают, что среди 20 независимых компонент тензора кривизны только 9 отличны от нуля:

$$\begin{aligned} R_{0303} = -R_{0202} = R_{1313} = -R_{1212} = R_{0212} = -R_{0313} &= \frac{1}{2}\epsilon A e^{i(t-x^1)}, \\ R_{0203} = R_{1213} = -R_{0213} &= -\frac{1}{2}\epsilon B e^{i(t-x^1)}. \end{aligned}$$

При преобразовании координат компоненты тензора кривизны ведут себя ковариантным образом – на то он и тензор. Однако в линейном приближении они не просто ковариантны, а инвариантны. Нетрудно проверить, что выражение (23.124) действительно инвариантно относительно преобразований (23.108) с произвольным вектором u .

Из явного выражения для нетривиальных компонент тензора кривизны вытекает, что амплитуды волн A и B нельзя обратить в нуль никаким преобразованием координат. Следовательно, они описывают физические распространяющиеся степени свободы.

Поскольку вакуумные уравнения Эйнштейна в линейном приближении линейны, то им будет удовлетворять произвольная суперпозиция плоских волн. В частности, поправки к метрике вида

$$h_{\alpha\beta} = H_{\alpha\beta}[f(x^1 - t) + g(x^1 + t)],$$

где f и g – произвольные функции, описывающие распространение волн вдоль оси x^1 в положительную и отрицательную стороны, также удовлетворяют линейризованным уравнениям Эйнштейна в радиационной калибровке. Отсюда, в частности, следует, что для однозначного задания волнового решения уравнений Эйнштейна необходимо задать четыре функции на поверхности Коши: по две для каждой волны. При этом функции можно задать произвольным образом. Таким образом, вакуумные уравнения Эйнштейна без космологической постоянной описывают распространение двух физических степеней свободы, которые порождают нетривиальную кривизну пространства-времени и не устраняются никаким преобразованием координат. Данный подсчет степеней свободы приводит к тому же результату, что и общий подход, основанный на гамильтоновом формализме, который будет рассмотрен позже в главе 24.

23.12 Сплошная среда в общей теории относительности

В правой части уравнений Эйнштейна (23.2) находится тензор энергии-импульса материи $T_m^{\alpha\beta}$. В случае скалярного, электромагнитного и других полей, уравнения движения которых следуют из вариационного принципа, правая часть уравнений Эйнштейна определяется вариацией соответствующего действия по метрике. В этом случае вопросов с определением тензора энергии-импульса материи не возникает. Некоторые из этих тензоров будут рассмотрены в дальнейшем.

В то же время в общей теории относительности существует ряд важных моделей (особенно в космологии), для которых тензор энергии-импульса материи не следует из вариационного принципа. В настоящем разделе мы определим тензор энергии-импульса материи $T_m^{\alpha\beta}$, которая рассматривается как сплошная среда, и изучим

некоторые из его свойств. При этом мы не будем опираться на вариационный принцип.

Пусть пространство-время $(\mathbb{M}, g)$ топологически тривиально $\mathbb{M} \approx \mathbb{R}^{1,3}$ и покрыто одной картой. Мы предполагаем, что координаты $\{x^\alpha\} = \{x^0, x^\mu\}$ выбраны таким образом, что координата x^0 является временем, т.е. $g_{00} > 0$. Кроме того, мы считаем, что все сечения $x^0 = \text{const}$ – пространственноподобны.

Можно привести ряд физических аргументов [169], глава IV, §35 в пользу того, что тензор энергии-импульса материи, которая рассматривается как сплошная среда, имеет вид

$$T_{\mathbb{M}}^{\alpha\beta} := (\mathcal{E} + \mathcal{P})u^\alpha u^\beta - \mathcal{P}g^{\alpha\beta}, \quad (23.125)$$

где $\mathcal{E}(x)$ и $\mathcal{P}(x)$ – плотность энергии и давление материи в точке $x \in \mathbb{M}$, и

$$u^\alpha := dx^\alpha/ds, \quad ds := \sqrt{|ds^2|}, \quad ds^2 := g_{\alpha\beta}dx^\alpha dx^\beta,$$

– четырехмерная скорость материи в точке $x \in \mathbb{M}$, которая удовлетворяет тождеству $u^\alpha u_\alpha = 1$. Здесь мы предполагаем, что каждая точка материи движется вдоль времениподобной мировой линии $x^\alpha(s)$ в будущее, т.е. $u^0 > 0$. Ясно, что мировые линии точек материи – это интегральные кривые векторного поля скорости u .

При рассмотрении моделей сплошной среды задаются, как правило, не линии тока $x^\alpha(s)$, а векторное поле скоростей $u^\alpha(s)$. Мы предполагаем, что поле скоростей является достаточно гладким, и через каждую точку многообразия проходит одна и только одна линия тока. Математически это означает, что система уравнений

$$\frac{dy^\alpha}{ds} = u^\alpha(y),$$

с начальным условием $y^\alpha(0) = x^\alpha$ имеет единственное решение для всех $x \in \mathbb{M}$, которое определено для всех $t \in \mathbb{R}$.

Поскольку u^α и $g^{\alpha\beta}$ являются, соответственно, компонентами вектора и тензора относительно преобразований координат, то мы считаем, что плотность энергии $\mathcal{E}$ и давление материи $\mathcal{P}$ являются достаточно гладкими скалярными полями на пространстве-времени $\mathbb{M}$. В этом случае правая часть равенства (23.125) представляет собой ковариантный симметричный тензор второго ранга. Для обычной (наблюдаемой) материи плотность энергии предполагается положительной, $\mathcal{E} > 0$, а давление – неотрицательным, $\mathcal{P} \geq 0$ ¹.

В используемых обозначениях, $\kappa = 1$, размерность тензора энергии-импульса совпадает с размерностью тензора Риччи:

$$[T_{\mathbb{M}\alpha\beta}] = [R_{\alpha\beta}] = l^{-2}.$$

Поскольку компоненты метрики и скорости безразмерны, то плотность энергии и давления имеют такие же размерности:

$$[\mathcal{E}] = [\mathcal{P}] = l^{-2}.$$

Мы не обсуждаем физических аргументов, приводящих к тензору энергии-импульса (23.125), отсылая читателя к монографии [169], глава IV, §35. В настоящем разделе мы рассматриваем, в основном, математические свойства данного определения.

¹ Давление, в принципе, может быть отрицательным. Примером является резина. Для нее увеличение объема по сравнению с состоянием равновесия приводит к увеличению давления. В современных космологических моделях предполагается наличие темной энергии. Этот вид материи имеет положительную плотность энергии, но отрицательное давление (см., раздел 35.3.4).

Замечание. В нерелятивистской гидро- и газодинамике все уравнения записываются таким образом, что в них входит не сама энергия и давление, а только их градиенты. Это означает, что энергия и давление определены с точностью до добавления произвольной постоянной. В общей теории относительности ситуация отличается принципиально, т.к. уравнения меняются, если к $\mathcal{E}$ или $\mathcal{P}$ добавить постоянную. В частности, наличие космологической постоянной Λ можно интерпретировать как среду с постоянной плотностью энергии $\mathcal{E} = \Lambda$ и постоянным давлением $\mathcal{P} = -\mathcal{E} = -\Lambda$. Если $\mathcal{E} > 0$, то давление отрицательно, $\mathcal{P} < 0$. Поэтому космологическую постоянную можно интерпретировать, как некоторую среду, заполняющую все пространство-время со свойствами обыкновенной резины. $\square$

Из общих физических представлений следует, что след тензора энергии-импульса для обычной материи должен быть неотрицательным [169], глава IV, §34:

$$T_{\text{M}\alpha}^{\alpha} = \mathcal{E} - 3\mathcal{P} \geq 0. \quad (23.126)$$

Этим свойством обладает, в частности, тензор энергии-импульса для произвольного распределения точечных частиц (23.99). Отсюда вытекает ограничение на давление

$$\mathcal{P} \leq \frac{\mathcal{E}}{3}. \quad (23.127)$$

Поскольку давление, по предположению, неотрицательно, то с учетом равенства (23.127) существует два крайних случая. Если материя, которой заполнена вселенная, настолько разрежена, что давление можно считать равным нулю, то говорят, что материя пылевидна. Максимальное возможное давление, $\mathcal{P} = \mathcal{E}/3$, соответствует газу ультрарелятивистских частиц, скорости которых близки к скорости света (см. раздел 23.9.3). В этом случае говорят, что вселенная заполнена газом излучения или, просто, излучением.

$$\begin{aligned} \mathcal{P} &= 0 && \text{— пыль,} \\ \mathcal{P} &= \frac{\mathcal{E}}{3} && \text{— излучение.} \end{aligned} \quad (23.128)$$

Для обычной материи $0 \leq \mathcal{P} \leq \mathcal{E}/3$.

Пример 23.12.1 (Нерелятивистская гидродинамика). Рассмотрим пространство Минковского $\mathbb{R}^{1,3}$ в декартовой системе координат с метрикой Лоренца $\eta_{\alpha\beta} = \text{diag}(+ - - -)$. Пусть пространство-время заполнено идеальной (без вязкости) жидкостью. Течение жидкости описывается плотностью ρ , давлением $\mathcal{P}$ и трехмерной скоростью u^μ , $\mu = 1, 2, 3$. Тензор энергии-импульса идеальной жидкости в общей теории относительности, по определению, имеет вид (23.125). Покажем, что уравнения движения нерелятивистской идеальной жидкости (если не считать уравнения состояния) следуют из закона сохранения четырехмерного тензора энергии-импульса, $\partial_\beta T_{\text{M}}^{\alpha\beta} = 0$.

В нерелятивистском приближении мы считаем, что пространственные компоненты скорости малы: $u^0 \approx 1$, $u^\mu \sim \epsilon \ll 1$, где $\{u^\alpha\} = \{u^0, u^\mu\}$, $\alpha = 0, 1, 2, 3$, и давление мало, $\mathcal{P} \sim \epsilon^2 \ll \mathcal{E}$. Кроме того положим $\mathcal{E} = \rho$, где ρ — плотность жидкости, имея в виду формулу $\mathcal{E} = \rho c^2$ при $c = 1$. Тогда в низшем приближении компоненты тензора энергии-импульса равны:

$$\begin{aligned} T_{\text{M}}^{00} &= (\mathcal{E} + \mathcal{P})u^0u^0 - \mathcal{P} && \approx \rho, \\ T_{\text{M}}^{0\mu} &= T_{\text{M}}^{\mu 0} = (\mathcal{E} + \mathcal{P})u^0u^\mu && \approx \rho u^\mu, \\ T_{\text{M}}^{\mu\nu} &= (\mathcal{E} + \mathcal{P})u^\mu u^\nu - \mathcal{P}\eta^{\mu\nu} && \approx \rho u^\mu u^\nu - \mathcal{P}\eta^{\mu\nu}. \end{aligned} \quad (23.129)$$

Рассмотрим закон сохранения энергии-импульса $\partial_\beta T_M^{\alpha\beta} = 0$. Нулевая компонента этого равенства в главном приближении имеет вид

$$\partial_\beta T_M^{0\beta} = \partial_0 T_M^{00} + \partial_\mu T_M^{0\mu} \approx \partial_0 \rho + \partial_\mu (\rho u^\mu) = 0. \quad (23.130)$$

Полученное уравнение совпадает с *уравнением непрерывности*. Пространственные компоненты закона сохранения энергии-импульса в главном приближении приводят к равенству

$$\partial_\beta T_M^{\mu\beta} = \partial_0 T_M^{\mu 0} + \partial_\nu T_M^{\mu\nu} \approx \rho \partial_0 u^\mu + u^\mu [\partial_0 \rho + \partial_\nu (\rho u^\nu)] + \rho u^\nu \partial_\nu u^\mu - \eta^{\mu\nu} \partial_\nu \mathcal{P} = 0,$$

что, с учетом уравнения непрерывности (23.130), дает *уравнение Эйлера*

$$\partial_0 u^\mu + u^\nu \partial_\nu u^\mu = \frac{1}{\rho} \eta^{\mu\nu} \partial_\nu \mathcal{P}. \quad (23.131)$$

Если дополнить уравнение непрерывности и уравнение Эйлера уравнением состояния идеальной жидкости $\mathcal{P} = \mathcal{P}(\rho)$, связывающим давление и плотность, то получим полную систему уравнений для идеальной жидкости (пять уравнений на пять неизвестных). Таким образом, уравнения движения нерелятивистской идеальной жидкости следуют из закона сохранения четырехмерного тензора энергии-импульса (23.125), дополненного уравнением состояния.

Эта же система уравнений (23.130), (23.131) описывает движение идеального газа. Разница заключается только в уравнении состояния. Для идеального газа уравнение состояния имеет вид

$$\mathcal{P} = \frac{\rho}{\mu} RT, \quad (23.132)$$

где μ , R и T есть, соответственно, молекулярный вес, универсальная газовая постоянная и абсолютная температура. При постоянной температуре $T = \text{const}$ давление идеального газа прямо пропорционально плотности. $\square$

Рассмотренный пример показывает, что в нерелятивистском пределе ковариантное сохранение тензора энергии-импульса сплошной среды

$$\nabla_\beta T_M^{\alpha\beta} = 0$$

сводится к уравнениям нерелятивистской гидродинамики идеальной жидкости. При этом плотность энергии $\mathcal{E}$ мы отождествили с плотностью жидкости ρ . Напомним, что в релятивистской механике точечных частиц энергия и импульс частицы являются компонентами четырехмерного вектора, причем энергия представляет собой нулевую компоненту и не является скаляром. Для сплошной среды $\mathcal{E}(x)$ и $\mathcal{P}(x)$ – скалярные поля.

Наличие в пространстве-времени метрики и времениподобного векторного поля u , $u^2 = 1$, позволяет определить проекционные операторы (4.25):

$$\Pi_\alpha^\beta := u_\alpha u^\beta, \quad \Pi_\alpha^\beta := \delta_\alpha^\beta - u_\alpha u^\beta.$$

В каждой точке $x \in \mathbb{M}$ эти операторы проектируют тензорные поля, соответственно, на направление вектора скорости u и перпендикулярную гиперплоскость в касательном пространстве $\mathbb{T}_x(\mathbb{M})$. Например, проекция метрики имеет вид

$$\begin{aligned} g_{\alpha\beta}^L &:= \Pi_\alpha^\gamma \Pi_\beta^\delta g_{\gamma\delta} = u_\alpha u_\beta, \\ g_{\alpha\beta}^T &:= \Pi_\alpha^\gamma \Pi_\beta^\delta g_{\gamma\delta} = g_{\alpha\beta} - u_\alpha u_\beta. \end{aligned}$$

Ясно также, что

$$u^{\perp\alpha} := u^\beta \Pi_\beta^{\perp\alpha} = u^\alpha, \quad u^{\text{T}} := u^\beta \Pi_\beta^{\text{T}\alpha} = 0.$$

Поэтому тензор энергии-импульса (23.125) можно переписать с помощью проекционных операторов

$$T_{\text{M}}^{\alpha\beta} = \mathcal{E} g^{\perp\alpha\beta} - \mathcal{P} g^{\text{T}\alpha\beta}. \quad (23.133)$$

Поскольку тензор энергии-импульса сплошной среды (23.125) не был получен из вариационного принципа, то на него необходимо наложить дополнительное условие

$$\nabla_\beta T_{\text{M}}^{\beta\alpha} = 0, \quad (23.134)$$

которое является условием совместности уравнений Эйнштейна. Более подробно

$$(\mathcal{E} + \mathcal{P}) u^\beta \nabla_\beta u^\alpha + u^\alpha \nabla_\beta [(\mathcal{E} + \mathcal{P}) u^\beta] - g^{\alpha\beta} \partial_\beta \mathcal{P} = 0,$$

где мы воспользовались условием метричности связности Леви-Чивиты $\nabla_\beta g^{\gamma\alpha} = 0$. Проекция этого уравнения на вектор u и перпендикулярную гиперплоскость имеют следующий вид

$$(\mathcal{E} + \mathcal{P}) \nabla_\alpha u^\alpha + u^\alpha \partial_\alpha \mathcal{E} = 0, \quad (23.135)$$

$$(\mathcal{E} + \mathcal{P}) u^\beta \nabla_\beta u^\alpha - (g^{\alpha\beta} - u^\alpha u^\beta) \partial_\beta \mathcal{P} = 0, \quad (23.136)$$

где мы воспользовались уравнением $u_\alpha \nabla_\beta u^\alpha = 0$, которое следует из условия $u^2 = 0$ после дифференцирования. Легко проверить, что свертка уравнений (23.136) с ковектором u_α тождественно обращается в нуль. Следовательно, только четыре уравнения из (23.135), (23.136) являются независимыми, и они эквивалентны условию ковариантного сохранения тензора энергии-импульса $\nabla_\beta T_{\text{M}}^{\beta\alpha} = 0$.

Уравнение (23.135) является ковариантным обобщением уравнения непрерывности для нерелятивистской жидкости (23.130), а уравнение (23.136) – ковариантным обобщением уравнения Эйлера (23.131). Эти уравнения представляют собой систему уравнений *релятивистской гидродинамики*.

Система уравнений (23.135), (23.136) вместе с уравнениями Эйнштейна не образует полной системы уравнений релятивистской гидродинамики. Ее необходимо дополнить уравнением состояния. Широкий класс моделей описывается уравнением состояния $\mathcal{P} = \mathcal{P}(\mathcal{E})$, связывающим давление с плотностью энергии в каждой точке пространства-времени. Такие жидкости называются *баротропными*.

Второе слагаемое в уравнении Эйлера (23.136) после опускания индекса имеет вид

$$\Pi_\alpha^{\text{T}\beta} \partial_\beta \mathcal{P} = (\delta_\alpha^\beta - u_\alpha u^\beta) \partial_\beta \mathcal{P}.$$

Если оно равно нулю, т.е. градиент давления параллелен вектору скорости, то уравнение Эйлера упрощается $u^\beta \nabla_\beta u^\alpha = 0$. Это есть уравнение экстремалей. В этом случае точки жидкости движутся так же, как и точечные частицы под действием только гравитационных сил.

Для пылевидной материи давление равно нулю и система уравнений релятивистской гидродинамики существенно упрощается:

$$\nabla_\alpha (\mathcal{E} u^\alpha) = 0, \quad u^\beta \nabla_\beta u^\alpha = 0. \quad (23.137)$$

Мы видим, что пылевидная материя движется вдоль экстремалей, как множество точечных частиц.

При анализе общих свойств решений уравнений Эйнштейна на тензор энергии-импульса материи накладываются определенные *энергетические условия*. Перечислим три наиболее распространенных.

Определение. Говорят, что источники (поля материи) удовлетворяют:

- 1) *слабому энергетическому условию*, если для любого времениподобного векторного поля X выполнено неравенство

$$T_{\alpha\beta}X^\alpha X^\beta \geq 0; \quad (23.138)$$

- 2) *сильному энергетическому условию*, если для любого времениподобного векторного поля X выполнено неравенство

$$\rho_{\alpha\beta}X^\alpha X^\beta \geq 0, \quad \text{где} \quad \rho_{\alpha\beta} := T_{\alpha\beta} - \frac{1}{n-2}g_{\alpha\beta}T_{\mu\gamma}{}^\gamma, \quad (23.139)$$

где n – размерность пространства-времени. Это условие называется также *условием положительности Риччи* (см. уравнение (23.4));

- 3) *доминантному энергетическому условию*, если для произвольного времениподобного векторного поля X , направленного в будущее, векторное поле с компонентами $T_{\mu\beta}X^\beta$ также времениподобно и направлено в будущее. $\square$

В данном определении векторное поле X времениподобно, т.е. $X^2 := X^\alpha X_\alpha > 0$. Поэтому, не ограничивая общности, можно считать, что оно нормировано на единицу, $X^2 = 1$. Кроме того, всегда можно выбрать такую систему координат, в которой произвольное времениподобное векторное поле имеет только одну нетривиальную компоненту $X = (1, 0, \dots, 0)$. Поэтому слабое и сильное энергетические условия можно переформулировать. В произвольной системе координат, в которой координата x^0 является временем, должны выполняться равенства: $T_{00} \geq 0$ или $\rho_{00} \geq 0$, соответственно.

Ясно, что если выполнено доминантное энергетическое условие, то $T_{\alpha\beta}X^\alpha X^\beta > 0$. Следовательно, из условия 3) следует 1). Кроме того, если след тензора энергии-импульса материи положителен, как, например, для точечных частиц (см. раздел 23.9.3), то из сильного энергетического условия вытекает слабое, причем получается строгое неравенство $T_{\alpha\beta}X^\alpha X^\beta > 0$. Если же след тензора энергии-импульса материи отрицателен, то, наоборот, слабое энергетическое условие сильнее, чем сильное энергетическое условие.

Предложение 23.12.1. *Если выполнены условия $\mathcal{E} > 0$, $\mathcal{P} \geq 0$, то тензор энергии-импульса сплошной среды удовлетворяет слабому и сильному энергетическим условиям.*

Доказательство. Зафиксируем систему координат, в которой времениподобное векторное поле имеет вид $X = (1, 0, 0, 0)$. Тогда слабое энергетическое условие эквивалентно неравенству

$$T_{00} = \mathcal{E}u_0^2 + \mathcal{P}(u_0^2 - g_{00}) \geq 0.$$

Первое слагаемое в этом выражении, очевидно, неотрицательно (в принципе оно может обратиться в нуль, т.к. условия $u^0 > 0$ недостаточно для положительности нулевой ковариантной компоненты скорости u_0). Поскольку

$$u_0 := g_{00}u^0 + g_{0\mu}u^\mu,$$

то второе слагаемое принимает вид

$$\mathcal{P}(g_{00}^2(u^0)^2 + 2g_{00}g_{0\mu}u^0u^\mu + (g_{0\mu}u^\mu)^2 - g_{00}) = -\mathcal{P}g_{00}\left(g_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}}\right)u^\mu u^\nu,$$

где мы воспользовались тождеством $u^2 = 1$. Согласно теореме 4.2.2 на лоренцевом многообразии матрица в скобках в правой части равенства отрицательно определена, и поэтому слабое энергетическое условие выполнено.

Для тензора энергии-импульса сплошной среды справедливо равенство

$$\rho_{\alpha\beta} = (\mathcal{E} + \mathcal{P})u_\alpha u_\beta - \mathcal{P}g_{\alpha\beta} - \frac{1}{n-2}(\mathcal{E} - 3\mathcal{P})g_{\alpha\beta}.$$

Отсюда следует выражение для нулевой компоненты

$$\rho_{00} = \mathcal{E}(u_0^2 - g_{00}) + \mathcal{P}(u_0^2 - g_{00}) + \frac{n-3}{n-2}\mathcal{E}g_{00} + \frac{3}{n-2}\mathcal{P}g_{00}.$$

Первые два слагаемых неотрицательны в силу предыдущих рассуждений. Последние два слагаемых положительны, т.к. координата x^0 является временем и, следовательно, $g_{00} > 0$. $\square$

Замечание. В доказательстве предложения мы не использовали неравенство $\mathcal{P} \leq \mathcal{E}/3$, обеспечивающее неотрицательность следа тензора энергии-импульса. $\square$

Предложение 23.12.2. Если выполнено условие $|\mathcal{P}| < |\mathcal{E}|$, то тензор энергии-импульса сплошной среды удовлетворяет условию энергодоминантности.

Доказательство. В системе координат, где $X = (1, 0, 0, 0)$, условие энергодоминантности имеет вид

$$T_{m0}{}^\alpha T_{m0\alpha} > 0.$$

Подставляя в это неравенство выражение для тензора энергии-импульса (23.125), получим соотношение

$$(\mathcal{E}^2 - \mathcal{P}^2)u_0^2 + \mathcal{P}^2 g_{00} > 0.$$

Очевидно, что неравенство $|\mathcal{P}| < |\mathcal{E}|$ достаточно для выполнения условия энергодоминантности. $\square$

Таким образом, для обычной материи все три энергетических условия выполнены, т.к. давление удовлетворяет неравенствам $0 \leq \mathcal{P} \leq \mathcal{E}/3$.

23.12.1 Акустические фононы в нерелятивистской гидродинамике

В физике всегда большой интерес вызывали аналогии между явлениями из разных областей. В настоящем разделе мы покажем, что некоторые явления в нерелятивистской гидродинамике и общей теории относительности описываются уравнениями, которые обладают рядом одинаковых свойств.

Уравнения движения для акустических фононов следуют из классических нерелятивистских уравнений гидродинамики следующим образом. Рассмотрим некоторое точное решение уравнений гидродинамики. Тогда уравнение для малых возмущений (фононов) вблизи этого решения сводится к уравнению Даламбера с нетривиальной эффективной четырехмерной метрикой лоренцевой сигнатуры, в которой роль скорости света играет скорость звука в жидкости. Отличие от общей теории относительности сводится к тому, что эффективная метрика для фононов определяется уравнениями гидродинамики, а не уравнениями Эйнштейна. Тем не менее уравнение для фононов задается нетривиальной четырехмерной метрикой, для которой тензор

кривизны отличен от нуля. Другими словами, фононы двигаются на многообразии с нетривиальной геометрией. При этом возможно возникновение горизонтов, когда скорость течения жидкости превышает скорость звука, и, следовательно, образование акустических аналогов черных дыр.

В настоящем разделе мы следуем выводу уравнения для фононов, предложенному в [170, 171] (см. также [172]).

Рассмотрим 4-мерное галилеево пространство-время с декартовой системой координат $\{x^\alpha\}$, $\alpha = 0, 1, 2, 3$, которые мы будем обозначать индексами из начала греческого алфавита $\alpha, \beta, \dots$. Координату $x^0 \in \mathbb{R}$ мы отождествляем с временем, $x^0 = t$. Пространственные координаты $\{x^\mu\} \in \mathbb{R}^3$ мы будем обозначать индексами из середины греческого алфавита $\mu, \nu, \dots$. Жидкость без вязкости называется *идеальной* и описывается плотностью $\rho(x)$, давлением $\mathcal{P}(x)$ и вектором скорости $\mathbf{u} = \{u^\mu(x)\}$, где $x = \{x^\alpha\}$. Движение идеальной жидкости или идеального газа в пространстве определяется следующей замкнутой системой из пяти нелинейных уравнений для пяти переменных (см., например, [173])

$$\rho \dot{\mathbf{u}} + \rho(\mathbf{u} \nabla) \mathbf{u} = -\nabla \mathcal{P} + \mathbf{f}, \quad (23.140)$$

$$\dot{\rho} + \operatorname{div}(\rho \mathbf{u}) = 0, \quad (23.141)$$

$$\mathcal{P} = \mathcal{P}(\rho), \quad (23.142)$$

где точка обозначает дифференцирование по времени и ∇ – градиент. Уравнение (23.140) называется уравнением Эйлера (23.131) и представляет собой второй закон Ньютона для элемента объема жидкости. Здесь $\mathbf{f}(x)$ – плотность внешних сил. Например, в гравитационном поле $\mathbf{f} = -\rho \nabla \varphi$, где $\varphi(x)$ – потенциал гравитационного поля. В дальнейшем мы ограничимся только этим случаем. Уравнение (23.141) является уравнением непрерывности (23.130). Последнее уравнение (23.142) является уравнением состояния жидкости, которое характеризует саму жидкость и считается заданным. Здесь мы предполагаем, что давление жидкости зависит только от ее плотности, т.е. она является баротропной.

Для несжимаемой жидкости $\rho = \text{const}$ и уравнение непрерывности принимает вид $\operatorname{div} \mathbf{u} = 0$.

Уравнения (23.140)–(23.142) записаны в стандартном для гидродинамики виде, где не делается различие между верхними и нижними индексами.

Если уравнение состояния задано, то давление $\mathcal{P}$ (или плотность ρ) можно исключить из системы уравнений движения. Для этого заметим, что

$$\nabla \mathcal{P} = \frac{d\mathcal{P}}{d\rho} \nabla \rho = c^2 \nabla \rho,$$

где введена скорость звука $c(\rho)$

$$c^2 := \frac{d\mathcal{P}}{d\rho}. \quad (23.143)$$

Для обычных жидкостей с увеличением плотности давление увеличивается, и поэтому $c^2 > 0$. Тогда уравнение Эйлера можно переписать в виде

$$\dot{\mathbf{u}} + (\mathbf{u} \nabla) \mathbf{u} = -c^2 \frac{\nabla \rho}{\rho} - \nabla \varphi, \quad (23.144)$$

где скорость звука $c = c(\rho)$ рассматривается, как заданная функция от плотности жидкости. Уравнение Эйлера (23.144) вместе с уравнением непрерывности (23.141)

однозначно описывают движение такой жидкости при заданных начальных и граничных условиях.

Воспользовавшись тождеством

$$\frac{1}{2}\nabla \mathbf{u}^2 = [\mathbf{u}, \text{rot } \mathbf{u}] + (\mathbf{u}\nabla)\mathbf{u},$$

где квадратные скобки обозначают векторное произведение, уравнение Эйлера (23.144) можно переписать в виде

$$\dot{\mathbf{u}} - [\mathbf{u}, \text{rot } \mathbf{u}] = -\nabla \left(h + \frac{\mathbf{u}^2}{2} + \varphi \right),$$

где введена *энтальпия* жидкости

$$h(\mathcal{P}) := \int_0^{\mathcal{P}} \frac{d\mathcal{P}'}{\rho(\mathcal{P}')}.$$

Предположим, что движение жидкости является безвихревым:

$$\text{rot } \mathbf{u} = 0 \quad \Leftrightarrow \quad \partial_\mu u_\nu - \partial_\nu u_\mu = 0. \quad (23.145)$$

На геометрическом языке это означает, что 1-форма $dx^\mu u_\mu$ является замкнутой на пространственном сечении $x^0 = \text{const}$. Тогда локально (во всем $\mathbb{R}^3$) существует потенциальное поле $\psi(x)$ (потенциал) такое, что

$$u_\mu = -\partial_\mu \psi. \quad (23.146)$$

Для безвихревой жидкости уравнение Эйлера эквивалентно *уравнению Бернулли*

$$-\dot{\psi} + h + \frac{(\nabla \psi)^2}{2} + \varphi = F(t), \quad (23.147)$$

где $F(t)$ – произвольная функция времени. Поскольку потенциал ψ определен с точностью до добавления произвольной функции времени, то, без ограничения общности, положим $F(t) = 0$.

Теперь можно приступить к изучению фононов в нерелятивистской гидродинамике. Допустим, что нам известно точное решение уравнений гидродинамики $\rho_0(x)$, $\mathcal{P}_0(x)$ и $\mathbf{u}_0(x) = -\nabla \psi_0$. Получим уравнение, описывающее распространение акустических возбудений (фононов) вблизи этого решения. Пусть

$$\begin{aligned} \rho &\approx \rho_0 + \epsilon \rho_1, \\ \mathcal{P} &\approx \mathcal{P}_0 + \epsilon \mathcal{P}_1, \\ u^\mu &\approx u_0^\mu + \epsilon u_1^\mu, \\ \psi &\approx \psi_0 + \epsilon \psi_1. \end{aligned} \quad (23.148)$$

где $\epsilon \ll 1$ – малый параметр разложения. При этом мы считаем внешние силы заданными $\varphi = \varphi_0$.

В дальнейшем мы будем использовать обозначения, принятые в дифференциальной геометрии, и различать верхние и нижние индексы. Пространственные индексы в декартовой системе координат поднимаются и опускаются с помощью метрики $\eta_{\mu\nu} = -\delta_{\mu\nu}$ и ее обратной, которая отличается от евклидовой метрики знаком. В наших обозначениях $u_\mu = \partial_\mu \psi$, $u^\mu = \eta^{\mu\nu} \partial_\nu \psi = -\partial_\mu \psi$.

Уравнение Бернулли (23.147) в нулевом и первом порядке по ϵ имеет вид

$$\epsilon^0 : \quad -\partial_0\psi_0 + h_0 - \frac{1}{2}\eta^{\mu\nu}\partial_\mu\psi_0\partial_\nu\psi_0 + \varphi_0 = 0, \quad (23.149)$$

$$\epsilon^1 : \quad -\partial_0\psi_1 + \frac{\mathcal{P}_1}{\rho_0} - u_0^\mu\partial_\mu\psi_1 = 0, \quad (23.150)$$

где учтено разложение для энтальпии

$$h(\mathcal{P}) \simeq h(\mathcal{P}_0) + \epsilon \frac{\mathcal{P}_1}{\rho_0}.$$

Учтем, что

$$\rho_1 = \frac{d\rho}{d\mathcal{P}}\mathcal{P}_1 = \frac{\mathcal{P}_1}{c^2},$$

и найдем поправку к плотности ρ_1 из уравнения (23.150)

$$\rho_1 = \frac{\rho_0}{c^2}(\partial_0\psi_1 + u_0^\mu\partial_\mu\psi_1). \quad (23.151)$$

Уравнение непрерывности в нулевом и первом порядке по ϵ имеет вид

$$\epsilon^0 : \quad \partial_0\rho_0 + \partial_\mu(\rho_0 u_0^\mu) = 0, \quad (23.152)$$

$$\epsilon^1 : \quad \partial_1\rho_1 + \partial_\mu(\rho_0 u_1^\mu + \rho_1 u_0^\mu) = 0. \quad (23.153)$$

Подставим во второе уравнение решение для поправки к плотности (23.151). В результате получим уравнение для поправки к потенциалу скорости:

$$\partial_0 \left[\frac{\rho_0}{c^2}(\partial_0\psi_1 + u_0^\mu\partial_\mu\psi_1) \right] + \partial_\mu \left[\rho_0\partial^\mu\psi_1 + \frac{\rho_0}{c^2}u_0^\mu(\partial_0\psi_1 + u_0^\nu\partial_\nu\psi_1) \right] = 0. \quad (23.154)$$

Это волновое уравнение для $\psi_1(x)$ полностью определяет распространение акустических колебаний в движущейся жидкости, описываемой плотностью $\rho_0(x)$ и полем скоростей $u_0^\mu(x)$, с заданным уравнением состояния $s = s(\rho)$. В случае, когда скорость жидкости равна нулю, $u_0 = 0$, а плотность ρ_0 и величина s постоянны, уравнение (23.154) сводится к уравнению Даламбера, которое описывает распространение акустических возмущений со скоростью s . Это оправдывает введенное выше понятие скорости звука (23.143).

Если решение для $\psi_1(x)$ известно, то поправка к плотности ρ_1 однозначно определяется формулой (23.151).

Уравнение (23.154), как легко проверить, можно записать в матричных обозначениях

$$\partial_\alpha(f^{\alpha\beta}\partial_\beta\psi_1) = 0,$$

где

$$f^{\alpha\beta} := \frac{\rho_0}{c^2} \begin{pmatrix} 1 & u_0^\nu \\ u_0^\mu & c^2\eta^{\mu\nu} + u_0^\mu u_0^\nu \end{pmatrix}$$

Введем метрику в галилеевом пространстве

$$g_{\alpha\beta} := \frac{\rho_0}{c} \begin{pmatrix} c^2 + u_0^\mu u_{0\mu} & -u_{0\nu} \\ -u_{0\mu} & \eta_{\mu\nu} \end{pmatrix} \quad (23.155)$$

и ее обратную

$$g^{\alpha\beta} = \frac{1}{\rho_0 c} \begin{pmatrix} 1 & u_0^\nu \\ u_0^\mu & c^2 \eta^{\mu\nu} + u_0^\mu u_0^\nu \end{pmatrix} \quad (23.156)$$

Интервал, соответствующий метрике (23.155), можно записать в виде

$$ds^2 = \frac{\rho_0}{c} [c^2 dt^2 + \eta_{\mu\nu} (dx^\mu - u_0^\mu dt)(dx^\nu - u_0^\nu dt)]. \quad (23.157)$$

“Эффективная” метрика (23.155) имеет лоренцеву сигнатуру $(+---)$, и ее определитель равен

$$g := \det g_{\alpha\beta} = -\frac{\rho_0^4}{c^2}.$$

Сравнение метрики (23.155) с АДМ параметризацией произвольной псевдоримановой метрики (24.5) дает следующее выражение для функций хода и сдвига:

$$N = \sqrt{\rho_0 c}, \quad N^\mu = -\frac{\rho_0 u_0^\mu}{c}.$$

Вернемся к уравнению для фононов. Обратная метрика $g^{\alpha\beta}$ отличается от матрицы $f^{\alpha\beta}$ простым множителем

$$g^{\alpha\beta} = \frac{c}{\rho_0^2} f^{\alpha\beta}.$$

Теперь уравнение для акустических фононов можно переписать в инвариантном относительно общих преобразований координат виде

$$\frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} g^{\alpha\beta} \partial_\beta \psi) = 0, \quad (23.158)$$

где мы, для простоты обозначений, отбросили индекс у поправки к потенциалу скорости.

Таким образом, распространение фононов в движущейся жидкости описывается инвариантным волновым уравнением в четырехмерном пространстве-времени с нетривиальной метрикой лоренцевой сигнатуры (23.155). Эта метрика определяется плотностью ρ_0 , скоростью звука c и полем скоростей $\mathbf{u}_0$, которые удовлетворяют исходным уравнениям (23.149), (23.153). Подчеркнем, что движение самой жидкости происходит в плоском галилеевом пространстве-времени, а распространение акустических возмущений в этой движущейся жидкости описывается волновым уравнением на псевдоримановом пространстве-времени с нетривиальной “эффективной” метрикой.

Мы считаем, что уравнение состояния жидкости задано, и, следовательно, задана скорость звука в жидкости, как функция плотности. Тогда эффективная метрика определяется четырьмя функциями $\rho_0(x)$ и $\mathbf{u}_0(x)$, которые удовлетворяют уравнениям гидродинамики. При постановке задачи Коши для однозначного определения этих функций необходимо задать четыре произвольные функции в качестве начальных условий. В общей теории относительности метрика удовлетворяет уравнениям Эйнштейна и имеет две распространяющиеся степени свободы. При постановке задачи Коши для уравнений Эйнштейна также необходимо задать четыре произвольные функции на пространственноподобном сечении: по две на каждую степень свободы, т.к. уравнения движения второго порядка.

Нулевая компонента метрики g_{00} в (23.155) меняет знак в тех точках пространства-времени, где течение жидкости становится сверхзвуковым: $c^2 = \mathbf{u}^2 := -u^\mu u_\mu$. Эти поверхности в пространстве соответствуют горизонтам черных дыр. Действительно, поскольку скорость фононов ограничена скоростью звука в жидкости, то они не могут покинуть область сверхзвукового течения. Следовательно, в быстро текущей жидкости для фононов могут образовываться аналоги черных дыр в общей теории относительности. На рис.23.1 показана качественная картинка образования

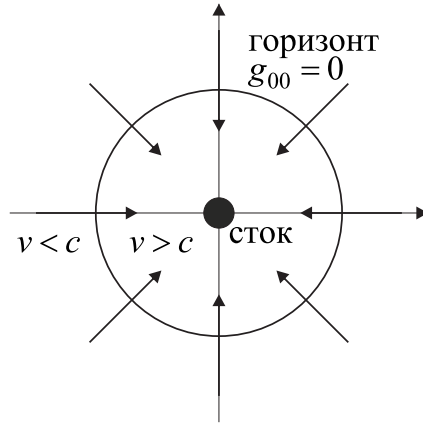


Рис. 23.1: Черная дыра для акустических фононов. Горизонт определяется равенством модуля скорости жидкости $v := \sqrt{-u_0^\mu u_{0\mu}}$ скорости звука c .

черной дыры в жидкости. Представим, что все трехмерное пространство $\mathbb{R}^3$ заполнено несжимаемой жидкостью, и в центре декартовой системы координат находится точечный сток, в который жидкость засасывается. Предположим, что жидкость на бесконечности покоится, и ее движение радиально. Поскольку скорость жидкости стремится к бесконечности при приближении к стоку, то существует сфера некоторого радиуса, на которой скорость жидкости равна скорости звука. Эта сфера называется горизонтом. Если фонов испущен в некоторой точке, лежащей вне горизонта, то он может либо попасть в сток, либо уйти на бесконечность. Если же фонов испущен из точки, лежащей внутри горизонта, то он неминуемо попадет в сток, т.к. его скорости недостаточно для прохождения через горизонт. Описанная ситуация является аналогом черных дыр в общей теории относительности. Надо только скорость звука заменить на скорость света. Аналогом фононов является электромагнитное излучение – фотоны.

23.13 Выбор системы координат

Уравнения общей теории относительности ковариантны относительно общих преобразований координат. Эту свободу можно использовать для выбора подходящей системы отсчета, которая может упростить уравнения Эйнштейна. В настоящем разделе будут описаны несколько широко распространенных способа фиксирования системы координат.

23.13.1 Сопутствующая система координат

Рассмотрим уравнения Эйнштейна

$$\kappa \left(R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R \right) = -\frac{1}{2} T_{\alpha\beta} \quad (23.159)$$

для сплошной среды с тензором энергии-импульса (см. раздел 23.12)

$$T_{\alpha\beta} = (\mathcal{E} + \mathcal{P}) u^\alpha u^\beta - \mathcal{P} g^{\alpha\beta}. \quad (23.160)$$

Для получения замкнутой системы уравнений уравнения Эйнштейна необходимо дополнить законом сохранения (уравнениями релятивистской гидродинамики)

$$\nabla_\beta T_{\alpha}^{\beta} = 0 \quad (23.161)$$

и уравнением состояния среды

$$\mathcal{P} = \mathcal{P}(\mathcal{E}), \quad (23.162)$$

предполагая среду баротропной. Система уравнений (23.159), (23.161) и (23.162) образуют полную систему для неизвестных функций: $g_{\alpha\beta}$, u^α , $\mathcal{P}$ и $\mathcal{E}$. Нетрудно проверить, что число уравнений равно числу неизвестных (напомним, что на вектор скорости u наложено условие $u^2 = 1$ и число его независимых компонент на единицу меньше размерности пространства-времени). Если $n = 4$, то число уравнений и число неизвестных равно пятнадцати.

Поскольку плотность энергии $\mathcal{E}(x)$ и давление $\mathcal{P}(x)$ являются скалярными полями, то уравнение состояния (23.162) является корректным.

По построению, все уравнения ковариантны. Поэтому преобразования координат можно использовать для упрощения системы уравнений. Обычно преобразования координат используют для фиксирования части компонент метрики. Однако для системы уравнений (23.159), (23.161) и (23.162) существует другая естественная возможность. Если размерность пространства-времени равна n , то в нашем распоряжении имеется n функций, которых достаточно для фиксирования векторного поля скорости. Тем самым число неизвестных функций уменьшится, и задача упростится. Такой подход часто используется в космологии.

Опишем этот способ задания системы координат. Рассмотрим псевдориманово многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, с метрикой g лоренцевой сигнатуры. Пусть на нем задано произвольное достаточно гладкое времениподобное векторное поле $u = u^\alpha(x) \partial_\alpha$, всюду отличное от нуля, $u^2 \neq 0$. Не ограничивая общности, будем считать, что

$$u^2 := g_{\alpha\beta} u^\alpha u^\beta = 1.$$

В противном случае можно просто заменить $u^\alpha \mapsto u^\alpha / u^2$.

Пример 23.13.1. Пусть пространство-время $\mathbb{M}$ заполнено сплошной средой. Тогда задано векторное поле скорости каждой точки среды u , для которого $u^2 = 1$. $\square$

Согласно общему рассмотрению в разделе 2.6.5 в окрестности произвольной точки существует такая система координат, в которой все компоненты векторного поля, кроме одной, равны нулю. Для единичного времениподобного векторного поля нетривиальная компонента равна единице:

$$\{u^\alpha\} = \{1, 0, \dots, 0\}. \quad (23.163)$$

Определение. Система координат, в которой для единичного времениподобного векторного поля выполнено условие (23.163), называется *сопутствующей векторному полю u* . $\square$

В этой системе координат ковариантная производная векторного поля u равна символам Кристоффеля с одним временным индексом,

$$\nabla_\alpha u^\beta = \partial_\alpha u^\beta + \Gamma_{\alpha\gamma}^\beta u^\gamma = \Gamma_{\alpha 0}^\beta.$$

В сопутствующей системе координат $\partial_\alpha u^\beta = 0$, поэтому при бесконечно малых преобразованиях координат вариация формы компонент вектора (2.100) и ковектора (2.101) содержит только одно слагаемое. Аналогично преобразуются компоненты произвольного тензора. Отсюда вытекает, что производная Ли от любого тензора Y типа (r, s) вдоль векторного поля u совпадает с производной по времени x^0 :

$$\mathcal{L}_u Y^{\alpha_1 \dots \alpha_r}_{\beta_1 \dots \beta_s} = \partial_0 Y^{\alpha_1 \dots \alpha_r}_{\beta_1 \dots \beta_s}$$

и не зависит от связности, как, впрочем, любая производная Ли.

С физической точки зрения сопутствующую систему координат можно представить следующим образом. Допустим, что некоторая среда заполняет все пространство-время. Тогда с каждой точкой среды связана мировая линия $x(s)$ (линия тока). Мы предполагаем, что касательные векторы к мировым линиям образуют достаточно гладкое времениподобное векторное поле (вектор скорости)

$$u = u^\alpha \partial_\alpha, \quad u^\alpha := \frac{dx^\alpha}{ds}, \quad ds := \sqrt{g_{\alpha\beta} dx^\alpha dx^\beta},$$

на многообразии $\mathbb{M}$. Выберем произвольное сечение $\mathbb{S}$, которое пересекает все линии тока один раз, и зададим произвольную систему координат x^μ , $\mu = 1, \dots, n-1$, на $\mathbb{S}$. Это сечение совсем не обязано быть пространственноподобным. Тогда сопутствующими координатами произвольной точки $y \in \mathbb{M}$ является набор чисел $\{x^0 := s, x^\mu\}$, где x^μ – координаты точки пересечения поверхности $\mathbb{S}$ с кривой $x(s)$, проходящей через точку y . Для определенности будем считать, что каждая линия тока пересекает поверхность $\mathbb{S}$ при $s = 0$.

Замечание. В предыдущем разделе мы установили, что пылевидная материя движется вдоль экстремалей (23.137). Это значит, что в общем случае при наличии давления или других негравитационных сил линии тока среды отличаются от экстремалей. $\square$

Если производная Ли от некоторого тензора вдоль векторного поля скорости u равна нулю, то в сопутствующей системе координат компоненты этого тензора могут зависеть только от координат x^μ , $\mu = 1, \dots, n-1$ на сечении $\mathbb{S}$. Это значит, что соответствующий тензор жестко связан с движущейся средой и движется вместе с ней. В дальнейшем мы будем предполагать, что все сечения $s = \text{const}$ и, в частности, сечение $\mathbb{S}$ пространственноподобны.

Сопутствующая векторному полю система координат определена неоднозначно. Действительно, совершив преобразование координат $x^\alpha \mapsto x^{\alpha'}(x)$. Тогда компоненты скорости преобразуются по тензорному закону:

$$u^\alpha \mapsto u^{\alpha'} := \frac{\partial x^{\alpha'}}{\partial x^\alpha} u^\alpha.$$

Если до и после преобразования координат система координат является сопутствующей, то функции преобразования координат должны удовлетворять следующей системе уравнений

$$1 = \frac{\partial x^{0'}}{\partial x^0}, \quad 0 = \frac{\partial x^{\mu'}}{\partial x^0}.$$

Общее решение данной системы уравнений имеет вид

$$x^0 \mapsto x^{0'} = x^0 + f(\mathbf{x}), \quad x^\mu \mapsto x^{\mu'} = x^\mu + f^\mu(\mathbf{x}), \quad (23.164)$$

где f, f^μ – n произвольных функций координат на сечении $\mathbb{S} \subset \mathbb{M}$ и $\mathbf{x} = \{x^\mu\}$. Функция f соответствует произволу в выборе сечения $x^0 = \text{const}$, и функции f^μ – свободе в выборе координат x^μ на данных сечениях.

Таким образом мы устранили n неизвестных функций в полной системе уравнений (23.159), (23.161) и (23.162). В этой системе координат тензор энергии-импульса (23.160) принимает вид

$$T_{\text{м}}^{00} = (\mathcal{E} + \mathcal{P}) - \mathcal{P}g^{00}, \quad T_{\text{м}}^{0\mu} = -\mathcal{P}g^{0\mu}, \quad T_{\text{м}}^{\mu\nu} = -\mathcal{P}g^{\mu\nu}.$$

В общем случае ни он, ни тензор энергии-импульса с одним опущенным индексом индексом $T_{\text{м}\alpha}^\beta$ не будут диагональны.

Теперь определим другую систему координат, которую также естественно назвать сопутствующей.

Если задано единичное времениподобное векторное поле u , то в каждой точке пространства-времени $x \in \mathbb{M}$ в касательном пространстве $T_x(\mathbb{M})$ его можно дополнить $n - 1$ линейно независимыми векторами e_μ , $\mu = 1, \dots, n - 1$, которые перпендикулярны вектору u . Тогда совокупность векторов $\{u, e_\mu\}$ образует в каждой точке репер. Ясно, что векторы e_μ пространственноподобны, и их можно выбрать достаточно гладкими. Тогда они задают $(n - 1)$ -мерное распределение векторных полей на $\mathbb{M}$ (см. раздел 2.11). Согласно теореме Фробениуса для этого распределения существуют интегральные подмногообразия тогда и только тогда, когда векторные поля e_μ находятся в инволюции. В общем случае это не так (это зависит от метрики). Отсюда следует, что остаточного произвола в выборе сопутствующей системы координат (23.164) недостаточно для того, чтобы выбрать секущую поверхность $\mathbb{S}$ таким образом, чтобы вектор u был к ней всюду ортогонален.

При определении системы координат, сопутствующей векторному полю, за основу взяты мировые линии точек среды. Теперь мы введем другое понятие сопутствующей системы координат, где за основу определения будет взята система гиперповерхностей, а не векторное поле.

Рассмотрим систему локальных координат x^α на $\mathbb{M}$, где координата x^0 является временем, и все сечения $x^0 = \text{const}$ пространственноподобны. По предположению, на $\mathbb{M}$ задана метрика лоренцевой сигнатуры, для которой мы будем использовать АДМ параметризацию (24.5). Выберем базис n, e_μ в касательном пространстве, состоящий из пространственных векторов $e_\mu = e_\mu^\alpha \partial_\alpha = \partial_\mu$ (символ $e_\mu^\alpha := \delta_\mu^\alpha$ введен для удобства последующих выкладок), касательных к гиперповерхностям $x^0 = \text{const}$, и векторное поле

$$n := \frac{1}{N}(\partial_0 - N^\mu \partial_\mu), \quad (23.165)$$

где N – функция хода и N^μ – функции сдвига. Это векторное поле перпендикулярно пространственным гиперповерхностям:

$$(n, e_\mu) := n^\alpha e_\mu^\beta g_{\alpha\beta} = n^0 e_\mu^\nu g_{0\nu} + n^\rho e_\mu^\nu g_{\rho\nu} = 0,$$

и имеет единичную длину, $n^2 = 1$.

В общем случае коммутатор векторных полей $[n, e_\mu]$ отличен от нуля. Поэтому пространственные координаты x^μ нельзя дополнить времениподобной координатой $x^{0'}$ так, чтобы вектор n был касателен к соответствующей координатной линии: $n = \partial_{0'}$. То есть базис n, e_μ является неголономным.

Нормальному вектору соответствует ортонормальная 1-форма

$$n = dx^\alpha n_\alpha = dx^0 N, \quad (23.166)$$

где $n_\alpha := n^\beta g_{\beta\alpha}$, для которой мы будем использовать то же обозначение.

Для векторов и 1-форм справедливы разложения на перпендикулярную и касательные составляющие:

$$X^\alpha = X^\perp n^\alpha + \tilde{X}^\mu e_\mu{}^\alpha, \quad X_\alpha = X_\perp n_\alpha + \tilde{X}_\mu e^\mu{}_\alpha,$$

где

$$\begin{aligned} X^\perp &= X^0 N, & \tilde{X}^\mu &= X^0 N^\mu + X^\mu, \\ X_\perp &= \frac{1}{N}(X_0 - N^\mu X_\mu), & \tilde{X}_\mu &= X_\mu. \end{aligned}$$

Поскольку $X_0 := X^\alpha g_{\alpha 0}$ и $X_\mu := X^\alpha g_{\alpha\mu}$, то нетрудно проверить, что $X_\perp = X^\perp$ и $\tilde{X}_\mu = \tilde{X}^\nu g_{\nu\mu}$.

Аналогично раскладываются тензоры произвольного ранга.

Определение. Система координат называется *сопутствующей*, если в этой системе координат скорость $\{u^\alpha\}$ каждой точки сплошной среды имеет вид

$$\{u^\alpha\} = \{n^\alpha\} = \left\{ \sqrt{g^{00}}, \frac{g^{0\mu}}{\sqrt{g^{00}}} \right\} = \left\{ \frac{1}{N}, -\frac{N^\mu}{N} \right\}. \quad \square \quad (23.167)$$

То есть единичный нормальный вектор к пространственным сечениям совпадает с векторным полем скорости.

В правой части равенства (23.167), определяющего сопутствующую систему координат, стоят определенные компоненты метрики, которые не образуют компонент вектора. Следовательно, равенство (23.167) нековариантно и действительно фиксирует систему координат.

Название сопутствующая система координат оправдано следующим образом. Поскольку вектор нормали n к пространственному сечению $x^0 = \text{const}$ совпадает с вектором скорости u , то репер n, e_μ привязан к среде и движется вместе с ней.

Если метрика имеет блочно диагональный вид 4.20, т.е. $N^\mu = 0$ и $N = 1$, и, следовательно, выбрана временная калибровка (см. следующий раздел), то вектор скорости материи в каждой точке пространства-времени имеет только временную составляющую $\{u^\alpha\} = (1, 0, 0, 0)$. В этом случае сопутствующая система координат совпадает с системой координат, сопутствующей векторному полю скорости u , которая была введена в начале раздела. Это значит, что в сопутствующей системе координат каждая точка материи покоится. Другими словами, система координат движется вместе с материей. При этом каждая точка среды движется по времениподобной экстремали (геодезической).

При наличии сил негравитационного происхождения, например, давления, точки материи могут двигаться не по экстремалиям, и метрика в общем случае не будет

блочно-диагональна. Математически это означает, что в общем случае нельзя одновременно удовлетворить условию (23.167) и условию блочной диагональности метрики (4.20). Поэтому данные выше два определения сопутствующей системы координат не эквивалентны.

Обозначим тензорные индексы по отношению к базису $\{e_a\} := \{n, e_\mu\}$ латинскими буквами $a, b, \dots$. Тогда они принимают значения $\{a\} = \{\perp, \mu\} = (\perp, 1, \dots, n-1)$. Отличие этого базиса от координатного базиса в касательном пространстве заключается в том, что он не голоном, т.е. в общем случае не существует такой системы координат $y^a(x)$, в которой были бы выполнены условия:

$$\frac{\partial x^\alpha}{\partial y^0} = n^\alpha, \quad \frac{\partial x^\nu}{\partial y^\mu} = \delta_\mu^\nu$$

(см. раздел 6.9). В базисе e_a метрика имеет блочно диагональный вид

$$g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix}, \quad (23.168)$$

поскольку базисные векторы n были выбраны единичными и перпендикулярными к гиперповерхностям. В каждой точке пространства-времени векторы скорости (23.167) в базисе n, ∂_μ имеют только одну отличную от нуля компоненту,

$$\{u^a\} = (1, 0, 0, 0).$$

Это значит, что в рассматриваемом базисе материя покоится, что оправдывает название сопутствующая.

Понятие сопутствующей системы координат полезно, т.к. позволяет исключить из тензора энергии-импульса сплошной среды компоненты скорости u^α , заменив их на компоненты метрики с помощью определения (23.167). Не трудно проверить, что тензор энергии-импульса материи с одним контравариантным и одним ковариантным индексом в сопутствующей системе координат (23.167), и, следовательно, относительно базиса n, ∂_μ является диагональным. Например, в четырехмерном пространстве-времени положим

$$T_{ma}{}^b = \begin{pmatrix} \mathcal{E} & 0 & 0 & 0 \\ 0 & -\mathcal{P} & 0 & 0 \\ 0 & 0 & -\mathcal{P} & 0 \\ 0 & 0 & 0 & -\mathcal{P} \end{pmatrix}. \quad (23.169)$$

Тогда в координатном базисе $\{\partial_\alpha\}$ его контравариантные компоненты имеют вид

$$T_m{}^{00} = \mathcal{E} \frac{1}{N^2}, \quad T_m{}^{0\mu} = T_m{}^{\mu 0} = -\mathcal{E} \frac{N^\mu}{N^2}, \quad T_m{}^{\mu\nu} = \mathcal{E} \frac{N^\mu N^\nu}{N^2} - \mathcal{P} \hat{g}^{\mu\nu}. \quad (23.170)$$

Тем самым компоненты контравариантного тензора энергии-импульса не зависят от векторного поля скоростей сплошной среды.

Предложение 23.13.1. *Сопутствующая система координат (23.167) определена по крайней мере с точностью до масштабных преобразований (растяжки, гомотетии) пространственных координат:*

$$x^0 \mapsto x'^0 = x^0, \quad x^\mu \mapsto x'^\mu = k(t)x^\mu, \quad (23.171)$$

где $k(t) \neq 0$ – достаточно гладкая функция времени $t := x^0$.

Доказательство. При преобразовании координат (23.171) координатный базис преобразуется по правилам:

$$\begin{aligned} \partial_0 &= \partial'_0 + \dot{k}x^\mu \partial'_\mu, & dx^0 &= dx'^0, \\ \partial_\mu &= k\partial'_\mu, & dx^\mu &= \frac{dx'^\mu}{k} - \frac{x'^\mu \dot{k}dt}{k^2}, \end{aligned}$$

где $\dot{k} := dk/dt$. Тогда в новой системе координат компоненты метрики примут вид

$$\begin{aligned} N' &= N, \\ N'_\mu &= \frac{N_\mu}{k} - \frac{\dot{k}x_\mu}{k^2}, & N'^\mu &= kN^\mu - \dot{k}x^\mu, \\ g'_{\mu\nu} &= \frac{g_{\mu\nu}}{k^2}, & \hat{g}'^{\mu\nu} &= k^2 \hat{g}^{\mu\nu}. \end{aligned}$$

Теперь из формул (23.170) следует, что плотность энергии и давление в сопутствующей системе координат не меняются при масштабном преобразовании (23.171)

$$\mathcal{E}'(x') = \mathcal{E}(x), \quad \mathcal{P}'(x') = \mathcal{P}(x).$$

При этом компоненты вектора скорости преобразуются по-правилам

$$u'^0 = u^0, \quad u'^\mu = ku^\mu + \dot{k}u^0 x^\mu.$$

То есть так же, как и правая часть определения (23.167). $\square$

Инвариантность плотности энергии и давления относительно масштабного преобразования следовало ожидать, т.к. плотность энергии и давление являются скалярными полями и инвариантны относительно любых преобразований координат пространства-времени x^α и, в частности, растяжений (23.171). Нетривиальность проведенного рассмотрения заключается в том, что в правой части равенства (23.167), определяющего сопутствующую систему координат, стоят определенные компоненты метрики, которые в общем случае не совпадают с компонентами никакого вектора.

Уравнения релятивистской гидродинамики (23.161), которые были выписаны в начале раздела, можно записать в сопутствующей системе координат (23.167). Прямые вычисления приводят к следующей системе уравнений:

$$\begin{aligned} \dot{\mathcal{E}} - N^\mu \partial_\mu \mathcal{E} - NK(\mathcal{E} + \mathcal{P}) &= 0, \\ \partial_\mu \mathcal{P} + \frac{\partial_\mu N}{N}(\mathcal{E} + \mathcal{P}) &= 0, \end{aligned} \tag{23.172}$$

где точка обозначает дифференцирование по времени $\dot{\mathcal{E}} := \partial \mathcal{E} / \partial t$ и

$$K = \frac{1}{2N}(2\hat{\nabla}^\mu N_\mu - \hat{g}^{\mu\nu} \dot{g}_{\mu\nu}) \tag{23.173}$$

– внешняя скалярная кривизна пространственной гиперповерхности (24.45).

23.13.2 Временная калибровка

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, на котором задана метрика лоренцевой сигнатуры $g_{\alpha\beta}(x)$, $\text{sign } g_{\alpha\beta} = (+ - \dots -)$.

Определение. Система координат, в которой метрика имеет блочно диагональный вид (4.20)

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix}, \quad (23.174)$$

где $g_{\mu\nu}$ – отрицательно определенная риманова метрика на пространственноподобных сечениях $x^0 = \text{const}$, называется *временной калибровкой*. Эту систему координат называют также *синхронной*, *гауссовой* или *полугеодезической*. $\square$

В синхронной системе отсчета координата x^0 является временем и явно выделена. Напомним, что греческие буквы из начала алфавита пробегают все значения индексов $(\alpha, \beta, \dots = 0, 1, \dots, n-1)$, а из середины – только пространственные $(\mu, \nu, \dots = 1, 2, \dots, n-1)$.

При переходе в синхронную систему отсчета n произвольных функций, параметризующих диффеоморфизмы, используются для фиксирования n компонент метрики:

$$g_{00} = 1, \quad g_{0\mu} = 0.$$

В АДМ параметризации метрики (см. раздел 24.2) временная калибровка соответствует условиям $N = 1$, $N_\mu = 0$.

Замечание. Названия гауссова или полугеодезическая система координат распространены в математической литературе, когда рассматриваются римановы пространства с положительно определенной метрикой. В физической литературе, где преимущественно рассматриваются многообразия с метрикой лоренцевой сигнатуры, чаще употребляют термины временная калибровка или синхронная система координат, потому что в этой системе отсчета координата x^0 действительно играет роль наблюдаемого времени. $\square$

Название синхронная система координат для метрики (23.174) оправдано следующим обстоятельством.

Синхронизация часов

Рассмотрим произвольную систему координат. В общем случае интервал между двумя близкими событиями $\{x^\alpha\}$ и $\{x^\alpha + dx^\alpha\}$ имеет вид

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta.$$

Предположим, что координата x^0 является наблюдаемым временем, т.е. $g_{00} > 0$, и все сечения $x^0 = \text{const}$ пространственноподобны. Если два события C и D произошли в данной системе координат в одной и той же точке пространства, то они имеют координаты C = $\{x_C^0, x^\mu\}$ и D = $\{x_D^0, x^\mu\}$. При этом данные события разделены интервалом собственного времени

$$\Delta s = \int_{x_C^0}^{x_D^0} dx^0 \sqrt{g_{00}}. \quad (23.175)$$

Этот интеграл равен длине времениподобной кривой

$$x^0 = x_C^0 + (x_D^0 - x_C^0)\tau, \quad x^\mu = \text{const}, \quad \tau \in [0, 1],$$

соединяющей события C и D. Конечно, в другой системе координат эти события могут произойти не только в разное наблюдаемое время, но и в разных точках пространства.

Таким образом, если два события, произошедшие в одной точке пространства в данной системе координат, разделены наблюдаемым временем $x_D^0 - x_C^0$, то они разделены интервалом собственного времени (23.175). При этом нулевая компонента метрики g_{00} определяет различие собственного и наблюдаемого времени для событий, произошедших в одной точке.

Теперь определим понятие одновременности для событий, которые произошли в двух разных, но близких точках пространства в данной фиксированной системе координат. Пусть событие А имеет пространственные координаты x^μ , а событие В – близкие координаты $x^\mu + dx^\mu$. На рис.23.2 сплошными линиями показаны временные оси, проходящие через точки А и В. Возникает следующий вопрос одновременности. Допустим, что событие А имеет координаты $\{x^0, x^\mu\}$. Какова временная координата $x^0 + \Delta x^0$ события, произошедшего в точке В, которое можно назвать одновременным с событием А?

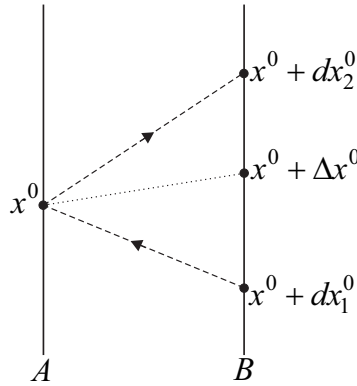


Рис. 23.2: Одновременность близких событий А и В.

Чтобы определить одновременность, испустим свет в точке В в некоторый момент времени $x^0 + dx_1^0$ (величина dx_1^0 отрицательна). Как только свет попадет в точку А, сразу отразим его. Допустим, что свет вернулся в точку В в момент времени $x^0 + dx_2^0$. Поскольку для света $ds^2 = 0$, то изменение наблюдаемого времени в обоих случаях должно удовлетворять уравнению

$$g_{00}(dx_{1,2}^0)^2 + 2g_{0\mu}dx_{1,2}^0dx^\mu + g_{\mu\nu}dx^\mu dx^\nu = 0.$$

Это квадратное уравнение имеет два решения:

$$\begin{aligned} dx_1^0 &= \frac{1}{g_{00}} \left[-g_{0\mu}dx^\mu - \sqrt{(g_{0\mu}g_{0\nu} - g_{\mu\nu}g_{00})dx^\mu dx^\nu} \right], \\ dx_2^0 &= \frac{1}{g_{00}} \left[-g_{0\mu}dx^\mu + \sqrt{(g_{0\mu}g_{0\nu} - g_{\mu\nu}g_{00})dx^\mu dx^\nu} \right]. \end{aligned}$$

Поскольку мы предположили, что $g_{00} > 0$ и метрика $g_{\mu\nu}$ отрицательно определена, то отсюда вытекает, что $dx_2^0 > 0$, а $dx_1^0 < 0$.

Определение. Назовем событие в точке В *одновременным* событию А = $\{x^0, x^\mu\}$, если его временная координата равна $x^0 + \Delta x^0$, где

$$\Delta x^0 := \frac{dx_1^0 + dx_2^0}{2} = -\frac{g_{0\mu}dx^\mu}{g_{00}},$$

т.е. лежит посередине между $x^0 + dx_2^0$ и $x^0 + dx_1^0$. □

Таким образом можно синхронизировать часы, расположенные в различных, но близких точках пространства. Этот процесс можно продолжить вдоль произвольной кривой в пространстве. Конечно, данная процедура синхронизации часов зависит от выбора системы координат (нековариантна) и зависит также от выбора кривой, соединяющей две точки пространства-времени.

Рассмотрим замкнутую кривую γ в пространстве-времени с началом и концом в точке A . Произведем синхронизацию часов вдоль кривой γ описанным выше способом. Тогда после возвращения в точку A временная координата получит приращение

$$\Delta x^0 := - \oint_{\gamma} \frac{g_{0\mu} dx^\mu}{g_{00}}.$$

Отсюда следует, что синхронизация часов в общем случае невозможна, т.к. приращение Δx^0 в исходной точке A может быть отлично от нуля. Кроме того, если мы хотим синхронизировать часы во всей области пространства-времени $\mathbb{U} \subset \mathbb{M}$, которая покрывается данной системой координат, то равенство $\Delta x^0 = 0$ должно также выполняться для любой замкнутой кривой γ , целиком лежащей в $\mathbb{U}$. Отсюда вытекает

Предложение 23.13.2. *Для того, чтобы в выбранной системе координат x^α , где x^0 – время и все сечения $x^0 = \text{const}$ пространственноподобны, покрывающей некоторую область $\mathbb{U} \subset \mathbb{M}$, можно было синхронизировать часы во всей области $\mathbb{U}$ необходимо и достаточно чтобы $g_{0\mu} = 0$.* $\square$

Если в некоторой области смешанные компоненты метрики равны нулю, $g_{0\mu} = 0$, то часы можно синхронизировать. В этом случае одновременными будут те события, которые происходят при одинаковом значении наблюдаемого времени x_0 .

Вернемся к рассмотрению синхронной системы координат и докажем теорему существования.

Предложение 23.13.3. *Пусть задано пространство-время $\mathbb{M}$ с метрикой лоренцевой сигнатуры. Тогда в некоторой окрестности каждой точки $x \in \mathbb{M}$ существует система координат, в которой метрика имеет блочно диагональный вид (23.174).*

Доказательство. Выберем в многообразии $\mathbb{M}$ произвольную достаточно гладкую пространственноподобную гиперповерхность $\mathbb{N}$, содержащую точку $x \in \mathbb{N} \hookrightarrow \mathbb{M}$. Пусть y^μ – некоторая система координат на гиперповерхности $\mathbb{N}$ в окрестности точки x . Построим на $\mathbb{N}$ векторное поле n , перпендикулярное к гиперповерхности. Через каждую точку $y \in \mathbb{N}$ в направлении n проведем экстремаль в обоих направлениях. Мы уже знаем, что такая экстремаль существует и единственна (см. раздел 16). Поскольку гиперповерхность пространственноподобна, то векторное поле и экстремали времениподобны. Выберем в качестве канонического параметра вдоль каждой экстремали ее длину t таким образом, чтобы гиперповерхность $\mathbb{N}$ задавалась уравнением $t = 0$. Тогда в некоторой окрестности $\mathbb{U}$ гиперповерхности, $\mathbb{U} \subset \mathbb{M}$, будет определена система координат $y = \{y^0 := t, y^\mu\} \in \mathbb{U}$. Это и есть искомая синхронная система координат.

Покажем это. По построению, координатная кривая $\{y^0 = t, y^\mu = \text{const}\}$, $t \in \mathbb{R}$, является экстремалью. Ее вектор скорости в построенной системе координат имеет одну отличную от нуля компоненту $\dot{y}^\alpha = \delta_0^\alpha$. Поскольку экстремаль удовлетворяет уравнению

$$\ddot{y}^\alpha = -\Gamma_{\beta\gamma}^\alpha \dot{y}^\beta \dot{y}^\gamma,$$

то в построенной системе координат на метрику наложены условия $\Gamma_{00}^\alpha = 0$. Опустив индекс α , получим уравнения на компоненты метрики:

$$\partial_0 g_{0\alpha} - \frac{1}{2} \partial_\alpha g_{00} = 0. \quad (23.176)$$

Поскольку в качестве параметра вдоль экстремали выбрана ее длина, то касательный вектор ∂_0 имеет единичную длину. Следовательно, в построенной системе координат $g_{00} = 1$. Тогда уравнения (23.176) примут вид $\partial_0 g_{0\alpha} = 0$, т.е. компоненты $g_{0\mu}$ не зависят от времени. Кроме того, вектор скорости, по построению, перпендикулярен гиперповерхности на $\mathbb{N}$. Это значит, что в начальный момент времени t пространственно-временные компоненты метрики равны нулю, $g_{0\mu}(t = 0) = 0$. Поскольку они не зависят от времени, то это равенство выполнено всюду в $\mathbb{U}$. Тем самым построенная система координат является синхронной. $\square$

Ниже мы докажем обратное утверждение: если метрика имеет блочно диагональный вид (23.174), то координатные линии, соответствующие времени, являются экстремалими. Это значит, что единственный произвол при построении синхронной системы отсчета – это выбор пространственного сечения $\mathbb{N}$, которое может быть произвольно, и выбор пространственных координат на $\mathbb{N}$.

Замечание. Если в качестве исходной гиперповерхности $\mathbb{N}$ для построения координат выбрать времениподобную гиперповерхность, то метрика примет вид

$$g_{\alpha\beta} = \begin{pmatrix} -1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix},$$

где $g_{\mu\nu}$ – метрика лоренцевой сигнатуры на сечениях $x^0 = \text{const}$. Вообще говоря, построенная система координат не зависит от сигнатуры метрики. Аналогичные системы координат рассматривались еще Гауссом в римановой геометрии. $\square$

Замечание. Построение синхронной системы координат начиналось с выбора гиперповерхности $\mathbb{N} \hookrightarrow \mathbb{M}$, которая имеет размерность $n - 1$. Аналогичные системы координат можно строить, стартуя с подмногообразия $\mathbb{N}$ произвольной размерности, $0 \leq \dim \mathbb{N} < n$. Детали построения таких систем координат, которые названы полу-геодезическими, можно найти в [122]. В частном случае, когда все экстремали стартуют из одной точки, $\dim \mathbb{N} = 0$, получаем нормальные координаты, рассмотренные в разделе 16.9. $\square$

Определение. Сечения пространства-времени в синхронной системе координат $x^0 = t = \text{const}$ образуют семейство гиперповерхностей, которые называются (*геодезически*) *параллельными* друг другу. $\square$

Для римановых многообразий понятие экстремали и геодезической совпадают. Поэтому построенная система координат иногда называется *полугеодезической*. Мы будем использовать принятый в физике термин “временная калибровка” для метрики.

Обратная метрика во временной калибровке имеет вид

$$g^{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & g^{\mu\nu} \end{pmatrix}, \quad (23.177)$$

где $g^{\mu\nu}$ – метрика, обратная к $g_{\mu\nu}$. Ясно, что требование блочной диагональности метрики (23.174) эквивалентно требованию блочной диагональности обратной метрики.

Чтобы найти явный вид преобразований координат к синхронной системе отсчета $x^\alpha \mapsto y^\alpha(x)$, необходимо проинтегрировать уравнения экстремалей с заданными начальными условиями на гиперповерхности $\mathbb{N} \hookrightarrow \mathbb{M}$. Это – обычная механическая задача, которую можно решить, интегрируя уравнение Гамильтона–Якоби (см. раздел 21.8). Обрисуем кратко этот подход.

При преобразовании к синхронной системе координат $x^\alpha \mapsto y^\alpha(x)$ временная компонента обратной метрики преобразуется по тензорному закону. Следовательно, функция перехода $y^0(x)$ должна удовлетворять уравнению

$$1 = g^{\alpha\beta} \frac{\partial y^0}{\partial x^\alpha} \frac{\partial y^0}{\partial x^\beta}. \quad (23.178)$$

Это уравнение на функцию $y^0(x)$ совпадает с укороченным уравнением Гамильтона–Якоби (16.55). Поэтому временную координату y^0 можно отождествить с функцией действия точечной частицы. Общий интеграл укороченного уравнения Гамильтона–Якоби зависит от n постоянных. Поскольку функция действия определена с точностью до добавления постоянной, то одна из постоянных интегрирования аддитивна. Поэтому общий интеграл уравнения Гамильтона–Якоби можно представить в виде

$$y^0 = f(x, c) + A(c),$$

где f – некоторая функция от точки многообразия и $n-1$ постоянных интегрирования $c = \{c^\mu\} = (c^1, \dots, c^{n-1})$, а $A(c)$ – аддитивная постоянная, которую можно рассматривать как некоторую функцию постоянных c . Уравнения $y^0 = \text{const}$ определяют гиперповерхность $\mathbb{N} \hookrightarrow \mathbb{M}$. Локально решение этого уравнения можно представить в виде $x^\alpha = x^\alpha(c)$. То есть постоянные c образуют систему координат на $\mathbb{N}$. Временные координатные линии y^0 определяются уравнениями $\partial y^0 / \partial c^\mu = 0$ или

$$\frac{\partial f}{\partial c^\mu} = -\frac{\partial A}{\partial c^\mu}.$$

Если зафиксировать точку на гиперповерхности $\mathbb{N}$ (постоянные c), то в правой части данной системы уравнений будут стоять некоторые константы. Соответствующее решение системы $x^\mu = x^\mu(x^0)$ определит кривую, где роль параметра вдоль кривой играет координата x^0 , проходящую через данную точку $c \in \mathbb{N}$. Эти кривые являются траекториями частиц и, следовательно, экстремалами.

Из вида обратной метрики (23.177) следует уравнение

$$\delta^{0\gamma} = g^{\alpha\beta} \frac{\partial y^0}{\partial x^\alpha} \frac{\partial y^\gamma}{\partial x^\beta}, \quad \gamma = 0, \dots, n-1,$$

которое обобщает уравнение (23.178). Умножив его на матрицу Якоби преобразования координат, получим равенство

$$g^{\alpha\beta} \frac{\partial y^0}{\partial x^\alpha} = \frac{\partial x^\beta}{\partial y^0}. \quad (23.179)$$

Вектор нормали к гиперповерхностям имеет вид

$$n = g^{\alpha\beta} \frac{\partial y^0}{\partial x^\beta} \partial_\alpha.$$

С учетом равенства (23.179) отсюда следует, что вектор нормали касателен к координатным линиям y^0 :

$$\frac{\partial}{\partial y^0} = \frac{\partial x^\alpha}{\partial y^0} \partial_\alpha = g^{\alpha\beta} \frac{\partial y^0}{\partial x^\beta} \partial_\alpha.$$

Таким образом, координатные линии y^0 являются экстремальными, перпендикулярными к гиперповерхностям $y^0 = \text{const}$.

Перейдем к вычислению явного вида основных геометрических объектов в синхронной системе координат. Прямые вычисления приводят к следующим выражениям для символов Кристоффеля (6.24):

$$\begin{aligned} \Gamma_{00}^0 &= \Gamma_{00}^\mu = \Gamma_{0\mu}^0 = \Gamma_{\mu 0}^0 = 0, \\ \Gamma_{0\mu}^\nu &= \Gamma_{\mu 0}^\nu = \frac{1}{2} g^{\nu\rho} \partial_0 g_{\mu\rho}, \\ \Gamma_{\mu\nu}^0 &= -\frac{1}{2} \partial_0 g_{\mu\nu}, \\ \Gamma_{\mu\nu}^\rho &= \hat{\Gamma}_{\mu\nu}^\rho, \end{aligned} \quad (23.180)$$

где $\hat{\Gamma}_{\mu\nu}^\rho$ – символы Кристоффеля на пространственноподобном сечении $x^0 = \text{const}$, построенные только по метрике $g_{\mu\nu}$. В настоящем разделе знак тильды, который мы используем для обозначения геометрических объектов римановой геометрии при нулевых тензорах кручения и неметричности, для простоты, опущен.

Несложные вычисления приводят к следующим выражениям для компонент тензора кривизны со всеми опущенными индексами (6.81)

$$\begin{aligned} R_{0\mu 0\nu} &= \frac{1}{2} \partial_{00}^2 g_{\mu\nu} - \frac{1}{4} g^{\rho\sigma} \partial_0 g_{\mu\rho} \partial_0 g_{\nu\sigma}, \\ R_{0\mu\nu\rho} &= -R_{\mu 0\nu\rho} = \frac{1}{2} (\hat{\nabla}_\nu \partial_0 g_{\mu\rho} - \hat{\nabla}_\rho \partial_0 g_{\mu\nu}), \\ R_{\mu\nu\rho\sigma} &= \hat{R}_{\mu\nu\rho\sigma} + \frac{1}{4} (\partial_0 g_{\mu\rho} \partial_0 g_{\nu\sigma} - \partial_0 g_{\nu\rho} \partial_0 g_{\mu\sigma}), \end{aligned} \quad (23.181)$$

где $\hat{\nabla}_\nu$ обозначает ковариантную производную на пространственноподобном сечении:

$$\hat{\nabla}_\nu \partial_0 g_{\mu\rho} := \partial_\nu \partial_0 g_{\mu\rho} - \hat{\Gamma}_{\nu\mu}^\sigma \partial_0 g_{\sigma\rho} - \hat{\Gamma}_{\nu\rho}^\sigma \partial_0 g_{\mu\sigma},$$

и $\hat{R}_{\mu\nu\rho\sigma}$ – тензор кривизны пространственноподобного сечения $t = \text{const}$, построенный только по метрике $g_{\mu\nu}$. Свертка с обратной метрикой дает соответствующие тензор Риччи и скалярную кривизну:

$$\begin{aligned} R_{00} &= \frac{1}{2} g^{\mu\nu} \partial_{00}^2 g_{\mu\nu} - \frac{1}{4} g^{\mu\nu} g^{\rho\sigma} \partial_0 g_{\mu\rho} \partial_0 g_{\nu\sigma}, \\ R_{0\mu} &= R_{\mu 0} = \frac{1}{2} g^{\nu\rho} (\hat{\nabla}_\mu \partial_0 g_{\nu\rho} - \hat{\nabla}_\rho \partial_0 g_{\nu\mu}), \\ R_{\mu\nu} &= \hat{R}_{\mu\nu} + \frac{1}{2} \partial_{00}^2 g_{\mu\nu} - \frac{1}{2} g^{\rho\sigma} \partial_0 g_{\mu\rho} \partial_0 g_{\nu\sigma} + \frac{1}{4} \partial_0 g_{\mu\nu} g^{\rho\sigma} \partial_0 g_{\rho\sigma}, \\ R &= \hat{R} + g^{\mu\nu} \partial_{00}^2 g_{\mu\nu} - \frac{3}{4} g^{\mu\nu} g^{\rho\sigma} \partial_0 g_{\mu\rho} \partial_0 g_{\nu\sigma} + \frac{1}{4} (g^{\mu\nu} \partial_0 g_{\mu\nu})^2. \end{aligned} \quad (23.182)$$

Уравнения для экстремалей $x^\alpha(\tau)$ во временной калибровке имеют вид

$$\ddot{x}^0 = \frac{1}{2} \partial_0 g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu, \quad (23.183)$$

$$\ddot{x}^\mu = -g^{\mu\nu} \partial_0 g_{\nu\rho} \dot{x}^0 \dot{x}^\rho - \hat{\Gamma}_{\nu\rho}^\mu \dot{x}^\nu \dot{x}^\rho, \quad (23.184)$$

где точка обозначает дифференцирование по каноническому параметру τ . Из вида уравнений сразу следует

Предложение 23.13.4. *Если выбрана синхронная система координат, то временные координатные линии $\{x^0 = t, x^\mu = \text{const}\}$ являются экстремальными.*

Это утверждение уже было доказано другим способом при рассмотрении укороченного уравнения Гамильтона–Якоби (23.178).

Из уравнений для времениподобных экстремалей, отличных от координатных линий x^0 , можно исключить временную компоненту скорости. Для этого воспользуемся законом сохранения (16.29)

$$(\dot{x}^0)^2 - \dot{\mathbf{x}}^2 = C_0 = \text{const}, \quad (23.185)$$

где $\dot{\mathbf{x}}^2 := -g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu > 0$, и исключим производную $\dot{x}^0$ из уравнения (23.184). В результате получим замкнутую систему уравнений только для пространственных координат экстремали:

$$\ddot{x}^\mu = -g^{\mu\nu}\partial_0 g_{\nu\rho}\dot{x}^\rho \sqrt{|C_0 + \dot{\mathbf{x}}^2|} - \hat{\Gamma}_{\nu\rho}{}^\mu \dot{x}^\nu \dot{x}^\rho.$$

Неоднозначность при извлечении корня несущественна, т.к. соответствует обращению времени $x^0 \mapsto -x^0$. Отсюда следует, что во временной калибровке пространственные компоненты экстремали $\{x^\mu(t)\}$ в общем случае не являются экстремальными для пространственной части метрики $g_{\mu\nu}$. В частном случае, когда пространственная метрика $g_{\mu\nu}$ не зависит от времени x^0 , проекция экстремали $\{x^\alpha(\tau)\} \mapsto \{c^0, x^\mu(\tau)\}$ на пространственное сечение $x^0 = c^0 = \text{const}$ является экстремалью для метрики $g_{\mu\nu}$ на этом сечении.

Рассмотрим еще один способ записи уравнений для экстремалей. Пусть $\dot{x}^0 \neq 0$. Тогда уравнения для экстремалей во временной калибровке (23.183), (23.184) можно переписать в эквивалентном виде, заменив канонический параметр τ на наблюдаемое время x^0 . Другими словами, рассмотрим траектории $x^\mu(x^0)$, т.е. форму экстремалей. Поскольку

$$\begin{aligned} \partial_0 x^\mu &= \frac{\dot{x}^\mu}{\dot{x}^0}, \\ \partial_{00}^2 x^\mu &= \frac{\ddot{x}^\mu \dot{x}^0 - \dot{x}^\mu \ddot{x}^0}{(\dot{x}^0)^3}, \end{aligned}$$

то, воспользовавшись уравнениями для экстремалей, получим равенство

$$\partial_{00}^2 x^\mu = -\hat{\Gamma}_{\nu\rho}{}^\mu \partial_0 x^\nu \partial_0 x^\rho - g^{\mu\nu} \partial_0 g_{\nu\rho} \partial_0 x^\rho - \frac{1}{2} \partial_0 x^\mu \partial_0 g_{\nu\rho} \partial_0 x^\nu \partial_0 x^\rho. \quad (23.186)$$

Это и есть уравнение для траекторий $x^\mu(x^0)$. Тем самым мы получили систему уравнений, определяющих форму экстремалей. Их количество на единицу меньше, чем число исходных уравнений. Зависимость канонического параметра τ от наблюдаемого времени x^0 можно определить из закона сохранения (23.185), который запишем в виде

$$(\dot{x}^0)^2 (1 + g_{\mu\nu} \partial_0 x^\mu \partial_0 x^\nu) = C_0.$$

Отсюда для ненулевых экстремалей, $C_0 \neq 0$, получаем выражение для производной

$$\dot{x}^0 = \pm \sqrt{\frac{C_0}{1 + g_{\mu\nu} \partial_0 x^\mu \partial_0 x^\nu}}.$$

Это уравнение позволяет определить зависимость $x^0(\tau)$ для заданной траектории.

Для нулевых экстремалей $C_0 = 0$ и $1 + g_{\mu\nu}\partial_0 x^\mu \partial_0 x^\nu = 0$. Последнее условие согласуется с уравнениями (23.186). Действительно, прямые вычисления приводят к равенству

$$\partial_0(1 + g_{\mu\nu}\partial_0 x^\mu \partial_0 x^\nu) = -(1 + g_{\mu\nu}\partial_0 x^\mu \partial_0 x^\nu)\partial_0 g_{\rho\sigma}\partial_0 x^\rho \partial_0 x^\sigma.$$

То есть, если условие $1 + g_{\mu\nu}\partial_0 x^\mu \partial_0 x^\nu = 0$ выполнено в начальный момент времени, то оно будет также выполнено во все последующие моменты. Это значит, что для нулевых экстремалей зависимость $x^0(\tau)$ не определяется уравнениями (23.186) и может быть произвольна. В этом случае канонический параметр можно просто отождествить с временем: $\tau := x^0$.

Теперь покажем, что синхронная система координат в общей теории относительности при наличии материи не может быть статична в том смысле, что пространственные компоненты метрики $g_{\mu\nu}$ обязательно зависят от x^0 . Для определенности, рассмотрим четырехмерное пространство-время. С этой целью введем обозначение для производных по времени от компонент метрики

$$p_{\mu\nu} := \partial_0 g_{\mu\nu}.$$

Это нечто, напоминающее импульсы, сопряженные к пространственным компонентам метрики. След новых переменных равен производной от определителя метрики

$$p := g^{\mu\nu} p_{\mu\nu} = \partial_0[\ln(-g)], \quad g := \det g_{\alpha\beta},$$

т.к. во временной калибровке $\det g_{\alpha\beta} = \det g_{\mu\nu}$. В новых переменных компоненты тензора Риччи (23.182) примут вид:

$$\begin{aligned} R_{00} &= \frac{1}{2}\partial_0 p + \frac{1}{4}p_{\mu\nu}p^{\mu\nu}, \\ R_{0\mu} &= \frac{1}{2}\hat{\nabla}_\mu p - \frac{1}{2}\hat{\nabla}_\nu p_\mu{}^\nu, \\ R_{\mu\nu} &= \hat{R}_{\mu\nu} + \frac{1}{2}\partial_0 p_{\mu\nu} - \frac{1}{2}p_\mu{}^\rho p_{\nu\rho} + \frac{1}{4}p_{\mu\nu}p. \end{aligned} \quad (23.187)$$

Запишем уравнения Эйнштейна (23.4) без космологической постоянной

$$R_{00} = -\frac{1}{2\kappa} \left(T_{00} - \frac{1}{2}T_M \right), \quad (23.188)$$

$$R_{0\mu} = -\frac{1}{2\kappa} T_{0\mu}, \quad (23.189)$$

$$R_{\mu\nu} = -\frac{1}{2\kappa} \left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T_M \right). \quad (23.190)$$

Эти уравнения для непрерывной среды с тензором энергии-импульса (23.125) принимают вид

$$R_{00} = -\frac{1}{2\kappa} \left((\mathcal{E} + \mathcal{P})(u_0)^2 - \frac{1}{2}\mathcal{E} + \frac{1}{2}\mathcal{P} \right), \quad (23.191)$$

$$R_{0\mu} = -\frac{1}{2\kappa} (\mathcal{E} + \mathcal{P})u_0 u_\mu, \quad (23.192)$$

$$R_{\mu\nu} = -\frac{1}{2\kappa} \left((\mathcal{E} + \mathcal{P})u_\mu u_\nu - \frac{1}{2}g_{\mu\nu}\mathcal{E} + \frac{1}{2}g_{\mu\nu}\mathcal{P} \right). \quad (23.193)$$

Предложение 23.13.5. Пусть пространство-время заполнено непрерывной средой с положительной плотностью энергии и давлением: $\mathcal{E} > 0$ и $\mathcal{P} > 0$. Тогда уравнения Эйнштейна (23.191)–(23.193) без космологической постоянной в синхронной системе координат не имеют статических решений.

Доказательство. В синхронной системе координат $u^2 = (u_0)^2 + g^{\mu\nu}u_\mu u_\nu = 1$, и поэтому $u_0 \geq 1$. Статичность метрики означает, что $p_{\mu\nu} = 0$. В этом случае $R_{0\mu} = 0$, и из уравнения (23.192) следует равенство $u_\mu = 0$. Следовательно, $u_0 = 1$. Тогда уравнение (23.191) примет вид

$$0 = \frac{1}{2}\mathcal{E} + \frac{3}{2}\mathcal{P},$$

что противоречит сделанным предположениям. $\square$

Предложение 23.13.6. Вакуумные уравнения Эйнштейна без космологической постоянной в синхронной системе координат имеют только плоские статические решения, для которых тензор кривизны равен нулю.

Доказательство. При нулевом тензоре энергии-импульса материи, $T_{\mu\alpha\beta} = 0$, уравнения (23.188) и (23.189) для $p_{\mu\nu} = 0$ удовлетворяются автоматически. Уравнение (23.193) сводится к уравнению

$$\hat{R}_{\mu\nu} = 0.$$

Если пространство трехмерно, то полный тензор кривизны взаимно однозначно определяется тензором Риччи и, следовательно, полный тензор кривизны пространства равен нулю, $\hat{R}_{\mu\nu\rho\sigma} = 0$. Теперь из выражения для компонент тензора кривизны (23.181) следует, что полный тензор кривизны пространства-времени также обращается в нуль, $R_{\alpha\beta\gamma\delta} = 0$. В свою очередь, это значит, что локально существует такая система координат, в которой метрика четырехмерного пространства-времени является лоренцевой, $g_{\alpha\beta} = \eta_{\alpha\beta}$. Это и означает тривиальность решений уравнений Эйнштейна. Глобально соответствующее пространство-время представляет собой либо пространство Минковского, либо все возможные торы или цилиндры, получающиеся из пространства Минковского путем отображения всех или части декартовых координат на окружность. $\square$

В заключение раздела рассмотрим еще одно свойство решений уравнений Эйнштейна в синхронной системе отсчета. Предположим, что выполнено сильное энергетическое условие (23.139). Тогда из уравнения (23.191) следует неравенство

$$\frac{1}{2}\partial_0 p + \frac{1}{4}p_\mu{}^\nu p_\nu{}^\mu \leq 0.$$

Справедливо алгебраическое неравенство

$$p_\mu{}^\nu p_\nu{}^\mu \geq \frac{1}{3}p^2,$$

которое нетрудно доказать путем диагонализации матрицы $p_\mu{}^\nu$. Поэтому должно быть выполнено неравенство

$$\partial_0 p + \frac{1}{6}p^2 \leq 0.$$

Перепишем его в виде

$$\partial_0 \frac{1}{p} \geq \frac{1}{6}. \quad (23.194)$$

Допустим, что в некоторый момент времени $x^0 := t$ величина $1/p$ положительна. Тогда при уменьшении времени t за конечное время функция $1/p$ обратится в нуль, поскольку производная больше нуля. Это означает, что модуль определителя метрики $|g|$ обращается в нуль. Допустим, что это происходит в момент времени t_0 . Вблизи этой точки положим

$$-g = C(t - t_0)^k, \quad C = \text{const},$$

с некоторым показателем степени $k \in \mathbb{R}_+$. Тогда неравенство (23.194) ограничивает показатель степени $k \leq 6$. Отсюда следует, что модуль определителя метрики обращается в нуль не быстрее, чем $(t - t_0)^6$.

Если в начальный момент времени $t = 0$ величина $1/p$ отрицательна, то то же самое происходит при положительных временах.

Обращение в нуль определителя метрики отнюдь не означает, что возникает сингулярность в пространстве-времени. Например, определитель евклидовой метрики в сферической системе координат равен нулю в начале координат. Как правило, обращение в нуль определителя метрики связано с выбором системы отсчета. Выше было показано, что синхронная система координат возникает при построении семейства экстремалей, которые ортогональны некоторой пространственноподобной гиперповерхности. В общем случае на конечном расстоянии эти экстремали начинают пересекаться, образуя каустические поверхности. В точках пересечения экстремалей система координат вырождается, что может приводить к обращению в нуль определителя метрики.

23.13.3 Калибровка светового конуса

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, на котором задана метрика $g_{\alpha\beta}$ лоренцевой сигнатуры.

Определение. Система координат, в которой метрика имеет вид

$$g_{\alpha\beta} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & & & & \\ 0 & & & & \\ \vdots & & & g_{\mu\nu} & \\ 0 & & & & \end{pmatrix}, \quad \begin{aligned} \alpha, \beta &= 0, 1, \dots, n-1, \\ \mu, \nu &= 1, 2, \dots, n-1, \end{aligned} \quad (23.195)$$

называется калибровкой *светового конуса*. □

Ниже мы докажем, что локально калибровка светового конуса существует. При этом n произвольных функций, параметризующих преобразования координат, используются для фиксирования n компонент метрики:

$$g_{00} = g_{02} = g_{03} = \dots = g_{0n-1} = 0, \quad g_{01} = 1.$$

Для построения, системы координат, соответствующей калибровке светового конуса, введем несколько определений. Допустим, что гиперповерхность $\mathbb{S} \hookrightarrow \mathbb{M}$ задана параметрически $x^\alpha = x^\alpha(y^\mu)$, где y^μ – координаты на гиперповерхности. Вложение $\mathbb{S} \hookrightarrow \mathbb{M}$ индуцирует на гиперповерхности метрику

$$ds^2 = h_{\mu\nu} dy^\mu dy^\nu = g_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} dy^\mu dy^\nu. \quad (23.196)$$

В этом выражении индуцированная метрика $h_{\mu\nu}$ в общем случае не совпадает с блоком $g_{\mu\nu}$, входящим в выражение (23.195).

Определение. Если индуцированная метрика $h_{\mu\nu}$ на гиперповерхности невырождена, то гиперповерхность называется *неизотропной*. В противном случае, если $\det h_{\mu\nu} = 0$, гиперповерхность называется *изотропной*. $\square$

В дальнейшем мы увидим, что индуцированная метрика может быть вырождена только тогда, когда метрика на исходном многообразии $\mathbb{M}$ не является положительно или отрицательно определенной. Например, она может иметь лоренцеву сигнатуру.

Набор $n - 1$ векторных полей на $\mathbb{S}$, которые нумеруются индексом μ ,

$$e_\mu := \partial_\mu x^\alpha \partial_\alpha := \frac{\partial x^\alpha}{\partial y^\mu} \partial_\alpha$$

образуют базис касательного пространства к гиперповерхности $\mathbb{S} \hookrightarrow \mathbb{M}$. Вектор $m = m^\alpha \partial_\alpha$ ортогонален к гиперповерхности, если его скалярное произведение со всеми касательными векторами равно нулю:

$$(m, e_\mu) := g_{\alpha\beta} m^\alpha \partial_\mu x^\beta = m_\alpha \partial_\mu x^\alpha = 0, \quad \forall \mu = 1, 2, \dots, n - 1, \quad (23.197)$$

где скобки означают скалярное произведение. Поскольку ранги матриц $g_{\alpha\beta}$ и $\partial_\mu x^\alpha$ равны, соответственно, n и $n - 1$, то ранг матрицы $g_{\alpha\beta} \partial_\mu x^\beta$ равен $n - 1$. Это означает, что система линейных однородных уравнений на компоненты нормального вектора m^α имеет единственное нетривиальное решение с точностью до умножения на произвольную отличную от нуля постоянную.

Важную роль в дальнейшем будет играть следующее

Предложение 23.13.7. *Нормаль к гиперповерхности изотропна тогда и только тогда, когда поверхность изотропна.*

Доказательство. Индуцированную метрику $h_{\mu\nu}$ (23.196) можно рассматривать, как произведение двух прямоугольных матриц $\partial_\mu x^\alpha$ и $g_{\alpha\beta} \partial_\mu x^\beta$ размеров, соответственно, $(n - 1) \times n$ и $n \times (n - 1)$. Тогда из формулы Бине–Коши (37.13) следует, что

$$\det h_{\mu\nu} = \sum_{\gamma=1}^n \det_\gamma(\partial_\mu x^\alpha) \det_\gamma(g_{\alpha\beta} \partial_\mu x^\beta),$$

где $\det_\gamma$ обозначает определитель квадратной матрицы, полученной из матрицы размера $(n - 1) \times n$ вычеркиванием γ -того столбца. Поскольку $m_\alpha \partial_\mu x^\alpha = 0$, то между столбцами матрицы $\partial_\mu x^\alpha = 0$ имеется линейная зависимость. Предположим, что некоторая компонента ковектора $\{m_\alpha\}$ отлична от нуля, например, $m_{\delta_1} \neq 0$. Тогда, в силу свойств определителя и линейной зависимости столбцов, справедливо равенство

$$\det_\gamma(\partial_\mu x^\alpha) = m_\gamma \frac{(-1)^{\delta_1 - \gamma} \det_{\delta_1}(\partial_\mu x^\alpha)}{m_{\delta_1}}.$$

Аналогичная формула имеет место и для второго определителя под знаком суммы

$$\det_\gamma(g_{\alpha\beta} \partial_\mu x^\beta) = m^\gamma \frac{(-1)^{\delta_2 - \gamma} \det_{\delta_2}(g_{\alpha\beta} \partial_\mu x^\beta)}{m^{\delta_2}}$$

для некоторого индекса δ_2 . Таким образом, определитель индуцированной метрики пропорционален длине нормального вектора к гиперповерхности

$$\det h_{\mu\nu} = (-1)^{\delta_1 + \delta_2} \frac{\det_{\delta_1}(\partial_\mu x^\alpha) \det_{\delta_2}(g_{\alpha\beta} \partial_\mu x^\beta)}{m_{\delta_1} m^{\delta_2}} m_\gamma m^\gamma,$$

где в правой части предполагается суммирование только по индексу γ . Индексы δ_1 и δ_2 фиксированы и выбраны таким образом, что $m_{\delta_1} \neq 0$ и $m^{\delta_2} \neq 0$. Поскольку коэффициент пропорциональности между $\det h_{\mu\nu}$ и $m_\gamma m^\gamma$ отличен от нуля, то из этого равенства следует утверждение предложения. $\square$

Рассмотрим некоторые свойства изотропных гиперповерхностей.

Допустим, что изотропная гиперповерхность $\mathbb{S}$ задана уравнением

$$W(x) = \text{const}, \quad (23.198)$$

где $W(x)$ – некоторая достаточно гладкая функция координат. При этом мы считаем, что различным значениям постоянной соответствуют различные непересекающиеся изотропные гиперповерхности. Обозначим, как и ранее, координаты на $\mathbb{S}$ через y^μ . То есть функции $x^\alpha(y)$ удовлетворяют уравнению (23.198) для всех y . Продифференцируем это уравнение:

$$\frac{\partial W}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial y^\mu} = \frac{\partial W}{\partial x^\alpha} e_\mu^\alpha = 0.$$

Сравнение полученного равенства с определением нормали (23.197) показывает, что вектор нормали к гиперповерхности имеет вид

$$m = m^\alpha \partial_\alpha = g^{\alpha\beta} \frac{\partial W}{\partial x^\beta} \partial_\alpha. \quad (23.199)$$

Поскольку вектор нормали к изотропной гиперповерхности имеет нулевую длину, то выполнено равенство

$$m^2 = g^{\alpha\beta} \frac{\partial W}{\partial x^\alpha} \frac{\partial W}{\partial x^\beta} = 0. \quad (23.200)$$

Таким образом доказано

Предложение 23.13.8. *Поверхность $\mathbb{S}$, определяемая уравнением (23.198), изотропна тогда и только тогда, когда выполнено равенство (23.200).*

Следующее утверждение не имеет аналога в римановой геометрии.

Предложение 23.13.9. *Нормальный вектор к изотропной поверхности $\mathbb{S}$ является также и касательным к ней.*

Доказательство. Дополним координаты на гиперповерхности до системы координат в $\mathbb{M}$ еще одной координатой $y^0 := W$. В новой системе координат гиперповерхности определяются условиями $y^0 = \text{const}$, и уравнение (23.200) примет вид $g^{00} = 0$. Тогда нулевая компонента нормального вектора (23.199) обратится в нуль, $m^0 = 0$, т.к. $\partial W / \partial y^\mu = 0$. Это значит, что нормальный вектор является также и касательным. $\square$

На самом деле сделанное утверждение очевидно. Действительно, вектор k будет касательным к гиперповерхности $\mathbb{S}$ тогда и только тогда, когда его скалярное произведение с вектором нормали равно нулю, $(k, m) = 0$. Это и означает, что вектор нормали m является касательным к $\mathbb{S}$, т.к. $m^2 = 0$.

Рассмотрим интегральные кривые $x(u)$ векторного поля m :

$$\dot{x}^\alpha = m^\alpha, \quad m^\alpha := g^{\alpha\beta} \frac{\partial W}{\partial x^\beta}, \quad (23.201)$$

где точка обозначает дифференцирование по параметру $u \in \mathbb{R}$. Эти кривые изотропны, т.к. выполнено равенство (23.200). Кроме того, они перпендикулярны изотропной поверхности $\mathbb{S}$. С другой стороны, вдоль каждой кривой выполнено равенство

$$\frac{dW}{du} = \frac{\partial W}{\partial x^\alpha} \frac{dx^\alpha}{du} = g^{\alpha\beta} \frac{\partial W}{\partial x^\alpha} \frac{\partial W}{\partial x^\beta} = 0.$$

Отсюда следует, что, если интегральная кривая $x(u)$ начинается на изотропной поверхности $\mathbb{S}$, то она целиком лежит на этой поверхности.

Можно сказать по-другому. Поскольку векторное поле m касательно к $\mathbb{S}$, то его интегральные кривые лежат на $\mathbb{S}$.

Вычислим теперь ускорение данной кривой:

$$\dot{x}^\alpha \nabla_\alpha \dot{x}^\beta = \dot{x}^\alpha g^{\beta\gamma} \nabla_\alpha \frac{\partial W}{\partial x^\gamma} = g^{\alpha\delta} \frac{\partial W}{\partial x^\delta} g^{\beta\gamma} \nabla_\alpha \frac{\partial W}{\partial x^\gamma}.$$

С другой стороны, продифференцируем определяющее уравнение (23.200):

$$\nabla_\alpha \left(g^{\beta\gamma} \frac{\partial W}{\partial x^\beta} \frac{\partial W}{\partial x^\gamma} \right) = 2g^{\beta\gamma} \frac{\partial W}{\partial x^\beta} \nabla_\alpha \frac{\partial W}{\partial x^\gamma} = 0.$$

Поскольку $\nabla_\alpha \nabla_\gamma W = \nabla_\gamma \nabla_\alpha W$, то отсюда вытекает, что ускорение кривой (23.201) равно нулю, $\dot{x}^\alpha \nabla_\alpha \dot{x}^\beta = 0$. Таким образом, интегральные кривые векторного поля m являются экстремалими. Отсюда следует

Предложение 23.13.10. *Все изотропные (нулевые) экстремали, которые касаются изотропной поверхности $\mathbb{S}$ хотя бы в одной точке, целиком лежат на этой поверхности.*

Теперь докажем существование калибровки светового конуса.

Предложение 23.13.11. *Пусть задано пространство-время $\mathbb{M}$ с метрикой лоренцевой сигнатуры. Тогда в некоторой окрестности каждой точки $x \in \mathbb{M}$ существует система координат, в которой метрика имеет вид (23.195).*

Доказательство. Аналогично доказательству предложения 23.13.3. Выберем произвольную пространственноподобную поверхность $\mathbb{N} \hookrightarrow \mathbb{M}$ с координатами y^μ . В каждой точке $y \in \mathbb{N}$ строим нормальный вектор n . Он времениподобен. Выберем координаты y^μ на $\mathbb{N}$ таким образом, чтобы вектор

$$e_1 := \frac{\partial x^\alpha}{\partial y^1} \partial_\alpha$$

имел единичную длину, $e_1^2 = -1$, и был перпендикулярен всем остальным касательным векторам:

$$(e_1, e_A) = 0, \quad A = 2, \dots, n-1.$$

Это всегда возможно в силу предложения 23.13.3. То есть мы фиксируем временную калибровку на поверхности $\mathbb{N}$. Затем строим вектор

$$m = n - e_1.$$

По построению, этот вектор определен на поверхности $\mathbb{N}$, имеет нулевую длину, $m^2 = 0$, перпендикулярен всем касательным векторам e_A , и выполнено равенство

$(m, e_1) = 1$. Другими словами, в каждом световом конусе в точке $y \in \mathbb{N}$ мы выбираем единственный вектор m , скалярное произведение которого с касательным вектором e_1 равно единице, а с остальными касательными векторами равно нулю.

Затем через каждую точку $y \in \mathbb{N}$ в направлении m проводим экстремаль. Выбираем канонический параметр y^0 вдоль экстремалей таким образом, чтобы гиперповерхность $\mathbb{N}$ определялась уравнением $y^0 = 0$. Тогда в некоторой окрестности гиперповерхности $\mathbb{N}$ определена система координат $\{y^\alpha\} = \{y^0, y^\mu\}$.

По построению, координатные линии y^0 являются экстремальями. Поэтому на компоненты метрики в новой системе координат имеются ограничения (23.176). В рассматриваемом случае касательный вектор к координатной линии y^0 изотропен, и, следовательно, $g_{00} = 0$. Поэтому на компоненты метрики возникает уравнение (23.176), которое сводится к условию

$$\partial_0 g_{0\mu} = 0. \quad (23.202)$$

По построению, координаты y^μ на $\mathbb{N}$ и изотропный вектор m выбраны таким образом, что на поверхности $\mathbb{N}$ выполнены условия:

$$g_{00} = 0, \quad g_{01} = 1, \quad g_{02} = g_{03} = \dots = g_{0n-1} = 0.$$

Тогда из уравнений (23.202) следует, что эти условия будут выполнены во все последующие моменты времени. Следовательно, система координат, в которой выполнена калибровка светового конуса построена.

Аналогичное построение можно провести, если в качестве исходной поверхности $\mathbb{N}$ выбрать поверхность с метрикой лоренцевой сигнатуры. $\square$

Посмотрим теперь на построение калибровки светового конуса с точки зрения уравнения Гамильтона–Якоби. Предположим, что изотропная гиперповерхность $\mathbb{S}$ задана уравнением $W(x) = 0$. Тогда условие изотропности поверхности примет вид (23.200), которое должно быть выполнено на изотропной гиперповерхности $\mathbb{S}$. Изотропная гиперповерхность, задаваемая уравнением $W = 0$, является ни чем иным, как характеристикой волнового уравнения (16.56). Уравнение (23.200) представляет собой укороченное уравнение Гамильтона–Якоби для нулевых экстремалей, где $W(x)$ – укороченная функция действия.

Рассмотрим систему линейных дифференциальных уравнений первого порядка

$$g^{\alpha\beta} \partial_\alpha W \partial_\beta \varphi = 0$$

на функцию $\varphi(x)$. Известно, что эта система имеет $n-1$ функционально независимых решений, которые обозначим через $\varphi^\mu(x)$, $\mu = 1, \dots, n-1$. Среди этих решений содержится, очевидно, и $W(x)$. Не ограничивая общности, положим $\varphi^1(x) := W(x)$. Введем новую систему координат

$$y^0 := \psi(x), \quad y^\mu := \varphi^\mu(x),$$

где $\psi(x)$ – некоторая функция, функционально независимая от $\varphi^\mu(x)$, т.е.

$$\frac{\partial y^\alpha}{\partial x^\beta} \neq 0$$

где $\{y^\alpha\} = \{y^0, y^\mu\}$. В новой системе координат компоненты обратной метрики примут вид

$$g'^{\alpha\beta} = g^{\gamma\delta} \frac{\partial y^\alpha}{\partial x^\gamma} \frac{\partial y^\beta}{\partial x^\delta}.$$

По построению,

$$g'^{01} = g'^{10} = g^{\alpha\beta} \partial_\alpha W \partial_\beta \psi \neq 0, \quad g'^{1\mu} = g'^{\mu 1} = 0,$$

причем $g'^{01} \neq 0$, т.к. ψ функционально независима от φ^μ . Эту функцию всегда можно подобрать таким образом, что $g'^{01} = 1$. Таким образом, в новой системе координат обратная метрика примет вид

$$g'^{\alpha\beta} = \begin{pmatrix} g'^{00} & 1 & g'^{0B} \\ 1 & 0 & 0 \\ g'^{A0} & 0 & g'^{AB} \end{pmatrix}, \quad A, B = 2, 3, \dots, n-1.$$

Нетрудно проверить, что метрика $g'_{\alpha\beta}$, которая обратна к матрице $g'^{\alpha\beta}$, в новой системе координат имеет вид (23.195). Тем самым локальное существование калибровки светового конуса доказано еще раз. В этой калибровке изотропные гиперповерхности $\mathbb{S}$ задаются условиями $y^1 = \text{const}$. Векторы ∂_0 , касательные к координатным линиям y^0 и, следовательно, касательные также к изотропной гиперповерхности, имеют нулевую длину, $(\partial_0, \partial_0) = g_{00} = 0$. Метрика, индуцированная на гиперповерхностях $y^1 = \text{const}$, имеет вид

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & g'_{AB} & \\ 0 & & & \end{pmatrix}, \quad A, B = 2, 3, \dots, n-1,$$

и является, очевидно, вырожденной.

Вернемся к прежним обозначениям координат через x^α вместо y^α . То есть будем считать, что в координатах x^α зафиксирована калибровка светового конуса, и метрика имеет вид (23.195), или

$$ds^2 = 2dx^0 dx^1 + g_{\mu\nu} dx^\mu dx^\nu. \quad (23.203)$$

Наша интуиция основана на таких системах координат, в которых метрика диагональна. Поскольку в калибровке светового конуса метрика недиагональна, то полезно рассмотреть простой пример плоского пространства Минковского.

Пример 23.13.2. Рассмотрим трехмерное пространство Минковского $\mathbb{R}^{1,2}$. В инерциальной системе отсчета t, x, y метрика Лоренца имеет вид

$$ds^2 = dt^2 - dx^2 - dy^2.$$

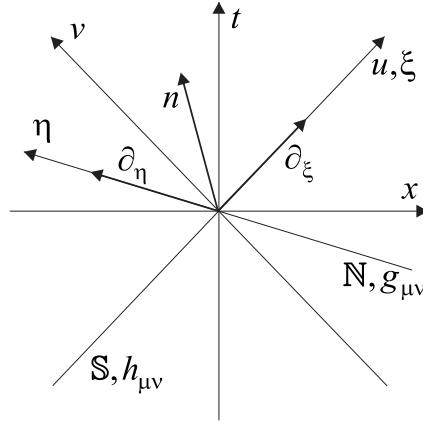
В плоскости t, x можно ввести координаты светового конуса (1.93)

$$u := \frac{1}{\sqrt{2}}(t + x), \quad v := \frac{1}{\sqrt{2}}(t - x).$$

Рассмотрим более общее преобразование координат

$$\xi = \frac{1}{\sqrt{2}}((2-a)t + ax), \quad \eta = \frac{1}{\sqrt{2}}(t - x), \quad a \in \mathbb{R}, \quad (23.204)$$

которое параметризуется одним вещественным параметром a . Эти координаты в плоскости t, x показаны на рис. 23.3. На рисунке показаны конусные оси v и u , ко-

Рис. 23.3: Координаты ξ, η в плоскости t, x .

которые определяются уравнениями, соответственно, $u = 0$ и $v = 0$. Ось ξ совпадает с осью u . Показана также ось η при $0 < a < 1$. В рассматриваемой системе координат, координата ξ по-прежнему является конусной, а координата η совпадает с конусной координатой v только при $a = 1$. Легко найти обратное преобразование координат

$$t = \frac{1}{\sqrt{2}}(\xi + a\eta), \quad x = \frac{1}{\sqrt{2}}(\xi + (a - 2)\eta).$$

Метрика Лоренца в координатах $(x^0, x^1, x^2) := (\xi, \eta, y)$ имеет вид

$$ds^2 = 2d\xi d\eta + 2(a - 1)d\eta^2 - dy^2.$$

Компоненты метрики и ее обратной в координатах ξ, η, y равны

$$g_{\alpha\beta} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2(a - 1) & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad g^{\alpha\beta} = \begin{pmatrix} -2(a - 1) & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Следовательно, координаты $(x^0, x^1, x^2) = (\xi, \eta, y)$ задают калибровку светового конуса.

Отметим, что метрика $g_{\mu\nu}$ на плоскостях $\xi = \text{const}$ является отрицательно определенной при $a < 1$ и имеет лоренцеву сигнатуру при $a > 1$.

В трехмерном пространстве Минковского существует два типа изотропных гиперповерхностей: плоскости и конусы (характеристики в примере 16.7.1). Изотропные плоскости – это все плоскости, имеющие угол $\pi/4$ с осью t .

Сравним координаты светового конуса данного примера при $0 < a < 1$ с системой координат, построенной в предложении (23.13.11) (см. рис.23.3). Пространственноподобная поверхность $\mathbb{N}$ натянута на векторы $\partial_\eta, \partial_y$ и определяется уравнением $x^0 = \xi = 0$. Изотропные поверхности $\mathbb{S}$ натянуты на векторы ∂_ξ, ∂_y и определяются уравнениями $x^1 = \eta = \text{const}$. Вектор

$$\partial_\xi = \frac{1}{\sqrt{2}}\partial_t + \frac{1}{\sqrt{2}}\partial_x$$

касается изотропной поверхности $\mathbb{S}$. Вектор

$$\partial_\eta = \frac{a}{\sqrt{2}}\partial_t - \frac{2 - a}{\sqrt{2}}\partial_x$$

касается поверхности $\mathbb{N}$. Нормированный вектор e_1 пропорционален ∂_η :

$$e_1 = \frac{1}{\sqrt{2(1-a)}}\partial_\eta = \frac{a}{2\sqrt{1-a}}\partial_t - \frac{2-a}{2\sqrt{1-a}}\partial_x.$$

Нормированный вектор нормали n к пространственноподобной поверхности $\mathbb{N}$ имеет вид

$$n = \frac{2-a}{2\sqrt{1-a}}\partial_t - \frac{a}{2\sqrt{1-a}}\partial_x.$$

Вычислим изотропный вектор m :

$$m := n - e_1 = \sqrt{1-a}\partial_t + \sqrt{1-a}\partial_x = \sqrt{2(1-a)}\partial_\xi.$$

Как и следовало ожидать, он пропорционален вектору ∂_ξ . □

Рассмотренный пример показывает, что калибровка светового конуса (23.203) определена неоднозначно и может быть зафиксирована таким образом, что метрика $g_{\mu\nu}$ на гиперповерхностях $x^0 = \text{const}$ будет либо отрицательно определена, либо иметь лоренцеву сигнатуру в произвольной конечной области $\mathbb{U}$.

Глава 24

Гамильтонова формулировка общей теории относительности

В настоящей главе мы перепишем вакуумные (без полей материи) уравнения Эйнштейна в гамильтоновом виде. Эта задача важна как для исследования классических уравнений движения, в частности, при рассмотрении задачи Коши, так и для построения квантовой теории гравитации.

Каноническая формулировка общей теории относительности была впервые дана Дираком в формализме второго порядка [147]. Он показал, что гамильтониан гравитационного поля равен линейной комбинации связей и нашел его явное выражение. Позже Арновитт, Дезер и Мизнер в серии статей, завершившихся обзором [174] (без ссылки на Дирака), существенно упростили вычисления и прояснили геометрический смысл канонических импульсов, выразив их через внешнюю кривизну пространственной гиперповерхности, вложенной в четырехмерное пространство-время. Выражение для гамильтониана было найдено ими в формализме первого порядка, когда метрика $g_{\alpha\beta}$ и симметричная аффинная связность $\Gamma_{\{\alpha\beta\}}^\gamma$ рассматриваются в качестве независимых переменных. По сути дела этот подход и упростил вычисления.

24.1 Лагранжиан Гильберта–Эйнштейна

Рассмотрим пространство-время $\mathbb{M}$ произвольной размерности, $\dim \mathbb{M} = n \geq 3$. Локальные координаты, как и ранее, мы нумеруем греческими буквами из начала алфавита x^α , $\alpha = 0, 1, \dots, n-1$. Мы предполагаем, что на $\mathbb{M}$ задана метрика $g_{\alpha\beta}$ лоренцевой сигнатуры, а кручение и тензор неметричности тождественно равны нулю (псевдориманова геометрия). Как и ранее мы предполагаем, что координата x^0 является временем, и все сечения $x^0 = \text{const}$ пространственноподобны. Уравнения движения для метрики в общей теории относительности при отсутствии полей материи следуют из вариационного принципа для действия Гильберта–Эйнштейна (23.7)

$$S_{\text{HE}} = \int dx \sqrt{|g|} R, \quad \sqrt{|g|} := \det e_\alpha^a = \sqrt{|\det g_{\alpha\beta}|}, \quad (24.1)$$

где R – псевдориманова скалярная кривизна, построенная по метрике $g_{\alpha\beta}$. В настоящей главе мы опускаем знак тильды для геометрических объектов псевдоримановой геометрии, чтобы упростить обозначения. Гравитационную постоянную перед действием Гильберта–Эйнштейна мы положили равной единице, $\kappa = 1$, поскольку уравнения не меняются при умножении полного действия на постоянную. Космологическую постоянную мы пока положим равной нулю, $\Lambda = 0$, т.к. ее учет приводит

только к изменению потенциала для метрики и не вызывает трудностей при канонической формулировке общей теории относительности.

Действие Гильберта–Эйнштейна в виде (24.1) содержит вторые производные от метрики, что создает определенные трудности для перехода к каноническому формализму. Однако, из явного вида тензора кривизны (6.88) следует, что вторые производные входят линейно. Поэтому от них можно избавиться, добавив к подинтегральному выражению полную производную. При этом мы потеряем явную ковариантность, зато действие будет зависеть только от первых производных метрики. С этой целью представим скалярную кривизну в виде двух слагаемых, что проверяется прямой проверкой,

$$R = \partial\Gamma - L_{\text{HE}}, \quad (24.2)$$

где

$$L_{\text{HE}} := g^{\alpha\beta}(\Gamma_{\alpha\beta}{}^{\gamma}\Gamma_{\gamma} - \Gamma_{\alpha\delta}{}^{\gamma}\Gamma_{\beta\gamma}{}^{\delta}), \quad (24.3)$$

$$\partial\Gamma := g^{\alpha\beta}(\partial_{\alpha}\Gamma_{\beta} - \partial_{\gamma}\Gamma_{\alpha\beta}{}^{\gamma}) = \frac{1}{\sqrt{|g|}}\partial_{\alpha}(\sqrt{|g|}g^{\alpha\beta}\Gamma_{\beta} - \sqrt{|g|}g^{\beta\gamma}\Gamma_{\beta\gamma}{}^{\alpha}) + 2L_{\text{HE}}. \quad (24.4)$$

Отсюда следует, что лагранжиан Гильберта–Эйнштейна $\sqrt{|g|}L_{\text{HE}}$ отличается от скалярной кривизны, умноженной на определитель репера, на полную производную:

$$\sqrt{|g|}R - \partial_{\alpha}(\sqrt{|g|}g^{\alpha\beta}\Gamma_{\beta} - \sqrt{|g|}g^{\beta\gamma}\Gamma_{\beta\gamma}{}^{\alpha}) = \sqrt{|g|}L_{\text{HE}},$$

и, следовательно, приводит к тем же уравнениям Эйлера–Лагранжа. Он квадратичен по символам Кристоффеля и поэтому зависит только от первых производных метрики.

24.2 АДМ параметризация метрики и репера

При анализе гамильтоновой структуры уравнений общей теории относительности Арновитт, Дезер и Мизнер (АДМ) [174] использовали специальную параметризацию метрики, которая существенно упростила вычисления. В настоящем разделе будет описана АДМ параметризация метрики и соответствующая параметризация репера, которая упрощает вычисления в переменных Картана.

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, с метрикой лоренцевой сигнатуры $(+ - \dots -)$. Пусть $\{x^{\alpha}\}$, $\alpha = 0, 1, \dots, n-1$, – система локальных координат. Выделим среди координат время $t = x^0$, тогда $\{x^{\alpha}\} = \{x^0, x^{\mu}\}$, $\mu = 1, \dots, n-1$. В дальнейшем буквы из начала греческого алфавита $\alpha, \beta, \dots$ будут пробегать все значения индексов, а из середины $\mu, \nu, \dots$ – только пространственные значения. Это правило легко запомнить по следующим включениям: $\{\mu, \nu, \dots\} \subset \{\alpha, \beta, \dots\}$ и $\{1, 2, \dots\} \subset \{0, 1, 2, \dots\}$. АДМ параметризация метрики имеет следующий вид

$$g_{\alpha\beta} = \begin{pmatrix} N^2 + N^{\rho}N_{\rho} & N_{\nu} \\ N_{\mu} & g_{\mu\nu} \end{pmatrix}, \quad (24.5)$$

где $g_{\mu\nu}$ – метрика на $(n-1)$ -мерных сечениях многообразия $x^0 = \text{const}$, которые предполагаются пространственноподобными. В выбранной параметризации вместо n компонент метрики, содержащих хотя бы один временной индекс, g_{00} и $g_{0\mu}$, введены n функций N и N_{μ} . Выше $N^{\rho} := \hat{g}^{\rho\mu}N_{\mu}$, где $\hat{g}^{\rho\mu} = (n-1) \times (n-1)$ матрица, обратная к $g_{\mu\nu}$:

$$\hat{g}^{\rho\mu}g_{\mu\nu} = \delta_{\nu}^{\rho},$$

которую мы называем обратной пространственной метрикой. В настоящей главе подъем пространственных индексов будет всегда осуществляться с помощью обратной метрики $\hat{g}^{\rho\mu}$, помеченной шляпкой, которая в общем случае не совпадает с “пространственной” частью метрики $g^{\alpha\beta}$, обратной к $g_{\alpha\beta}$: $\hat{g}^{\rho\mu} \neq g^{\rho\mu}$. Функция $N = N(x)$ называется *функцией хода*, а функции $N_\mu = N_\mu(x)$ – *функциями сдвига*. Не ограничивая общности, можно считать, что функция хода положительна $N > 0$. В этом случае АДМ параметризация метрики (24.5) является взаимно однозначной.

Интервал, соответствующий параметризации (24.5), можно записать в виде

$$ds^2 = N^2 dt^2 + g_{\mu\nu}(dx^\mu + N^\mu dt)(dx^\nu + N^\nu dt).$$

Вообще, квадрат произвольного вектора $X = X^\alpha \partial_\alpha$ в АДМ параметризации метрики (24.5) имеет вид суммы двух квадратов:

$$X^2 := g_{\alpha\beta} X^\alpha X^\beta = N^2 (X^0)^2 + g_{\mu\nu} (X^\mu + N^\mu X^0)(X^\nu + N^\nu X^0),$$

где выделен квадрат временной компоненты. Первое слагаемое в этом выражении положительно, а второе – отрицательно, т.к. метрика на пространственноподобных сечениях отрицательно определена.

Мы предполагаем, что координата $x^0 = t$ является временем, т.е. касательный вектор ∂_0 к координатной линии x^0 времениподобен. Формально это условие записывается в виде

$$(\partial_0, \partial_0) = g_{00} = N^2 + N^\rho N_\rho > 0. \quad (24.6)$$

В этом случае метрика $g_{\alpha\beta}$ имеет лоренцеву сигнатуру тогда и только тогда, когда матрица (теорема 4.2.2)

$$g_{\mu\nu} - \frac{N_\mu N_\nu}{N^2 + N^\rho N_\rho} \quad (24.7)$$

отрицательно определена. Отметим, что сама метрика $g_{\mu\nu}$, индуцированная на сечении $x^0 = \text{const}$, может и не быть отрицательно определенной. Это значит, что сечение $x^0 = \text{const}$ в общем случае не является пространственноподобным. В дальнейшем мы будем дополнительно предполагать, что координаты выбраны таким образом, чтобы все сечения $x^0 = \text{const}$ были пространственноподобны, т.е. метрика $g_{\mu\nu}$ также отрицательно определена. Это удобно для постановки задачи Коши, когда начальные данные задаются на пространственноподобной поверхности, и рассматривается их эволюция во времени.

Замечание. Аналогичным образом можно параметризовать метрику и на римановом многообразии с положительно определенной метрикой, для этого вместо времени достаточно явно выделить произвольную координату. $\square$

Метрика, обратная к (24.5), имеет вид

$$g^{\alpha\beta} = \begin{pmatrix} \frac{1}{N^2} & -\frac{N^\nu}{N^2} \\ -\frac{N^\mu}{N^2} & \hat{g}^{\mu\nu} + \frac{N^\mu N^\nu}{N^2} \end{pmatrix}. \quad (24.8)$$

Предложение 24.2.1. Пространственная матрица в правом нижнем блоке

$$g^{\mu\nu} = \hat{g}^{\mu\nu} + \frac{N^\mu N^\nu}{N^2}, \quad (24.9)$$

отрицательно определена.

Доказательство. Легко проверить, матрица (24.9) является обратной к метрике (24.7). Это значит, что отрицательная определенность метрики (24.7) эквивалентна отрицательной определенности матрицы $g^{\mu\nu}$. $\square$

Заметим, что, если метрика на многообразии $\mathbb{M}$ имеет лоренцеву сигнатуру, то условие пространственноподобности всех сечений $x^0 = \text{const}$ эквивалентно условию $N^2 > 0$. Действительно, из отрицательной определенности $g_{\mu\nu}$ следует отрицательная определенность обратной матрицы $\hat{g}^{\mu\nu}$. Тогда из уравнения (24.9) вытекает отрицательная определенность матрицы

$$g^{\mu\nu} - \frac{N^\mu N^\nu}{N^2}.$$

Это, в свою очередь, эквивалентно условию $g^{00} > 0$ или $N^2 > 0$.

Чтобы продемонстрировать тонкости, которые могут возникнуть при АДМ параметризации метрики, приведем простой

Пример 24.2.1. Рассмотрим двумерное пространство-время Минковского $\mathbb{R}^{1,1}$ с декартовыми координатами t, x . Введем новую систему координат ξ, η , зависящую от двух вещественных параметров a и b (см. рис.24.1)

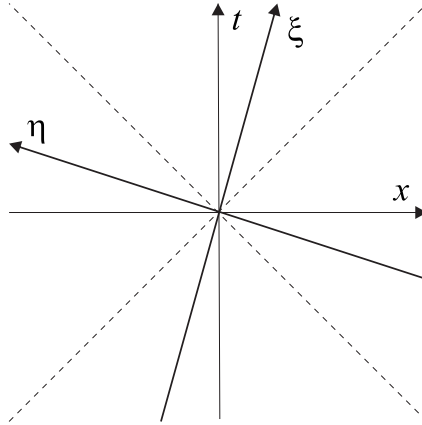


Рис. 24.1: Выбор координат на плоскости Минковского. Рисунок соответствует следующим значениям параметров: $0 < a < 1$ и $b > 1$.

$$\xi = t + ax, \quad \eta = t - bx, \quad |a| \neq 1, \quad |b| \neq 1, \quad a + b \neq 0.$$

Легко получить формулы обратного преобразования

$$t = \frac{b\xi + a\eta}{a + b}, \quad x = \frac{\xi - \eta}{a + b}.$$

В новых координатах лоренцева метрика имеет вид

$$ds^2 = dt^2 - dx^2 = \frac{1}{(a + b)^2} [(b^2 - 1)d\xi^2 + 2(ab + 1)d\xi d\eta + (a^2 - 1)d\eta^2].$$

Векторы, касательные к координатным линиям ξ и η имеют следующие компоненты:

$$\partial_\xi = \frac{b}{a + b}\partial_t + \frac{1}{a + b}\partial_x, \quad \partial_\eta = \frac{a}{a + b}\partial_t - \frac{1}{a + b}\partial_x.$$

Проанализируем АДМ параметризацию метрики в координатах $x^0 = \xi$, $x^1 = \eta$. Метрика имеет следующие компоненты:

$$g_{00} = \frac{b^2 - 1}{(a + b)^2}, \quad g_{01} = \frac{ab + 1}{(a + b)^2}, \quad g_{11} = \frac{a^2 - 1}{(a + b)^2}.$$

Функции хода и сдвига имеют вид

$$N^2 = -\frac{1}{a^2 - 1}, \quad N_1 = \frac{ab + 1}{(a + b)^2}.$$

Из условий $g_{00} > 0$ и $g_{11} < 0$ следует, соответственно, что $|b| > 1$ и $|a| < 1$. Мы видим, что эти условия являются необходимыми и достаточными для того, чтобы координатная линия ξ была времениподобной, а η – пространственноподобной. Нетрудно также проверить эквивалентность условий:

$$g_{00} > 0 \quad \Leftrightarrow \quad g_{11} - \frac{N_1 N_1}{N^2 + N^1 N_1} = -\frac{1}{b^2 - 1} < 0,$$

а также условий:

$$g^{00} > 0 \quad \Leftrightarrow \quad \hat{g}^{11} = g^{11} - \frac{N^1 N^1}{N^2} = \frac{(a + b)^2}{a^2 - 1} < 0. \quad \square$$

Обратная метрика (24.8) удобна в приложениях тем, что квадрат произвольного ковектора (1-формы) $A = dx^\alpha A_\alpha$ имеет вид суммы двух квадратов:

$$A^2 := g^{\alpha\beta} A_\alpha A_\beta = \frac{1}{N^2} (A_0 - N^\mu A_\mu)^2 + \hat{g}^{\mu\nu} A_\mu A_\nu,$$

где выделен квадрат пространственных компонент. Здесь первое слагаемое положительно, а второе отрицательно.

Используя формулу (37.11) для определителя блочных матриц, получаем выражение для определителя метрики (24.5)

$$\det g_{\alpha\beta} = N^2 \det g_{\mu\nu}. \quad (24.10)$$

Отсюда следует выражение для элемента объема

$$\sqrt{|g|} = N \sqrt{|\hat{g}|}, \quad \sqrt{|g|} := \sqrt{|\det g_{\alpha\beta}|}, \quad \sqrt{|\hat{g}|} := \sqrt{|\det g_{\mu\nu}|}. \quad (24.11)$$

Последняя формула является обобщением хорошо известной из школьного курса геометрии правила: объем призмы равен произведению площади основания на высоту. В рассматриваемом случае площадью основания является $\sqrt{|\hat{g}|}$, а высотой – функция хода N .

При проведении вычислений оказываются полезными следующие формулы, которые проверяются прямой проверкой:

$$\begin{aligned} g^{00} g^{\mu\nu} - g^{0\mu} g^{0\nu} &= \frac{\hat{g}^{\mu\nu}}{N^2}, \\ g^{\sigma\mu} g^{0\nu} - g^{\sigma\nu} g^{0\mu} &= \frac{N^\mu \hat{g}^{\sigma\nu} - N^\nu \hat{g}^{\sigma\mu}}{N^2}, \\ g^{\mu\nu} g_{\nu\sigma} &= \delta_\sigma^\mu + \frac{N^\mu N_\sigma}{N^2}, \\ g^{\mu\nu} g_{\mu\nu} &= n - 1 + \frac{N^\mu N_\mu}{N^2}. \end{aligned} \quad (24.12)$$

Параметризация репера

В аффинной геометрии вместо метрики $g_{\alpha\beta}$ и аффинной связности $\Gamma_{\alpha\beta}^\gamma$ удобно использовать переменные Картана: репер e_α^a и $\mathbb{GL}(n, \mathbb{R})$ связность $\omega_{\alpha a}^b$, где $a, b, \dots = 0, 1, \dots, n-1$ (см. раздел 5.4). Мы ограничимся рассмотрением геометрии Римана–Картана. Тогда линейная $\mathbb{GL}(n, \mathbb{R})$ связность сужается до лоренцевой $\mathbb{SO}(1, n-1)$ связности. Здесь и в дальнейшем мы примем следующее обозначение. Если латинский индекс a принимает значение нуль, то мы будем писать его курсивом: $a = 0$. Это правило принято потому что в дальнейшем значение нуль будут принимать как греческие, так и латинские индексы, что необходимо различать. Выделим в ортонормальном базисе касательного пространства временной вектор e_0 , $\{e_a\} = \{e_0, e_i\}$, тогда метрика Минковского примет вид

$$\eta_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & \eta_{ij} \end{pmatrix}, \quad i, j = 1, \dots, n-1,$$

где $\eta_{ij} = -\delta_{ij}$ – евклидова метрика с обратным знаком для пространственных компонент тензоров. Как и раньше, мы принимаем следующее соглашение. Латинские индексы из начала алфавита $a, b, \dots$ пробегают все значения, а индексы из середины алфавита $i, j, \dots$ – только пространственные. Поскольку репер содержит большее число независимых компонент, чем метрика, то его параметризация, соответствующая выделенной временной координате, является более громоздкой:

$$e_\alpha^a = \begin{pmatrix} N\sqrt{1-M^2} + (MN) & N^i - NM^i + \frac{\sqrt{1-M^2}-1}{M^2}(MN)M^i \\ M_\mu & \hat{e}_\mu^i + \frac{\sqrt{1-M^2}-1}{M^2}M_\mu M^i \end{pmatrix}, \quad (24.13)$$

где M_μ – дополнительное $(n-1)$ -компонентное пространственноподобное ковекторное поле (относительно трехмерных диффеоморфизмов). Ниже будет показано, что это поле пропорционально трехмерному вектору скорости, параметризующему преобразования Лоренца. Репер со шляпкой в (24.13) $\hat{e}_\mu^i$ на сечении $x^0 = \text{const}$ определяется трехмерной метрикой:

$$g_{\mu\nu} = \hat{e}_\mu^i \hat{e}_\nu^j \eta_{ij}, \quad (24.14)$$

В параметризации репера (24.13) использованы также обозначения $(MN) := M^\mu N_\mu = M^i N_i$ и $M^2 := M^\mu M_\mu = M^i M_i < 0$. Переход от пространственных латинских индексов к пространственным греческим индексам осуществляется с помощью репера $\hat{e}_\mu^i$ и его обратного $\hat{e}^\mu_i$, которые мы пометили шляпкой,

$$\hat{g}^{\mu\nu} = \hat{e}^\mu_i \hat{e}^\nu_j \eta^{ij}.$$

Например, $M^i := M_\mu \hat{e}^{\mu i}$. Как и для метрики, реперы со шляпкой и без нее необходимо различать: $\hat{e}_\mu^i \neq e_\mu^i$, $\hat{e}^\mu_i \neq e^\mu_i$. Из определения репера со шляпкой (24.14) следует равенство для определителей

$$\det \hat{e}_\mu^i = \sqrt{|\det g_{\mu\nu}|},$$

что согласуется с обозначением в (24.11).

С помощью прямых вычислений нетрудно проверить, что приведенная параметризация репера соответствует АДМ параметризации метрики (24.5)

$$N^2 + N^\rho N_\rho = e_0^a e_{0a}, \quad N_\mu = e_0^a e_{\mu a}, \quad g_{\mu\nu} = e_\mu^a e_{\nu a}.$$

Обратный репер имеет вид

$$e^\alpha_a = \begin{pmatrix} \frac{\sqrt{1-M^2}}{N} & -\frac{M_i}{N} \\ M^\mu - \frac{\sqrt{1-M^2}}{N}N^\mu & \hat{e}^\mu_i + \frac{N^\mu M_i}{N} + \frac{\sqrt{1-M^2}-1}{M^2}M^\mu M_i \end{pmatrix} \quad (24.15)$$

Чтобы понять геометрический смысл параметров M_i , рассмотрим преобразования касательного пространства. Связная компонента единицы группы Лоренца $\mathbb{SO}_0(1, n-1)$ имеет подгруппу вращений $\mathbb{SO}(n-1)$, действующую на пространственные компоненты репера $\hat{e}_\mu^i$. Эта подгруппа параметризуется $(n-1)(n-2)/2$ параметрами. Остальные $n-1$ параметров соответствуют лоренцевым бустам, которые параметризуются функциями M_i . В этом нетрудно убедиться. Представим репер (24.13) в виде

$$e_\alpha^a = \overset{\circ}{e}_\alpha^b S_b^a,$$

где

$$\overset{\circ}{e}_\alpha^a = \begin{pmatrix} N & N^i \\ 0 & \hat{e}_\mu^i \end{pmatrix} \quad (24.16)$$

– фиксированный репер, который преобразуется с помощью лоренцевых бустов

$$S_b^a = \begin{pmatrix} \sqrt{1-M^2} & -M^j \\ M_i & S_i^j \end{pmatrix} \in \mathbb{SO}_0(1, n-1), \quad (24.17)$$

где введено обозначение для пространственной части матрицы лоренцевых вращений

$$S_i^j = \delta_i^j + \frac{\sqrt{1-M^2}-1}{M^2}M_i M^j \quad (24.18)$$

Можно проверить, что матрица (24.17) действительно задает преобразования Лоренца. Репер, обратный к (24.16), имеет вид

$$\overset{\circ}{e}^\alpha_a = \begin{pmatrix} \frac{1}{N} & 0 \\ -\frac{N^\mu}{N} & \hat{e}^\mu_i \end{pmatrix}.$$

Формулы для фиксированного репера $\overset{\circ}{e}_\alpha^a$ и его обратного $\overset{\circ}{e}^\alpha_a$ получаются из (24.13) и (24.15) при $M_i = 0$.

Матрица пространственных вращений (24.18) часто используется в вычислениях, поэтому приведем для нее несколько полезных соотношений, которые проверяются прямой проверкой,

$$\begin{aligned} S^{-1}{}_i{}^j &= \delta_i^j - \frac{\sqrt{1-M^2}-1}{M^2\sqrt{1-M^2}}M_i M^j \\ S_i^j(M) &= S_i^j(-M), \quad S_{ij} = S_{ji}, \\ S_i^j M_j &= \sqrt{1-M^2}M_i, \\ S_i^k S_{jk} &= \eta_{ij} - M_i M_j, \\ S^{-1}{}_i{}^k S^{-1}{}_{jk} &= \eta_{ij} + \frac{M_i M_j}{\sqrt{1-M^2}}. \end{aligned}$$

Таким образом, вместо n^2 независимых компонент репера мы ввели функцию сдвига N (1 компонента), функции хода N^i ($n - 1$ компонент), лоренцевы бусты M_i ($n - 1$ компонент) и пространственный репер $\hat{e}_\mu^i$ ($(n - 1)^2$ компонент). Эта параметризация при $N > 0$ является взаимно однозначной.

Отметим, что единичный нормальный вектор к сечению $x^0 = \text{const}$ (23.165) относительно неголономного базиса $e_a = e^\mu_a \partial_\mu$ имеет вид

$$n = \sqrt{1 - M^2} e_0 - M^i e_i.$$

То есть его компоненты относительно неголономного базиса полностью определяются лоренцевыми бустами.

Приведем более привычную параметризацию лоренцевых вращений S_a^b с помощью вектора скорости. Вектор M^i взаимно однозначно параметризуется вектором скорости V^i :

$$M^i = \frac{V^i}{\sqrt{1 - \mathbf{V}^2}} \quad \Leftrightarrow \quad V^i = \frac{M^i}{\sqrt{1 - M^2}},$$

где $\mathbf{V}^2 := -V^i V_i$. По определению, $M^2 < 0$ и, следовательно, $0 < \mathbf{V}^2 < 1$. Тогда

$$S_a^b = \begin{pmatrix} \frac{1}{\sqrt{1 - \mathbf{V}^2}} & -\frac{V^j}{\sqrt{1 - \mathbf{V}^2}} \\ \frac{V_i}{\sqrt{1 - \mathbf{V}^2}} & \delta_i^j - \frac{V_i V^j}{\mathbf{V}^2} \left(\frac{1}{\sqrt{1 - \mathbf{V}^2}} - 1 \right) \end{pmatrix}.$$

Эта матрица для лоренцевых бустов была получена ранее (1.216).

Замечание. В римановом пространстве с положительно определенной метрикой репер можно параметризовать похожим образом, заменив матрицу лоренцевых вращений на ортогональную матрицу. $\square$

24.3 Геометрия гиперповерхностей

При гамильтоновой формулировке моделей гравитации мы рассматриваем пространство-время как семейство пространственноподобных гиперповерхностей, которое параметризуется временем. Другими словами, для каждого момента времени пространство представляет собой гиперповерхность, вложенную в пространство-время. Поскольку уравнения моделей гравитации определяют геометрию всего пространства-времени, то полезно знать, какая геометрия возникает при этом на пространственноподобных сечениях. В настоящем разделе мы подойдем к этому вопросу с общей точки зрения, предполагая, что на объемлющем многообразии задана аффинная геометрия общего вида, т.е. независимые метрика и связность, не предполагая лоренцевой сигнатуры метрики.

Рассмотрим $(n - 1)$ -мерную гиперповерхность $\mathbb{U}$, вложенную в n -мерное многообразие $\mathbb{M}$:

$$f : \quad \mathbb{U} \hookrightarrow \mathbb{M}. \quad (24.19)$$

Обозначим координаты на $\mathbb{M}$ и $\mathbb{U}$, соответственно, через x^α , $\alpha = 1, \dots, n$, и u^i , $i = 1, \dots, n - 1$. Тогда вложение $\mathbb{U}$ в $\mathbb{M}$ локально задается n функциями $x^\alpha(u)$, которые предполагаются достаточно гладкими. Произвольное векторное поле $X = \{X^i\} \in$

$\mathcal{X}(\mathbb{U})$, на гиперповерхности, отображается в векторное поле $Y = \{Y^\alpha\} \in \mathcal{X}(\mathbb{M})$ на $\mathbb{M}$ с помощью дифференциала отображения

$$f_* : \mathcal{X}(\mathbb{U}) \ni X = X^i \partial_i \mapsto Y = Y^\alpha \partial_\alpha \in \mathcal{X}(\mathbb{M}),$$

где

$$Y^\alpha := e^\alpha_i X^i, \quad e^\alpha_i := \partial_i x^\alpha.$$

Матрица Якоби e^α_i отображения f прямоугольна, имеет размер $n \times (n-1)$, ранг $n-1$ и, естественно, необратима. Возврат отображения f каждому ковекторному полю (1-форме) на $\mathbb{M}$ ставит в соответствие ковекторное поле на $\mathbb{U}$:

$$f^* : \Lambda_1(\mathbb{M}) \ni A = dx^\alpha A_\alpha \mapsto B = du^i B_i \in \Lambda_1(\mathbb{U}),$$

где

$$B_i := A_\alpha e^\alpha_i.$$

В дальнейшем мы будем отождествлять гиперповерхность $\mathbb{U}$ с ее образом: $\mathbb{U} = f(\mathbb{U}) \subset \mathbb{M}$.

1-форма $n = dx^\alpha n_\alpha$, определяемая системой алгебраических уравнений

$$n_\alpha e^\alpha_i = 0, \quad i = 1, \dots, n-1, \quad (24.20)$$

ортогональна гиперповерхности $\mathbb{U}$ и задает на многообразии $\mathbb{M}$ распределение $(n-1)$ -мерных подпространств, касательных к $\mathbb{U}$. Уравнения (24.20) имеют единственное нетривиальное решение с точностью до умножения на произвольную отличную от нуля функцию, поскольку из определения вложения следует, что ранг матрицы Якоби равен $n-1$.

Матрица Якоби e^α_i задает в касательном пространстве $\mathbb{T}_x(\mathbb{M})$, где $x \in \mathbb{U} \subset \mathbb{M}$, набор из $n-1$ векторов $e_i = e^\alpha_i \partial_\alpha$, которые образуют базис касательного пространства к гиперповерхности $\mathbb{U} \subset \mathbb{M}$.

Это все, что можно сказать о гиперповерхности $\mathbb{U}$, если задано только вложение (24.19). Теория становится намного более содержательной, если на многообразии $\mathbb{M}$ заданы дополнительные структуры. Остановимся на этом вопросе подробно.

Пусть на $\mathbb{M}$ задана аффинная геометрия, т.е. метрика $g_{\alpha\beta}$ и аффинная связность $\Gamma_{\alpha\beta}^\gamma$. Рассмотрим, какая геометрия возникает на гиперповерхности $\mathbb{U} \hookrightarrow \mathbb{M}$. Возврат отображения f^* индуцирует на гиперповерхности единственную метрику:

$$f^* : g_{\alpha\beta} \mapsto g_{ij} := g_{\alpha\beta} e^\alpha_i e^\beta_j. \quad (24.21)$$

Для многообразий с метрикой лоренцевой сигнатуры мы предполагаем, что гиперповерхность $\mathbb{U}$ пространственноподобна и, следовательно, индуцированная метрика невырождена и отрицательно определена. Наличие двух метрик $g_{\alpha\beta}$ и g_{ij} , соответственно, на $\mathbb{M}$ и $\mathbb{U}$ позволяет опускать и поднимать индексы у матрицы Якоби:

$$e_\alpha^i := g_{\alpha\beta} e^\beta_j g^{ij}.$$

Эта матрица проектирует произвольный вектор из $\mathbb{T}_{f(u)}(\mathbb{M})$ в касательное пространство к гиперповерхности $\mathbb{T}_u(\mathbb{U})$

$$\mathbb{T}_{f(u)}(\mathbb{M}) \ni \{X^\alpha\} \mapsto \{X^i := X^\alpha e_\alpha^i\} \in \mathbb{T}_u(\mathbb{U}). \quad (24.22)$$

Теперь определим связность на гиперповерхности $\mathbb{U} \hookrightarrow \mathbb{M}$, спроектировав ковариантную производную в $\mathbb{M}$ на гиперповерхность $\mathbb{U}$:

$$\hat{\nabla}_i X^k := (\nabla_\alpha X^\beta) e^\alpha_i e_\beta^k, \quad (24.23)$$

где компоненты X^i определены отображением (24.22).

Замечание. Из приведенного определения нельзя выразить ковариантную производную $\nabla_\alpha X^\beta$ в $\mathbb{M}$ через ковариантную производную $\hat{\nabla}_i X^k$ на $\mathbb{U}$, т.к. оригинал нельзя восстановить по его проекции. $\square$

Раскрытие равенства (24.23) приводит к выражению для компонент индуцированной связности на гиперповерхности $\mathbb{U}$:

$$\hat{\Gamma}_{ij}{}^k = (\partial_{ij}^2 x^\gamma + \Gamma_{\alpha\beta}{}^\gamma e^\alpha{}_i e^\beta{}_j) e_\gamma{}^k. \quad (24.24)$$

Эта связность единственна. Отметим, что если исходная связность $\Gamma_{\alpha\beta}{}^\gamma$ симметрична, то индуцированная связность $\hat{\Gamma}_{ij}{}^k$ также симметрична. Кроме того, связность на $\mathbb{U}$ определяется единственным образом только в том случае, если на $\mathbb{M}$ помимо связности задана также метрика.

Из уравнения (24.24) следует, что тензор кручения на $\mathbb{M}$ индуцирует кручение на гиперповерхности:

$$T_{ij}{}^k = T_{\alpha\beta}{}^\gamma e^\alpha{}_i e^\beta{}_j e_\gamma{}^k. \quad (24.25)$$

Прямые вычисления дают следующее выражение для ковариантной производной от индуцированной метрики на гиперповерхности

$$\hat{\nabla}_i g_{jk} = \partial_i g_{jk} - \hat{\Gamma}_{ij}{}^l g_{lk} - \hat{\Gamma}_{ik}{}^l g_{jl} = (\nabla_\alpha g_{\beta\gamma}) e^\alpha{}_i e^\beta{}_j e_\gamma{}^k.$$

Отсюда вытекает выражение для тензора неметричности на гиперповерхности

$$Q_{ijk} = Q_{\alpha\beta\gamma} e^\alpha{}_i e^\beta{}_j e_\gamma{}^k.$$

В частности, если связность $\Gamma_{\alpha\beta}{}^\gamma$ на $\mathbb{M}$ является метрической, то и индуцированная связность $\hat{\Gamma}_{ij}{}^k$ на $\mathbb{U}$ также метрическая.

Таким образом, метрика $g_{\alpha\beta}$ и связность $\Gamma_{\alpha\beta}{}^\gamma$ на $\mathbb{M}$ индуцируют единственные метрику g_{ij} и связность $\hat{\Gamma}_{ij}{}^k$ на гиперповерхности $\mathbb{U} \subset \mathbb{M}$. Обратное утверждение, конечно, неверно. Если метрика и связность заданы на гиперповерхности $\mathbb{U}$, то они не индуцируют на $\mathbb{M}$ геометрию единственным образом. Это понятно, поскольку размерность гиперповерхности меньше размерности самого многообразия.

В дальнейшем все геометрические объекты, относящиеся только к гиперповерхности и построенные по индуцированной метрике g_{ij} и связности $\hat{\Gamma}_{ij}{}^k$ мы будем отмечать шляпкой.

Наличие метрики $g_{\alpha\beta}$ позволяет построить единичное векторное поле $n = n^\alpha \partial_\alpha$, ортогональное к гиперповерхности. Как уже отмечалось, система уравнений $n_\alpha e^\alpha{}_i = 0$ определяет 1-форму $dx^\alpha n_\alpha$ с точностью до умножения на произвольную функцию. Воспользуемся этим произволом для того, чтобы в каждой точке вектор с компонентами $n^\alpha := g^{\alpha\beta} n_\beta$ имел единичную длину $(n, n) = n^\alpha n^\beta g_{\alpha\beta} = 1$. По построению, этот вектор ортогонален всем векторам, касательным к гиперповерхности:

$$(n, e_i) = n^\alpha e^\beta{}_i g_{\alpha\beta} = n_\alpha e^\alpha{}_i = 0.$$

Если на многообразии задана гиперповерхность, то естественно рассматривать базис $\{n, e_i\}$ в касательном пространстве $\mathbb{T}_x(\mathbb{M})$, $x \in \mathbb{U}$, определяемый этой гиперповерхностью. Ему соответствует сопряженный базис $\{n := dx^\alpha n_\alpha, e^i := dx^\alpha e_\alpha{}^i\}$ в кокасательном пространстве $\mathbb{T}_x^*(\mathbb{M})$. Тогда произвольный вектор X и 1-форма A раскладываются по этим базисам:

$$\begin{aligned} X^\alpha &= X^\perp n^\alpha + X^i e^\alpha{}_i, & X^\perp &:= X^\alpha n_\alpha, & X^i &:= X^\alpha e_\alpha{}^i, \\ A_\alpha &= A_\perp n_\alpha + A_i e_\alpha{}^i, & A_\perp &:= A_\alpha n^\alpha, & A_i &:= A_\alpha e^\alpha{}_i. \end{aligned}$$

Аналогично можно разложить тензор произвольного ранга. В частности, разложение ковариантного тензора второго ранга имеет вид

$$A_{\alpha\beta} = A_{\perp\perp} n_\alpha n_\beta + A_{\perp i} n_\alpha e_\beta^i + A_{i\perp} e_\alpha^i n_\beta + A_{ij} e_\alpha^i e_\beta^j,$$

где

$$A_{\perp\perp} = A_{\alpha\beta} n^\alpha n^\beta, \quad A_{\perp i} = A_{\alpha\beta} n^\alpha e_\beta^i, \quad A_{i\perp} = A_{\alpha\beta} e_\alpha^i n_\beta, \quad A_{ij} = A_{\alpha\beta} e_\alpha^i e_\beta^j.$$

Нетрудно проверить, что разложение для метрики гораздо проще:

$$g_{\alpha\beta} = n_\alpha n_\beta + e_\alpha^i e_\beta^j g_{ij}. \quad (24.26)$$

Разложение для обратной метрики имеет аналогичный вид:

$$g^{\alpha\beta} = n^\alpha n^\beta + e^\alpha_i e^\beta_j g^{ij}. \quad (24.27)$$

Из определения обратной метрики $g^{\alpha\beta} g_{\beta\gamma} = \delta_\gamma^\alpha$ следует правило суммирования матрицы Якоби по латинским индексам:

$$e^\alpha_i e_\beta^i = \delta_\beta^\alpha - n^\alpha n_\beta. \quad (24.28)$$

Нетрудно также проверить, что из определения обратной индуцированной метрики $g^{ij} g_{jk} = \delta_k^i$ следует равенство

$$e^\alpha_i e_\alpha^j = \delta_i^j, \quad (24.29)$$

где суммирование проводится по греческим индексам. С учетом этого правила из равенства (24.27) следует представление для обратной индуцированной метрики

$$g^{ij} = g^{\alpha\beta} e_\alpha^i e_\beta^j.$$

Индуцированные метрика (24.21) и связность (24.24) задают внутреннюю геометрию гиперповерхности $\mathbb{U} \hookrightarrow \mathbb{M}$. В частности, индуцированная связность задает тензор внутренней кривизны гиперповерхности

$$\hat{R}_{ijk}{}^l(\hat{\Gamma}) = \partial_i \hat{\Gamma}_{jk}{}^l - \hat{\Gamma}_{ik}{}^m \hat{\Gamma}_{jm}{}^l - (i \leftrightarrow j).$$

Вложение гиперповерхности $f : \mathbb{U} \hookrightarrow \mathbb{M}$ позволяет определить еще один важный объект, характеризующий то, как гиперповерхность $\mathbb{U}$ вложена в $\mathbb{M}$.

Определение. Тензор второго ранга с компонентами

$$K_{ij} := -(\nabla_\alpha n_\beta) e^\alpha_i e^\beta_j, \quad (24.30)$$

определен на $\mathbb{U} \subset \mathbb{M}$ и называется *внешней кривизной* гиперповерхности. Он равен ковариантной производной нормали, спроектированной на касательное пространство к гиперповерхности и взятой с обратным знаком. $\square$

В отличие от тензора внутренней кривизны внешняя кривизна является тензором второго, а не четвертого ранга и в общем случае никакой симметрии по индексам не имеет. Этот тензор характеризует изменение нормали при ее параллельном переносе вдоль кривой на гиперповерхности.

Раскрывая определение тензора внешней кривизны, с учетом равенства (24.20) получим, что

$$K_{ij} = n_\alpha (\partial_{ij}^2 x^\alpha + \Gamma_{\beta\gamma}^\alpha e^\beta_i e^\gamma_j).$$

Отсюда следует, что антисимметричная часть тензора внешней кривизны определяется тензором кручения

$$K_{[ij]} = \frac{1}{2} n_\alpha T_{\beta\gamma}^\alpha e^\beta_i e^\gamma_j = \frac{1}{2} T_{ij}^\perp. \quad (24.31)$$

Это доказывает следующее

Предложение 24.3.1. *Внешняя кривизна гиперповерхности $\mathbb{U} \hookrightarrow \mathbb{M}$ симметрична тогда и только тогда, когда сужение кручения связности $\Gamma_{\beta\gamma}^\alpha$ на $\mathbb{U} \subset \mathbb{M}$ удовлетворяет условию $T_{ij}^\perp = 0$.*

Вычислим ковариантную производную от матрицы Якоби

$$\nabla_i e^\alpha_j = e^\beta_i (\partial_\beta e^\alpha_j + \Gamma_{\beta\gamma}^\alpha e^\gamma_j) - \hat{\Gamma}_{ij}^k e^\alpha_k. \quad (24.32)$$

Разложение этого равенства (индекс α) по базису n, e_i показывает, что эта ковариантная производная имеет только нормальную составляющую и пропорциональна внешней кривизне:

$$\nabla_i e^\alpha_j = n^\alpha K_{ij}. \quad (24.33)$$

Введем специальное обозначение для “половины” ковариантной производной (24.32), которая содержит связность только для греческих индексов,

$$\tilde{\nabla}_i e^\alpha_j := e^\beta_i (\partial_\beta e^\alpha_j + \Gamma_{\beta\gamma}^\alpha e^\gamma_j).$$

Она ковариантна относительно преобразований координат x^α в $\mathbb{M}$, но не на $\mathbb{U}$. Разлагая ее по базису n, e_i , получим равенство

$$\tilde{\nabla}_i e^\alpha_j = n^\alpha K_{ij} + \hat{\Gamma}_{ij}^k e^\alpha_k. \quad (24.34)$$

Это соотношение известно как *формула Гаусса–Вейнгартена*.

Из равенств (24.33) и (24.34) следует ещё одно представление для тензора внешней кривизны

$$K_{ij} = n_\alpha \nabla_i e^\alpha_j = n_\alpha \tilde{\nabla}_i e^\alpha_j. \quad (24.35)$$

Полный тензор кривизны $R_{\alpha\beta\gamma\delta}$, спроектированный на гиперповерхность, можно выразить через тензор внутренней кривизны $\hat{R}_{ijkl}$, построенный только по индуцированной метрике (24.21) и связности (24.24), и тензор внешней кривизны. Для этого рассмотрим коммутатор ковариантных производных ковекторного поля, который определяется тензором кривизны и кручения (6.94)

$$\begin{aligned} [\nabla_\alpha, \nabla_\beta] X_\gamma &= -R_{\alpha\beta\gamma}^\delta X_\delta - T_{\alpha\beta}^\delta \nabla_\delta X_\gamma = \\ &= -R_{\alpha\beta\gamma i} X^i - R_{\alpha\beta\gamma\perp} X^\perp - T_{\alpha\beta}^i \nabla_i X_\gamma - T_{\alpha\beta}^\perp \nabla_\perp X_\gamma, \end{aligned} \quad (24.36)$$

где в правой части сначала вычисляются ковариантные производные, а затем проектируются на гиперповерхность и ортогональное направление:

$$\nabla_i X_\gamma := e^\alpha_i \nabla_\alpha X_\gamma, \quad \nabla_\perp X_\gamma := n^\alpha \nabla_\alpha X_\gamma.$$

Для того, чтобы спроектировать равенство (24.36) на гиперповерхность, спроектируем сначала ковариантную производную:

$$\nabla_i X_j := e^\alpha_i (\nabla_\alpha X_\beta) e^\beta_j = e^\alpha_i \nabla_\alpha (X_\beta e^\beta_j) - e^\alpha_i X_\beta \nabla_\alpha e^\beta_j = \hat{\nabla}_i X_j - X_\perp K_{ij},$$

где мы воспользовались равенством (24.33) и

$$\hat{\nabla}_i X_j := \partial_i X_j - \hat{\Gamma}_{ij}^k X_k$$

– $(n-1)$ -мерная ковариантная производная на гиперповерхности. Аналогично проектируется вторая ковариантная производная:

$$\begin{aligned} \nabla_i \nabla_j X_k &= \hat{\nabla}_i (\nabla_j X_k) - \nabla_\perp X_k K_{ij} - \nabla_j X_\perp K_{ik} = \\ &= \hat{\nabla}_i \hat{\nabla}_j X_k - \hat{\nabla}_i X_\perp K_{jk} - X_\perp \hat{\nabla}_i K_{jk} - \nabla_\perp X_k K_{ij} - \hat{\nabla}_j X_\perp K_{ik} - X^l K_{lj} K_{ik}, \end{aligned}$$

где

$$\nabla_j X_\perp := e^\alpha_j (\nabla_\alpha X_\beta) n^\beta = \hat{\nabla}_j X_\perp + X^l K_{lj}.$$

Антисимметризация полученного выражения по индексам i, j дает проекцию коммутатора (24.36) на гиперповерхность:

$$\begin{aligned} [\nabla_i, \nabla_j] X_k &= -R_{ijkl} X^l - R_{ijk\perp} X^\perp - T_{ij}^l \nabla_l X_k - T_{ij}^\perp \nabla_\perp X_k = \\ &= -R_{ijkl} X^l - R_{ijk\perp} X^\perp - T_{ij}^l \hat{\nabla}_l X_k + T_{ij}^l X_\perp K_{lk} - T_{ij}^\perp \nabla_\perp X_k. \end{aligned}$$

Учитывая независимость компонент X^l и $X^\perp$ и выражения для компонент тензора кручения (24.25), (24.31), получаем выражение для проекций полного тензора кривизны на гиперповерхность:

$$R_{ijkl} = \hat{R}_{ijkl} + K_{ik} K_{lj} - K_{jk} K_{li}, \quad (24.37)$$

$$R_{ijk\perp} = \hat{\nabla}_i K_{jk} - \hat{\nabla}_j K_{ik} + T_{ij}^l K_{lk}. \quad (24.38)$$

Полученные соотношения являются обобщением *уравнений Гаусса–Петерсона–Кодацци* на случай, когда на объемлющем многообразии $\mathbb{M}$ задана не риманова геометрия, а произвольная аффинная геометрия с ненулевым кручением и неметричностью.

В заключение настоящего раздела вычислим нормальные компоненты $G_{\perp\perp}$ и $G_{\perp i}$ тензора Эйнштейна

$$G_{\alpha\beta} := \tilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \tilde{R},$$

которые играют важную роль при анализе уравнений движения общей теории относительности. Знак тильды в правой части, как и ранее, означает, что кручение и неметричность связности, заданной на $\mathbb{M}$, равны нулю. В этом случае тензор кривизны обладает дополнительной симметрией относительно перестановки индексов (см. раздел 6.8) и тензор внешней кривизны симметричен: $K_{ij} = K_{ji}$.

Ниже, для простоты, мы опустим знак тильды.

Сначала вычислим скалярную кривизну

$$R = g^{\alpha\gamma} g^{\beta\delta} R_{\alpha\beta\gamma\delta} = 2R_{\perp\perp} + g^{ik} g^{jl} R_{ijkl},$$

где $R_{\perp\perp} = g^{ij} R_{i\perp j\perp}$ – нормальная составляющая тензора Риччи и учтено представление для обратной метрики (24.27). С учетом уравнений Гаусса–Петерсона–Кодацци (24.37) получаем равенство

$$g^{ik} g^{jl} R_{ijkl} = \hat{R} + K^2 - K^{ij} K_{ij},$$

где $\hat{R}$ – скалярная внутренняя кривизна гиперповерхности, $K := g^{ij}K_{ij}$ – скалярная внешняя кривизна гиперповерхности. Отсюда следуют выражения для нормальных компонент тензора Эйнштейна:

$$\begin{aligned} G_{\perp\perp} &= -\frac{1}{2}(\hat{R} + K^2 - K^{ij}K_{ij}), \\ G_{\perp i} &= \hat{\nabla}_j K_i^j - \hat{\nabla}_i K. \end{aligned} \quad (24.39)$$

Важным обстоятельством является то, что эти компоненты тензора Эйнштейна вообще не содержат нормальных производных к гиперповерхности $\nabla_\perp$ от индуцированной метрики и тензора внешней кривизны. На гамильтоновом языке это означает отсутствие производных по времени, и что вакуумные уравнения Эйнштейна $G_{\perp\perp} = 0$ и $G_{\perp i} = 0$ представляют собой связи, поскольку компоненты тензора внешней кривизны K^{ij} , как будет показано в следующем разделе, пропорциональны импульсам, канонически сопряженным компонентам индуцированной метрики g_{ij} .

24.4 Кривизна в АДМ параметризации метрики

АДМ параметризация метрики (24.5) удобна для канонической формулировки общей теории относительности, в которой исходными независимыми переменными являются компоненты метрики $g_{\alpha\beta}$ (обобщенные координаты) и канонически сопряженные импульсы $p^{\alpha\beta}$. Однако, чтобы перейти от лагранжиана к гамильтониану, необходимо произвести довольно громоздкие вычисления, чем мы и займемся в настоящем разделе.

Для существенного упрощения вычислений следует воспользоваться результатами раздела 24.3, в котором описана геометрия гиперповерхностей. Будем считать, что пространство-время является топологическим произведением $\mathbb{M} \approx \mathbb{R} \times \mathbb{U}$, где первый сомножитель соответствует времени x^0 . Тогда сечения $x^0 = \text{const}$ пространства-времени $\mathbb{M}$ задают семейство гиперповерхностей $\mathbb{U} \subset \mathbb{M}$, которые, по предположению, являются пространственноподобными. В качестве координат на гиперповерхностях выберем пространственные координаты:

$$\{u^i\} \mapsto \{x^\mu\}.$$

При этом мы теряем возможность независимого преобразования координат в пространстве-времени $\mathbb{M}$ и пространственноподобной гиперповерхности $\mathbb{U}$, зато многие формулы упрощаются.

Матрица Якоби вложения гиперповерхности в рассматриваемом случае имеет вид

$$\{e^\alpha_i\} \mapsto \{0_\mu, \delta^\nu_\mu\}, \quad \{e^\alpha_i\} \mapsto \{N^\mu, \delta^\mu_\nu\},$$

где символ 0_μ обозначает строку, состоящую из $n - 1$ нулей. Вложение индуцирует на гиперповерхностях метрику $g_{\mu\nu}$ согласно формуле (24.21).

Построим векторное поле $n = n^\alpha \partial_\alpha$, ортогональное семейству пространственноподобных гиперповерхностей $x^0 = \text{const}$. Из условия ортогональности $(n, X) = 0$, где $X = X^\mu \partial_\mu$ – произвольный вектор, касательный к сечению $x^0 = 0$, следует, что

$$n = n^0(\partial_0 - N^\mu \partial_\mu),$$

где n^0 – произвольная отличная от нуля функция. Если, кроме того, положить $n^0 = 1/N$, то длина нормального вектора будет равна единице, $n^2 = 1$. Таким образом,

единичный вектор, нормальный к сечениям $x^0 = \text{const}$ имеет вид

$$n = \frac{1}{N}(\partial_0 - N^\mu \partial_\mu) \quad (24.40)$$

и всегда является времениподобным. Ему соответствует ортонормальная 1-форма

$$n = dx^0 N. \quad (24.41)$$

Заметим, что если функции хода равны нулю, $N_\mu = 0$, то нормальный вектор пропорционален вектору, касательному к оси времени, $n \sim \partial_0$.

Произвольный тензор на $\mathbb{M}$ можно разложить по базису $\{n, e_\mu := \partial_\mu\}$. В частности, для векторов и 1-форм справедливы разложения:

$$X^\alpha = X^\perp n^\alpha + \tilde{X}^\mu e_\mu^\alpha, \quad X_\alpha = X_\perp n_\alpha + \tilde{X}_\mu e_\alpha^\mu,$$

где

$$\begin{aligned} X^\perp &:= X^0 N, & \tilde{X}^\mu &:= X^0 N^\mu + X^\mu, \\ X_\perp &:= \frac{1}{N}(X_0 - N^\mu X_\mu), & \tilde{X}_\mu &:= X_\mu. \end{aligned} \quad (24.42)$$

Нетрудно проверить, что

$$X^\perp = X_\perp \quad \text{и} \quad \tilde{X}^\mu = \hat{g}^{\mu\nu} \tilde{X}_\nu.$$

Представления для метрики (24.26) всего пространства-времени и ее обратной (24.27) принимает вид

$$\begin{aligned} g_{\alpha\beta} &= n_\alpha n_\beta + g_{\mu\nu} e_\alpha^\mu e_\beta^\nu, \\ g^{\alpha\beta} &= n^\alpha n^\beta + \hat{g}^{\mu\nu} e_\mu^\alpha e_\nu^\beta. \end{aligned} \quad (24.43)$$

Обозначим тензорные индексы по отношению к базису $\{n, e_\mu\}$ латинскими буквами $a, b, \dots$. Тогда они принимают значения $\{a\} = \{\perp, \mu\} = \{\perp, 1, 2, 3\}$. В этом базисе метрика и ее обратная имеют блочно-диагональный вид:

$$g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix}, \quad g^{ab} = \begin{pmatrix} 1 & 0 \\ 0 & \hat{g}^{\mu\nu} \end{pmatrix}.$$

Слово “неголономный” означает, что в общем случае такой системы координат, в которой базис n, ∂_μ был бы координатным, не существует.

Связность (24.24), индуцированная на гиперповерхностях – это символы Кристоффеля $\hat{\Gamma}_{\mu\nu}^\rho$, построенные по пространственной метрике $g_{\mu\nu}$.

Тензор внешней кривизны (24.30) гиперповерхности $x^0 = \text{const}$ в АДМ параметризации метрики имеет вид

$$K_{\mu\nu} = \Gamma_{\mu\nu}^0 N = \frac{1}{2N}(\hat{\nabla}_\mu N_\nu + \hat{\nabla}_\nu N_\mu - \dot{g}_{\mu\nu}), \quad (24.44)$$

где точка обозначает дифференцирование по времени, $\dot{g}_{\mu\nu} := \partial_0 g_{\mu\nu}$, и $\hat{\nabla}_\mu N_\nu := \partial_\mu N_\nu - \hat{\Gamma}_{\mu\nu}^\rho N_\rho$ – пространственная ковариантная производная. Тензор внешней кривизны симметричен, $K_{\mu\nu} = K_{\nu\mu}$, поскольку кручение в общей теории относительности равно нулю. В дальнейшем нам понадобится также след тензора внешней кривизны

$$K := K_\mu{}^\mu = \hat{g}^{\mu\nu} K_{\mu\nu} = \frac{1}{2N}(2\hat{\nabla}_\mu N^\mu - \hat{g}^{\mu\nu} \dot{g}_{\mu\nu}). \quad (24.45)$$

При вычислении тензора кривизны $R_{\alpha\beta\gamma\delta}$ пространства-времени $\mathbb{M}$ все производные по времени от пространственной части метрики $\dot{g}_{\mu\nu}$ удобно выражать через $K_{\mu\nu}$. Кроме этого, для исключения вторых производных по времени $\ddot{g}_{\mu\nu}$ нам понадобится производная по времени от тензора внешней кривизны,

$$\dot{K}_{\mu\nu} = \frac{1}{2N} \left[\hat{\nabla}_\mu \dot{N}_\nu + \hat{\nabla}_\nu \dot{N}_\mu - \ddot{g}_{\mu\nu} - N^\rho (\hat{\nabla}_\mu \dot{g}_{\nu\rho} + \hat{\nabla}_\nu \dot{g}_{\mu\rho} - \hat{\nabla}_\rho \dot{g}_{\mu\nu}) \right] - \frac{\dot{N}}{N} K_{\mu\nu},$$

где

$$\hat{\nabla}_\mu \dot{N}_\nu = \partial_\mu \dot{N}_\nu - \hat{\Gamma}_{\mu\nu}{}^\rho \dot{N}_\rho$$

и куда при вычислениях мы должны подставить выражение для $\dot{g}_{\mu\nu}$ через $K_{\mu\nu}$.

Приступим к вычислению тензора кривизны $R_{\alpha\beta\gamma\delta}$ для метрики, записанной в виде (24.5). Первый шаг состоит в вычислении символов Кристоффеля (6.24). Прямые выкладки приводят к следующим выражениям для линейно независимых символов Кристоффеля

$$\begin{aligned} \Gamma_{00}{}^0 &= \frac{1}{N} \left(\dot{N} + N^\rho \partial_\rho N + N^\rho N^\sigma K_{\rho\sigma} \right), \\ \Gamma_{00}{}^\mu &= \hat{g}^{\mu\nu} (\dot{N}_\nu - N \partial_\nu N - N^\rho \hat{\nabla}_\nu N_\rho) - \frac{N^\mu}{N} \left(\dot{N} + N^\rho \partial_\rho N + N^\rho N^\sigma K_{\rho\sigma} \right), \\ \Gamma_{0\mu}{}^0 &= \frac{1}{N} (\partial_\mu N + N^\nu K_{\mu\nu}), \\ \Gamma_{0\mu}{}^\nu &= \hat{\nabla}_\mu N^\nu - N K_\mu{}^\nu - \frac{N^\nu}{N} (\partial_\mu N + N^\rho K_{\mu\rho}), \\ \Gamma_{\mu\nu}{}^0 &= \frac{1}{N} K_{\mu\nu}, \\ \Gamma_{\mu\nu}{}^\rho &= \hat{\Gamma}_{\mu\nu}{}^\rho - \frac{N^\rho}{N} K_{\mu\nu}. \end{aligned} \tag{24.46}$$

В дальнейшем нам понадобятся два независимых следа символов Кристоффеля:

$$\Gamma_\alpha := \Gamma_{\alpha\beta}{}^\beta, \quad \Xi^\alpha := g^{\beta\gamma} \Gamma_{\beta\gamma}{}^\alpha.$$

Несложные вычисления приводят к формулам

$$\begin{aligned} \Gamma_0 &= \frac{\dot{N}}{N} + \hat{\nabla}_\mu N^\mu - N K, \\ \Gamma_\mu &= \hat{\Gamma}_\mu + \frac{\partial_\mu N}{N}, \\ \Xi^0 &= \frac{1}{N} K + \frac{1}{N^3} (\dot{N} - N^\mu \partial_\mu N), \\ \Xi^\mu &= \left(\hat{g}^{\rho\sigma} + \frac{N^\rho N^\sigma}{N^2} \right) \hat{\Gamma}_{\rho\sigma}{}^\mu - \frac{N^\mu}{N} K - \frac{N^\mu}{N^3} (\dot{N} - N^\rho \partial_\rho N) + \\ &\quad + \frac{1}{N^2} \hat{g}^{\mu\rho} (\dot{N}_\rho - N \partial_\rho N - N^\sigma \hat{\nabla}_\rho N_\sigma - 2N^\sigma \hat{\nabla}_\sigma N_\rho + 2N N^\sigma K_{\rho\sigma}). \end{aligned} \tag{24.47}$$

Выпишем также формулы дифференцирования по времени для символов Кристоффеля:

$$\begin{aligned} \partial_0 \hat{\Gamma}_{\mu\nu\rho} &= \frac{1}{2} (\hat{\nabla}_\mu \dot{g}_{\nu\rho} + \hat{\nabla}_\nu \dot{g}_{\mu\rho} - \hat{\nabla}_\rho \dot{g}_{\mu\nu}) + \hat{\Gamma}_{\mu\nu}{}^\sigma \dot{g}_{\rho\sigma}, \\ \partial_0 \hat{\Gamma}_{\mu\nu}{}^\sigma &= \frac{1}{2} \hat{g}^{\sigma\rho} (\hat{\nabla}_\mu \dot{g}_{\nu\rho} + \hat{\nabla}_\nu \dot{g}_{\mu\rho} - \hat{\nabla}_\rho \dot{g}_{\mu\nu}), \\ \partial_0 \hat{\Gamma}_\mu &= \frac{1}{2} \hat{g}^{\nu\rho} \hat{\nabla}_\mu \dot{g}_{\nu\rho}. \end{aligned} \tag{24.48}$$

В этих выражениях производные по времени $\dot{g}_{\mu\nu}$ также исключаются с помощью соотношения (24.44).

Теперь вычислим линейно независимые компоненты тензора кривизны:

$$\begin{aligned}
R_{0\mu 0\nu} &= -N\dot{K}_{\mu\nu} + \hat{R}_{\mu\rho\nu\sigma}N^\rho N^\sigma + NN^\rho(\hat{\nabla}_\mu K_{\nu\rho} + \hat{\nabla}_\nu K_{\mu\rho} - \hat{\nabla}_\rho K_{\mu\nu}) + \\
&\quad + N\hat{\nabla}_\mu \hat{\nabla}_\nu N + K_{\mu\nu}N^\rho N^\sigma K_{\rho\sigma} + N(K_\mu{}^\rho \hat{\nabla}_\nu N_\rho + K_\nu{}^\rho \hat{\nabla}_\mu N_\rho) - \\
&\quad - N^2 K_\mu{}^\rho K_{\nu\rho} - N^\rho N^\sigma K_{\mu\rho} K_{\nu\sigma}, \\
R_{\mu\nu\rho 0} &= \hat{R}_{\mu\nu\rho\sigma}N^\sigma + N(\hat{\nabla}_\mu K_{\nu\rho} - \hat{\nabla}_\nu K_{\mu\rho}) + (K_{\mu\rho}K_{\nu\sigma} - K_{\nu\rho}K_{\mu\sigma})N^\sigma, \\
R_{\mu\nu\rho\sigma} &= \hat{R}_{\mu\nu\rho\sigma} + K_{\mu\rho}K_{\nu\sigma} - K_{\mu\sigma}K_{\nu\rho},
\end{aligned} \tag{24.49}$$

где мы воспользовались формулой для коммутатора ковариантных производных:

$$(\hat{\nabla}_\mu \hat{\nabla}_\nu - \hat{\nabla}_\nu \hat{\nabla}_\mu)N_\rho = -\hat{R}_{\mu\nu\rho\sigma}N^\sigma.$$

Компоненты тензора кривизны, имеющие по крайней мере один временной индекс, относительно ортонормального базиса n, e_μ выглядят проще:

$$\begin{aligned}
R_{\perp\mu\perp\nu} &= \frac{1}{N} \left(-\dot{K}_{\mu\nu} + \hat{\nabla}_\mu \hat{\nabla}_\nu N + K_{\mu\rho} \hat{\nabla}_\nu N^\rho + K_{\nu\rho} \hat{\nabla}_\mu N^\rho - NK_{\mu\rho} K_{\nu}{}^\rho + N^\rho \hat{\nabla}_\rho K_{\mu\nu} \right), \\
R_{\mu\nu\rho\perp} &= \hat{\nabla}_\mu K_{\nu\rho} - \hat{\nabla}_\nu K_{\mu\rho}.
\end{aligned} \tag{24.50}$$

Фактически, компоненты тензора кривизны $R_{\mu\nu\rho\sigma}$ и $R_{\mu\nu\rho\perp}$ уже были получены в предыдущем разделе, см. формулы (24.37), (24.38), без прямых вычислений.

Тензор Риччи имеет следующие линейно независимые компоненты:

$$\begin{aligned}
R_{00} &= -N\hat{g}^{\mu\nu}\dot{K}_{\mu\nu} + \hat{R}_{\mu\nu}N^\mu N^\nu + NN^\mu(2\hat{\nabla}_\nu K^\nu{}_\mu - \partial_\mu K) + N\hat{\nabla}^\mu \hat{\nabla}_\mu N + \\
&\quad + N^\mu N^\nu K_{\mu\nu}K + 2NK^{\mu\nu}\hat{\nabla}_\mu N_\nu - N^2 K^{\mu\nu}K_{\mu\nu} - 2N^\mu N^\nu K_\mu{}^\rho K_{\nu\rho} + \\
&\quad + \frac{N^\mu N^\nu}{N} \left(-\dot{K}_{\mu\nu} + N^\rho \hat{\nabla}_\mu K_{\nu\rho} + \hat{\nabla}_\mu \hat{\nabla}_\nu N + 2K_\mu{}^\rho \hat{\nabla}_\nu N_\rho \right), \\
R_{0\mu} &= \frac{N^\nu}{N} \left(-\dot{K}_{\mu\nu} + N^\sigma \hat{\nabla}_\nu K_{\mu\sigma} + \hat{\nabla}_\mu \hat{\nabla}_\nu N + K_\mu{}^\rho \hat{\nabla}_\nu N_\rho + K_\nu{}^\rho \hat{\nabla}_\mu N_\rho \right) \\
&\quad + \hat{R}_{\mu\nu}N^\nu + N \left(\hat{\nabla}_\nu K^\nu{}_\mu - \partial_\mu K \right) + K_{\mu\nu}N^\nu K - 2K_\mu{}^\rho K_{\nu\rho}N^\nu, \\
R_{\mu\nu} &= \hat{R}_{\mu\nu} + \frac{1}{N} \left(-\dot{K}_{\mu\nu} + \hat{\nabla}_\mu \hat{\nabla}_\nu N + K_\mu{}^\rho \hat{\nabla}_\nu N_\rho + K_\nu{}^\rho \hat{\nabla}_\mu N_\rho + N^\rho \hat{\nabla}_\rho K_{\mu\nu} \right) + \\
&\quad + K_{\mu\nu}K - 2K_\mu{}^\rho K_{\nu\rho}.
\end{aligned} \tag{24.51}$$

Опять же, компоненты тензора Риччи относительно неголономного базиса n, e_μ существенно проще:

$$\begin{aligned}
R_{\perp\perp} &= -\frac{1}{N}\hat{g}^{\mu\nu}\dot{K}_{\mu\nu} + \frac{1}{N}\hat{\nabla}^\mu \hat{\nabla}_\mu N + \frac{2}{N}K^{\mu\nu}\hat{\nabla}_\mu N_\nu - K^{\mu\nu}K_{\mu\nu} + \frac{N^\mu}{N}\partial_\mu K, \\
R_{\perp\mu} &= \hat{\nabla}_\nu K^\nu{}_\mu - \partial_\mu K.
\end{aligned} \tag{24.52}$$

При этом пространственные компоненты тензора Риччи $R_{\mu\nu}$ относительно базиса n, ∂_μ остаются прежними, что следует из последнего равенства (24.42).

Далее, вычисляем скалярную кривизну

$$R = \hat{R} + \frac{2}{N} \left(-\hat{g}^{\mu\nu}\dot{K}_{\mu\nu} + \hat{\nabla}^\mu \hat{\nabla}_\mu N + 2K^{\mu\nu}\hat{\nabla}_\mu N_\nu + N^\mu \partial_\mu K \right) - 3K^{\mu\nu}K_{\mu\nu} + K^2. \tag{24.53}$$

Вакуумные уравнения Эйнштейна

$$G_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R = 0$$

также можно переписать в базисе n, e^μ . Выражения для $G_{\perp\perp}$ и $G_{0\perp}$ уже были получены в предыдущем разделе менее трудоемким образом (24.39). Выражение для $G_{\mu\nu}$ в АДМ параметризации метрики следует из выражения для тензора Риччи (24.51) и скалярной кривизны (24.53). Оно громоздко и мы не будем его выписывать, т.к. в дальнейшем оно нам не понадобится.

Если функции хода равны нулю, то справедливы равенства

$$R_{00} = N^2 R_{\perp\perp}, \quad R_{0\mu} = N R_{\perp\mu},$$

поскольку в этом случае метрика имеет блочно диагональный вид, и времениподобные векторы ∂_0 перпендикулярны пространственным гиперповерхностям.

Как уже отмечалось, тензор Эйнштейна удовлетворяет свернутым тождествам Бианки (6.121). Эти тождества полезно переписать в АДМ параметризации метрики. Четыре тождества (6.121) перепишем в следующем виде:

$$\nabla_\beta G^{\beta\alpha} = 0 \quad \Leftrightarrow \quad [(\nabla_\beta G^{\beta\alpha})n_\alpha = 0, \quad (\nabla_\beta G^{\beta\alpha})e_\alpha^\mu = 0].$$

После довольно утомительных вычислений получим равенства:

$$(\nabla_\beta G^{\beta\alpha})n_\alpha = \frac{1}{N}\dot{G}^{\perp\perp} - \frac{N^\mu}{N}\partial_\mu G^{\perp\perp} - G^{\perp\perp}K + \hat{\nabla}_\mu G^{\mu\perp} + 2\frac{\partial_\mu N}{N}G^{\mu\perp} + G^{\mu\nu}K_{\mu\nu}, \quad (24.54)$$

$$\begin{aligned} (\nabla_\beta G^{\beta\alpha})e_\alpha^\mu = & \frac{1}{N}\dot{G}^{\perp\mu} - \frac{N^\nu}{N}\hat{\nabla}_\nu G^{\perp\mu} - G^{\perp\mu}K + \hat{\nabla}_\nu G^{\nu\mu} + \frac{\partial_\nu N}{N}G^{\nu\mu} - \\ & - \frac{1}{N}G^{\perp\perp}\hat{g}^{\mu\nu}\partial_\nu N + \frac{1}{N}G^{\perp\nu}\hat{\nabla}_\nu N^\mu - 2G^{\perp\nu}K_\nu^\mu. \end{aligned} \quad (24.55)$$

Эти тождества будут полезны при анализе вторичных связей в общей теории относительности.

24.5 Гамильтониан

Скалярная кривизна содержит вторые производные от компонент метрики как по времени, так и по пространственным координатам и поэтому неудобна для канонической формулировки общей теории относительности. Чтобы исключить из лагранжиана все вторые производные, в разделе 24.1 к действию был добавлен граничный член. Однако для канонической формулировки достаточно исключить из лагранжиана только вторые производные по времени. Наиболее простой вид лагранжиан принимает после добавления следующего граничного члена

$$L_{\text{ADM}} = N\hat{e}R + 2\partial_0(\hat{e}K) - 2\partial_\mu(\hat{e}\hat{g}^{\mu\nu}\partial_\nu N), \quad (24.56)$$

где

$$\hat{e} := \det \hat{e}_\mu^i = \sqrt{|\hat{g}|} = \sqrt{|\det g_{\mu\nu}|}.$$

Прямые вычисления приводят к следующему простому выражению

$$L_{\text{ADM}} = N\hat{e} \left(K^{\mu\nu}K_{\mu\nu} - K^2 + \hat{R} \right). \quad (24.57)$$

Теперь нетрудно перейти к гамильтонову формализму. Во-первых, АДМ лагранжиан не содержит производных по времени от функции хода N и функций сдвига N_μ . Это значит, что теория содержит n первичных связей:

$$p^0 := \frac{\partial L_{\text{ADM}}}{\partial \dot{N}} = 0, \quad p^\mu := \frac{\partial L_{\text{ADM}}}{\partial \dot{N}_\mu} = 0, \quad (24.58)$$

число которых совпадает с числом независимых функций, параметризующих диффеоморфизмы. Импульсы, канонически сопряженные к пространственной метрике $g_{\mu\nu}$, имеют вид

$$p^{\mu\nu} = \frac{\partial L_{\text{ADM}}}{\partial \dot{g}_{\mu\nu}} = -\frac{1}{2N} \frac{\partial L_{\text{ADM}}}{\partial K_{\mu\nu}} = -\hat{e} (K^{\mu\nu} - g^{\mu\nu} K), \quad (24.59)$$

где мы воспользовались выражением для внешней кривизны (24.44). Отсюда следует, что обобщенные импульсы пропорциональны тензору внешней кривизны. Отметим, что импульсы являются не тензорами относительно преобразований координат x^μ , а тензорными плотностями степени -1 , как и определитель репера, степень которого равна

$$\deg \hat{e} = -1.$$

Поэтому ковариантная производная от импульсов в соответствии с правилом ковариантного дифференцирования (6.27) имеет вид

$$\hat{\nabla}_\mu p^{\nu\rho} = \partial_\mu p^{\nu\rho} + \hat{\Gamma}_{\mu\sigma}^\nu p^{\sigma\rho} + \hat{\Gamma}_{\mu\sigma}^\rho p^{\nu\sigma} - \hat{\Gamma}_\mu^{\nu\rho} p^{\nu\rho}.$$

Чтобы исключить из АДМ-лагранжиана скорости $\dot{g}_{\mu\nu}$, разложим импульсы на неприводимые компоненты, выделив из $p^{\mu\nu}$ след,

$$p^{\mu\nu} = \tilde{p}^{\mu\nu} + \frac{1}{n-1} p \hat{g}^{\mu\nu}, \quad (24.60)$$

где мы ввели след импульсов

$$p := p^{\mu\nu} g_{\mu\nu} = \hat{e} (n-2) K \quad (24.61)$$

и симметричную бесследовую часть

$$\tilde{p}^{\mu\nu} := \tilde{p}^{\nu\mu} = -\hat{e} \left(K^{\mu\nu} - \frac{1}{n-1} \hat{g}^{\mu\nu} K \right), \quad \tilde{p}^{\mu\nu} g_{\mu\nu} = 0.$$

Импульсы $p^{\mu\nu}$ и внешняя кривизна $K^{\mu\nu}$ находятся во взаимно однозначном соответствии. Из формулы (24.59) следует выражение внешней кривизны через импульсы

$$K^{\mu\nu} = -\frac{1}{\hat{e}} \left(p^{\mu\nu} - \frac{1}{n-2} g^{\mu\nu} p \right). \quad (24.62)$$

Теперь можно решить уравнение (24.59) относительно скоростей, воспользовавшись соотношением (24.44),

$$\dot{g}_{\mu\nu} = \frac{2N}{\hat{e}} \left(p_{\mu\nu} - \frac{1}{(n-2)} p g_{\mu\nu} \right) + \hat{\nabla}_\mu N_\nu + \hat{\nabla}_\nu N_\mu.$$

Несложные вычисления приводят к гамильтоновой плотности

$$H = p^{\mu\nu} \dot{g}_{\mu\nu} - L_{\text{ADM}} = N H_\perp + N^\mu H_\mu + 2\partial_\mu (p^{\mu\nu} N_\nu), \quad (24.63)$$

где

$$\begin{aligned} H_{\perp} &:= \frac{1}{\hat{e}} \left(p^{\mu\nu} p_{\mu\nu} - \frac{1}{(n-2)} p^2 \right) - \hat{e} \hat{R}, \\ H_{\mu} &:= -2\hat{\nabla}_{\nu} p^{\nu}_{\mu} = -2\partial_{\nu} p^{\nu}_{\mu} + \partial_{\mu} g_{\nu\rho} p^{\nu\rho}, \end{aligned} \quad (24.64)$$

и $p_{\mu\nu} := g_{\mu\rho} g_{\nu\sigma} p^{\rho\sigma}$.

Замечание. Ковариантная производная от импульсов в последнем выражении содержит только одно слагаемое с символами Кристоффеля, т.к. импульсы являются тензорными плотностями. $\square$

Отбрасывая в выражении для гамильтониана (24.63) дивергенцию, приходим к окончательному выражению для гамильтониана

$$\mathcal{H}_{\text{ADM}} = \int d\mathbf{x} H_{\text{ADM}} = \int d\mathbf{x} (N H_{\perp} + N^{\mu} H_{\mu}). \quad (24.65)$$

Для функционалов, полученных интегрированием по пространственным сечениям $x^0 = \text{const}$, будем использовать каллиграфический шрифт.

Перепишем выражение для $H_{\perp}$ в терминах неприводимых компонент импульсов (24.60):

$$H_{\perp} = \frac{1}{\hat{e}} \left[\tilde{p}^{\mu\nu} \tilde{p}_{\mu\nu} - \frac{1}{(n-1)(n-2)} p^2 \right] - \hat{e} \hat{R}.$$

Отсюда следует, что квадратичная форма импульсов в $H_{\perp}$ при $n \geq 3$ не является положительно определенной. Отсутствие положительной определенности квадратичной формы импульсов в гамильтониане представляет серьезные трудности для физической интерпретации моделей. В рассматриваемом случае не все компоненты импульсов являются физическими, т.к. общая теория относительности инвариантна относительно общих преобразований координат. После исключения нефизических степеней свободы с помощью связей и калибровочных условий квадратичная форма импульсов для физических степеней свободы будет положительно определена.

24.6 Вторичные связи

Для завершения построения гамильтонова формализма необходимо исследовать согласованность первичных связей (24.58) с уравнениями движения. Фазовое пространство общей теории относительности бесконечномерно и описывается $n(n+1)$ сопряженными координатами и импульсами: (N, p^0) , (N_{μ}, p^{μ}) , $(g_{\mu\nu}, p^{\mu\nu})$, в каждой точке пространства $\mathbf{x} \in \mathbb{U}$. На этом фазовом пространстве задана каноническая пуассонова структура с помощью одновременных скобок Пуассона:

$$[N, p'^0] = \delta(\mathbf{x} - \mathbf{x}'), \quad [N_{\mu}, p'^{\nu}] = \delta_{\mu}^{\nu} \delta(\mathbf{x} - \mathbf{x}'), \quad [g_{\mu\nu}, p'^{\rho\sigma}] = \delta_{\mu\nu}^{\rho\sigma} \delta(\mathbf{x} - \mathbf{x}'), \quad (24.66)$$

где штрих у полевой переменной означает, что она рассматривается в точке $\mathbf{x}' = (x'^1, \dots, x'^{n-1})$. Все поля рассматриваются в момент времени $t = x^0$. В правых частях скобок Пуассона для краткости использованы следующие обозначения для пространственной $(n-1)$ -мерной δ -функции и симметризованной комбинации символов Кронекера:

$$\begin{aligned} \delta(\mathbf{x} - \mathbf{x}') &:= \delta(x^1 - x'^1) \dots \delta(x^{n-1} - x'^{n-1}), \\ \delta_{\mu\nu}^{\rho\sigma} &:= \frac{1}{2} (\delta_{\mu}^{\rho} \delta_{\nu}^{\sigma} + \delta_{\mu}^{\sigma} \delta_{\nu}^{\rho}). \end{aligned} \quad (24.67)$$

Если заданы два функционала

$$\mathcal{F} = \int d\mathbf{x} F, \quad \mathcal{G} = \int d\mathbf{x} G,$$

где F и G – некоторые функции от пространственной метрики $g_{\mu\nu}$, сопряженных импульсов $p^{\mu\nu}$ и их производных, то скобка Пуассона этих функционалов определяется следующим интегралом

$$[\mathcal{F}, \mathcal{G}] := \int d\mathbf{x} \left(\frac{\delta \mathcal{F}}{\delta g_{\mu\nu}} \frac{\delta \mathcal{G}}{\delta p^{\mu\nu}} - \frac{\delta \mathcal{F}}{\delta p^{\mu\nu}} \frac{\delta \mathcal{G}}{\delta g_{\mu\nu}} \right). \quad (24.68)$$

Напомним, что сами компоненты метрики и импульсов также можно рассматривать в виде функционалов. Например,

$$g_{\mu\nu}(\mathbf{x}) = \int d\mathbf{x}' g_{\mu\nu}(\mathbf{x}') \delta(\mathbf{x} - \mathbf{x}').$$

Тогда последняя скобка Пуассона в (24.66) является следствием общего определения канонической пуассоновой структуры на множестве функционалов (24.68).

Аналогично определяется скобка Пуассона для функционалов, зависящих от полного набора канонических переменных: (N, p^0) , (N_μ, p^μ) , $(g_{\mu\nu}, p^{\mu\nu})$.

Рассмотрим теперь гамильтоновы уравнения движения для первичных связей (24.58):

$$\begin{aligned} \dot{p}^0 &= [p^0, \mathcal{H}_{\text{ADM}}] = -\frac{\delta \mathcal{H}_{\text{ADM}}}{\delta N} = -H_\perp, \\ \dot{p}^\mu &= [p^\mu, \mathcal{H}_{\text{ADM}}] = -\frac{\delta \mathcal{H}_{\text{ADM}}}{\delta N^\mu} = -H^\mu. \end{aligned}$$

Из условия согласованности первичных связей с уравнениями движения $\dot{p}^0 = 0$, $\dot{p}^\mu = 0$ следуют вторичные связи:

$$H_\perp = 0, \quad H_\mu = 0. \quad (24.69)$$

Отметим, что вторичные связи являются не тензорами, а тензорными плотностями степени -1 . Кроме того, вместо связей H^μ удобнее рассматривать эквивалентную систему связей с опущенным индексом $H_\mu := g_{\mu\nu} H^\nu$. В следующем разделе будет показано, что эти связи определяют генераторы преобразований координат на сечениях $x^0 = \text{const}$ и удовлетворяют более простой алгебре.

Связи H_μ линейны по импульсам и метрике. Связь $H_\perp$ квадратична по импульсам и неполиномиальна по метрике $g_{\mu\nu}$, поскольку зависит от корня из определителя метрики $\hat{e}$ и обратной метрики $\hat{g}^{\mu\nu}$. Последнее обстоятельство является существенной технической трудностью при построении теории возмущений.

Вторичные связи не зависят от канонических переменных (N, p^0) , (N_μ, p^μ) , и их можно исключить, рассматривая $[n(n-1) \times \infty^{n-1}]$ -мерное фазовое пространство переменных $g_{\mu\nu}$ и $p^{\mu\nu}$, на которые наложены связи (24.69). В этом случае функция хода N и функции сдвига N_μ рассматриваются, как множители Лагранжа в задаче на условный экстремум для действия

$$S = \int dx (p^{\mu\nu} \dot{g}_{\mu\nu} - H_{\text{ADM}}). \quad (24.70)$$

Поскольку гамильтониан общей теории относительности (24.65) представляет собой линейную комбинацию связей, то для исследования согласованности вторичных связей (24.69) с уравнениями движения необходимо вычислить скобки Пуассона связей между собой. Алгебра связей в общей теории относительности хорошо известна:

$$[H_{\perp}, H'_{\perp}] = -(H_{\mu}\hat{g}^{\mu\nu} + H'_{\mu}\hat{g}'^{\mu\nu})\delta_{\nu}, \quad (24.71)$$

$$[H_{\perp}, H'_{\mu}] = -H'_{\perp}\delta_{\mu}, \quad (24.72)$$

$$[H_{\mu}, H'_{\nu}] = -H_{\nu}\delta_{\mu} - H'_{\mu}\delta_{\nu}, \quad (24.73)$$

где введено сокращенное обозначение для производной δ -функции:

$$\delta_{\mu} := \frac{\partial}{\partial x'^{\mu}}\delta(\mathbf{x}' - \mathbf{x}).$$

Вид двух скобок Пуассона (24.72) и (24.73) можно найти, не прибегая к прямым вычислениям. С этой целью рассмотрим функционал

$$T_u := - \int d\mathbf{x} \, u^{\mu} H_{\mu},$$

где u^{μ} малое дифференцируемое векторное поле, которое мы рассматриваем как инфинитезимальный параметр преобразования. Вычисление скобок Пуассона координат фазового пространства $g_{\mu\nu}$, $p^{\mu\nu}$ с T_u приводит к равенствам:

$$\begin{aligned} \delta_u g_{\mu\nu} &:= [g_{\mu\nu}, T_u] = -\partial_{\mu} u^{\rho} g_{\rho\nu} - \partial_{\nu} u^{\rho} g_{\mu\rho} - u^{\rho} \partial_{\rho} g_{\mu\nu}, \\ \delta_u p^{\mu\nu} &:= [p^{\mu\nu}, T_u] = \partial_{\rho} u^{\mu} p^{\rho\nu} + \partial_{\rho} u^{\nu} p^{\mu\rho} - \partial_{\rho} (u^{\rho} p^{\mu\nu}). \end{aligned}$$

Это значит, что функционал T_u , который определяется связями H_{μ} , является генератором общих преобразований координат на гиперповерхностях $x^0 = \text{const}$. Напомним, что импульсы $p^{\mu\nu}$ являются тензорными плотностями степени -1 . Алгебра преобразований координат хорошо известна и задается скобкой Пуассона (24.73). Скобку Пуассона (24.72) также можно не вычислять явно. Ее вид следует из того, что связь $H_{\perp}$ является скалярной плотностью степени -1 . Таким образом, необходимо вычислить только скобку Пуассона (24.71). Эти вычисления очень громоздки, и впервые, по-видимому, были проделаны намного позже ДеВиттом [175]. В следующем разделе мы вычислим эту скобку после канонического преобразования, приводящего связи к полиномиальному виду, что существенно упрощает вычисления.

Ввиду того, что связи H_{μ} задают только пространственные диффеоморфизмы, мы будем называть их кинематическими. Они также не зависят от констант связи в действии, если таковые имеются. Связь $H_{\perp}$ называется динамической, т.к. она определяет развитие начальных данных во времени и существенно зависит от исходного действия, в частности, от констант связи.

Для сравнения приведем скобку Пуассона связей $H^{\mu} = 0$ с контравариантным индексом, которые эквивалентны связям $H_{\mu} = 0$,

$$[H^{\mu}, H'^{\nu}] = (\hat{g}^{\mu\nu} H^{\rho} + \hat{g}'^{\mu\nu} H'^{\rho}) \delta_{\rho} + (\hat{g}^{\mu\rho} \partial_{\rho} g_{\sigma\lambda} \hat{g}^{\nu\sigma} - \hat{g}'^{\nu\rho} \partial_{\rho} g_{\sigma\lambda} \hat{g}'^{\mu\sigma}) H^{\lambda} \delta.$$

Как видим, данная скобка Пуассона выглядит сложнее (24.73). Это говорит о том, что выбор вида связей из эквивалентных наборов является очень важным и может привести к существенному упрощению вычислений. К сожалению, общего рецепта как поступать в таких случаях не существует, и действует метод “пристального всматривания”.

Теперь посмотрим на вторичные связи с другой точки зрения. Сравнение выражения для связей (24.64) с тензором Эйнштейна (24.39) приводит к равенствам

$$\begin{aligned} H_{\perp} &= 2\hat{e}G_{\perp\perp} = \frac{2\hat{e}}{N^2}(G_{00} - 2N^{\mu}G_{0\mu} + N^{\mu}N^{\nu}G_{\mu\nu}), \\ H_{\mu} &= 2\hat{e}G_{\perp\mu} = \frac{2\hat{e}}{N}(G_{0\mu} - N^{\nu}G_{\nu\mu}). \end{aligned}$$

Мы видим, что вторичные связи (24.69) эквивалентны четырем уравнениям Эйнштейна

$$G_{\perp\perp} = 0, \quad G_{\perp\mu} = 0,$$

которые представляют собой определенные комбинации всех вакуумных уравнений Эйнштейна $G_{\alpha\beta} = 0$. Если выполнены все уравнения Эйнштейна, а также вторичные связи в начальный момент времени, то вторичные связи будут выполнены также во все последующие моменты времени. Это следует из того, что $\dot{G}_{\perp\perp} = 0$ и $\dot{G}_{\perp\mu} = 0$ как следствие свернутых тождеств Бианки (24.54), (24.55). Таким образом мы доказали, что вторичные связи в общей теории относительности согласованы с уравнениями движения, т.е. представляют собой связи первого рода. Для этого необязательно вычислять в явном виде алгебру связей (24.71)–(24.73). Однако знание алгебры связей полезно.

В заключение данного раздела выпишем канонические уравнения движения для пространственных компонент метрики и импульсов

$$\dot{g}_{\mu\nu} = \frac{2N}{\hat{e}}p_{\mu\nu} - \frac{2N}{\hat{e}(n-2)}g_{\mu\nu}p + \hat{\nabla}_{\mu}N_{\nu} + \hat{\nabla}_{\nu}N_{\mu}, \quad (24.74)$$

$$\begin{aligned} \dot{p}^{\mu\nu} &= \frac{N}{2\hat{e}}\hat{g}^{\mu\nu} \left(p^{\rho\sigma}p_{\rho\sigma} - \frac{1}{n-2}p^2 \right) - \frac{2N}{\hat{e}} \left(p^{\mu\rho}p^{\nu}_{\rho} - \frac{1}{n-2}p^{\mu\nu}p \right) - \\ &\quad - \hat{e}N \left(\hat{R}^{\mu\nu} - \frac{1}{2}\hat{g}^{\mu\nu}\hat{R} \right) + \hat{e}(\Delta N\hat{g}^{\mu\nu} - \hat{\nabla}^{\mu}\hat{\nabla}^{\nu}N) - \\ &\quad - p^{\mu\rho}\hat{\nabla}_{\rho}N^{\nu} - p^{\nu\rho}\hat{\nabla}_{\rho}N^{\mu} + \hat{\nabla}_{\rho}(N^{\rho}p^{\mu\nu}), \end{aligned} \quad (24.75)$$

где $\Delta := \hat{\nabla}^{\mu}\hat{\nabla}_{\mu}$ – оператор Лапласа–Бельтрами.

Уравнение (24.74), конечно, совпадает с определением импульсов (24.59). Уравнения для импульсов (24.75) эквивалентны уравнениям Эйнштейна $G_{\mu\nu} = 0$. Чтобы доказать это, запишем уравнения (24.75) в виде

$$\dot{p}^{\mu\nu} - (\dots)^{\mu\nu} = 0,$$

где точки обозначают все слагаемые в правой части (24.75). Тогда можно проверить, что выполнено равенство

$$g_{\mu\rho}g_{\nu\sigma}[\dot{p}^{\rho\sigma} - (\dots)^{\rho\sigma}] = N\hat{e}G_{\mu\nu}.$$

Таким образом, канонические уравнения движения (24.74), (24.75) с учетом уравнений связей (24.69) эквивалентны полной системе вакуумных уравнений Эйнштейна $G_{\alpha\beta} = 0$.

24.7 Полиномиальная гамильтонова форма

Гамильтониан для лагранжиана Гильберта–Эйнштейна (24.65), построенный ранее, полиномиален по импульсам $p^{\mu\nu}$, но неполиномиален по компонентам пространственной части метрики $g_{\mu\nu}$, т.к. содержит скалярную кривизну $\hat{R}$ и определитель репера,

которые входят в динамическую связь $H_\perp$. В настоящем разделе мы опишем каноническое преобразование, которое приводит к полиномиальной форме гамильтониана и, следовательно, к полиномиальной форме гамильтоновых уравнений движения. Это преобразование является аналогом полиномиальной лагранжевой формы действия Гильберта–Эйнштейна, рассмотренной в разделе 23.8.

Идея канонического преобразования состоит в следующем. Импульсы $p^{\mu\nu}$ приводимы и разлагаются на бесследовую часть $\tilde{p}^{\mu\nu}$ и след p в соответствии с формулой (24.60). При вычислениях, как правило, удобнее работать с неприводимыми компонентами, поскольку много слагаемых автоматически сокращаются. Поставим вопрос: “Нельзя ли совершить такое каноническое преобразование, после которого новыми импульсами будут неприводимые компоненты $\tilde{p}^{\mu\nu}$ и p ?”. Этот вопрос нетривиален, потому что разложение импульсов включает метрику, компоненты которой сами являются координатами фазового пространства. Ответ на поставленный вопрос отрицательный, потому что скобка Пуассона импульсов между собой отлична от нуля. Например, $[\tilde{p}^{\mu\nu}, p'] \neq 0$, где $p' := p(t, \mathbf{x}')$. Однако существует такое каноническое преобразование, что новые импульсы будут пропорциональны неприводимым компонентам $\tilde{p}^{\mu\nu}$ и p . Построением этого канонического преобразования мы и займемся в настоящем разделе.

Рассмотрим каноническое преобразование

$$(g_{\mu\nu}, p^{\rho\sigma}) \mapsto (k_{\mu\nu}, P^{\rho\sigma}), (\varrho, P), \quad (24.76)$$

к новым парам канонически сопряженных координат $k_{\mu\nu}$, ϱ и импульсов $P^{\mu\nu}$, P , где на координаты $k_{\mu\nu}$ и сопряженные им импульсы $P^{\mu\nu}$ наложены дополнительные условия

$$|\det k_{\mu\nu}| = 1, \quad P^{\mu\nu} k_{\mu\nu} = 0. \quad (24.77)$$

В качестве производящего функционала канонического преобразования выберем функционал

$$\mathcal{F} = - \int d\mathbf{x} \varrho^s k_{\mu\nu} p^{\mu\nu}, \quad s \in \mathbb{R}, \quad s \neq 0, \quad (24.78)$$

зависящий от новых координат ϱ , $k_{\mu\nu}$ и старых импульсов $p^{\mu\nu}$, а также от вещественного параметра s , который будет определен позже. Тогда старые координаты и новые импульсы определяются вариационными производными (см. раздел 21.12)

$$g_{\mu\nu} := - \frac{\delta \mathcal{F}}{\delta p^{\mu\nu}} = \varrho^s k_{\mu\nu}, \quad (24.79)$$

$$P^{\mu\nu} := - \frac{\delta \mathcal{F}}{\delta k_{\mu\nu}} = \varrho^s \tilde{p}^{\mu\nu}, \quad (24.80)$$

$$P := - \frac{\delta \mathcal{F}}{\delta \varrho} = \frac{s}{\varrho} p. \quad (24.81)$$

При вычислении вариационной производной по $k_{\mu\nu}$ учтено условие $|\det k_{\mu\nu}| = 1$, из которого вытекает ограничение на вариации $k^{\mu\nu} \delta k_{\mu\nu} = 0$, где $k^{\mu\nu}$ – тензорная плотность, обратная к $k_{\mu\nu}$: $k^{\mu\nu} k_{\nu\sigma} = \delta^\mu_\sigma$. Тем самым равенство нулю следа импульсов (24.77) автоматически следует из условия единичности определителя плотности $k_{\mu\nu}$ для производящего функционала (24.78). В выражении (24.81) учтено соотношение (24.79).

По сути дела, в качестве новой канонической переменной из метрики выделен ее определитель в некоторой степени, как следует из (24.79):

$$|\det g_{\mu\nu}| = \hat{e}^2 = \varrho^{s(n-1)}.$$

Поскольку метрика $g_{\mu\nu}$ невырождена, то отсюда, в частности, следует равенство $\varrho > 0$.

В дальнейшем симметричную тензорную плотность с единичным определителем $k_{\mu\nu}$ мы для краткости также будем называть метрикой.

Из формул (24.80), (24.81) следует, что новые импульсы $P^{\mu\nu}$ и P пропорциональны неприводимым компонентам старых импульсов $\tilde{p}^{\mu\nu}$ и p .

Отметим, что все новые канонические переменные являются тензорными плотностями следующих степеней:

$$\begin{aligned} \deg k_{\mu\nu} &= \frac{2}{n-1}, & \deg \varrho &= -\frac{2}{s(n-1)}, \\ \deg P^{\mu\nu} &= -\frac{2}{n-1} - 1, & \deg P &= \frac{2}{s(n-1)} - 1. \end{aligned}$$

Чтобы вычислить скалярную кривизну $\hat{R}$, входящую в связь (24.64), заметим, что каноническое преобразование (24.79) имеет вид преобразования Вейля $g_{\mu\nu} = e^{2\phi} k_{\mu\nu}$, где $\phi = s \ln \varrho / 2$. Поэтому можно воспользоваться полученной ранее формулой (17.40). Прямые вычисления приводят к следующему выражению для скалярной кривизны сечения $x^0 = \text{const}$ в новых координатах:

$$\hat{R} = \varrho^{-s-2} \left[\varrho^2 R^{(k)} + s(n-2) \varrho \partial_\mu (k^{\mu\nu} \partial_\nu \varrho) + s(n-2) \left(s \frac{n-3}{4} - 1 \right) (\partial \varrho)^2 \right], \quad (24.82)$$

где $(\partial \varrho)^2 := k^{\mu\nu} \partial_\mu \varrho \partial_\nu \varrho$. Скалярная кривизна, построенная по метрике $k_{\mu\nu}$, принимает удивительно простой вид

$$R^{(k)} = \partial_{\mu\nu}^2 k^{\mu\nu} + \frac{1}{2} k^{\mu\nu} \partial_\rho k_{\mu\sigma} \partial_\nu k^{\rho\sigma} - \frac{1}{4} k^{\mu\nu} \partial_\mu k_{\rho\sigma} \partial_\nu k^{\rho\sigma}. \quad (24.83)$$

Из единичности определителя метрики следует, что компоненты обратной метрики $k^{\mu\nu}$ являются полиномами степени $n-2$ по компонентам $k_{\mu\nu}$:

$$k^{\mu\nu} = \frac{1}{(n-2)!} \hat{\varepsilon}^{\mu\rho_1 \dots \rho_{n-2}} \hat{\varepsilon}^{\nu\sigma_1 \dots \sigma_{n-2}} k_{\rho_1\sigma_1} \dots k_{\rho_{n-2}\sigma_{n-2}},$$

где $\hat{\varepsilon}^{\mu_1 \dots \mu_{n-1}}$ — полностью антисимметричная тензорная плотность ранга $n-1$. Поэтому скалярная кривизна $R^{(k)}$ полиномиальна как по метрике $k_{\mu\nu}$, так и по обратной метрике $k^{\mu\nu}$.

Динамическая связь в новых переменных принимает вид

$$\begin{aligned} H_\perp &= \varrho^{-\frac{s(n-1)}{2}} \left[P^{\mu\nu} P_{\mu\nu} - \frac{\varrho^2}{s^2(n-1)(n-2)} P^2 \right] - \\ &- \varrho^{\frac{s(n-1)}{2} - s - 2} \left[\varrho^2 R^{(k)} + s(n-2) \varrho \partial_\mu (k^{\mu\nu} \partial_\nu \varrho) + s(n-2) \left(s \frac{n-3}{4} - 1 \right) (\partial \varrho)^2 \right], \end{aligned}$$

где опускание индексов у импульсов производится с помощью новой метрики, $P_{\mu\nu} := k_{\mu\rho} k_{\nu\sigma} P^{\rho\sigma}$.

Проанализируем возможность такого выбора постоянной s , чтобы динамическая связь имела полиномиальный вид. Оба выражения в квадратных скобках полиномиальны по всем динамическим переменным. Поскольку $n \geq 3$, то для положительности степени плотности ϱ перед первой квадратной скобкой необходимо выполнение неравенства $s < 0$. В этом случае степень ϱ перед второй квадратной скобкой

будет отрицательна. Таким образом, за счет выбора параметра s добиться полиномиальности самой связи $H_\perp$ нельзя. Однако, связь можно умножить целиком на произвольный множитель, отличный от нуля. При этом поверхность в фазовом пространстве, определяемая данной связью не изменится. Минимальная степень ϱ , на которую необходимо умножить $H_\perp$ будет тогда, когда степени ϱ перед квадратными скобками будут равны. Отсюда следует равенство

$$s = \frac{2}{n-2}.$$

Тогда, умножив динамическую связь на степень ρ ,

$$K_\perp := \varrho^{\frac{n-1}{n-2}} H_\perp, \quad (24.84)$$

получим эквивалентную полиномиальную связь

$$K_\perp = P^{\mu\nu} P_{\mu\nu} - \frac{n-2}{4(n-1)} \varrho^2 P^2 - \varrho^2 R^{(k)} - 2\varrho \partial_\mu (k^{\mu\nu} \partial_\nu \rho) + \frac{n-1}{n-2} (\partial \varrho)^2 = 0. \quad (24.85)$$

Кинематические связи в новых динамических переменных сохраняют свою полиномиальность:

$$H_\mu = -2\partial_\nu (P^{\nu\sigma} k_{\sigma\mu}) + P^{\nu\sigma} \partial_\mu k_{\nu\sigma} - \frac{n-2}{n-1} \partial_\mu (P \varrho) + P \partial_\mu \varrho = 0. \quad (24.86)$$

Исходя из явного выражения для новых канонических переменных (24.79)–(24.81), вычислим основные скобки Пуассона. Отличными от нуля являются только три скобки:

$$[k_{\mu\nu}, P'^{\varrho\sigma}] = \left(\delta_{\mu\nu}^{\rho\sigma} - \frac{1}{n-1} k_{\mu\nu} k^{\rho\sigma} \right) \delta(\mathbf{x}' - \mathbf{x}), \quad (24.87)$$

$$[P^{\mu\nu}, P'^{\rho\sigma}] = \frac{1}{n-1} (P^{\mu\nu} k^{\rho\sigma} - P^{\rho\sigma} k^{\mu\nu}) \delta(\mathbf{x}' - \mathbf{x}), \quad (24.88)$$

$$[\varrho, P'] = \delta(\mathbf{x}' - \mathbf{x}), \quad (24.89)$$

где

$$\delta_{\mu\nu}^{\rho\sigma} := \frac{1}{2} (\delta_\mu^\rho \delta_\nu^\sigma + \delta_\mu^\sigma \delta_\nu^\rho).$$

Скобка Пуассона (24.87) не имеет канонического вида для фазовых переменных. Появление дополнительного слагаемого в (24.87) связано с тем, что на поля $k_{\mu\nu}$ и $P^{\mu\nu}$ наложены дополнительные условия (24.77). По этой же причине отлична от нуля скобка Пуассона (24.88).

Нетрудно проверить, что пуассонова структура, определяемая скобками Пуассона (24.87)–(24.89) является вырожденной. Это значит, что все многообразие $\mathbb{N}$, задаваемое координатами $(k_{\mu\nu}, P^{\mu\nu})$ и (ϱ, P) является не симплектическим, а только пуассоновым многообразием (см. раздел 19.3). На пуассоновом многообразии $\mathbb{N}$ существуют две функции Казимира:

$$c^1 := \det k_{\mu\nu}, \quad c^2 := P^{\mu\nu} k_{\mu\nu}.$$

Действительно, из определения пуассоновой структуры (24.87)–(24.89) следует, что следующие скобки Пуассона равны нулю:

$$[c^{1,2}, k'_{\mu\nu}] = [c^{1,2}, P'^{\mu\nu}] = [c^{1,2}, \varrho'] = [c^{1,2}, P'] = 0.$$

Отсюда вытекает, что скобка Пуассона $[c^{1,2}, f'] = 0$, где $f \in \mathcal{C}^1(\mathbb{N})$ – произвольная дифференцируемая функция на $\mathbb{N}$. Пуассонова структура, ограниченная на сечения $\mathbb{V} \subset \mathbb{N}$, которые определяются уравнениями $c^1 = \text{const}$ и $c^2 = \text{const}$, невырождена. Следовательно, эти сечения являются симплектическими. Новым фазовым пространством общей теории относительности в рассматриваемом случае является подмногообразие $\mathbb{V} \subset \mathbb{N}$, определяемое двумя фиксированными значениями функций Казимира (24.77). Строго говоря, каноническое преобразование (24.76) является каноническим, т.е. сохраняющим вид скобок Пуассона, только между исходным фазовым пространством и подмногообразием $\mathbb{V}$. Полиномиальность связей достигнута за счет расширения исходного фазового пространства до пуассонова многообразия $\mathbb{N}$. Если решить дополнительные связи (24.77) явно, то полиномиальность будет нарушена. В этом нет ничего необычного. Например, электродинамика содержит связи, явное решение которых приводит даже к нелокальному действию для физических степеней свободы (см., например, [176]).

Рассмотрим алгебру связей. Поскольку вместо динамической связи $H_\perp$ мы ввели новую связь $K_\perp$, то алгебра связей изменится. Несложные вычисления приводят к равенствам

$$[K_\perp, K'_\perp] = -(\varrho^2 H^\mu + \varrho'^2 H'^\mu) \delta_\mu(\mathbf{x}' - \mathbf{x}), \quad (24.90)$$

$$[K_\perp, H'_\mu] = -(K_\perp + K'_\perp) \delta_\mu(\mathbf{x}' - \mathbf{x}), \quad (24.91)$$

$$[H_\mu, H'_\nu] = -H_\nu \delta_\mu(\mathbf{x}' - \mathbf{x}) - H'_\mu \delta_\nu(\mathbf{x}' - \mathbf{x}), \quad (24.92)$$

где

$$H_\mu := k_{\mu\nu} H^\nu, \quad \text{и} \quad \delta_\mu(\mathbf{x}' - \mathbf{x}) := \frac{\partial}{\partial x'^\mu} \delta(\mathbf{x}' - \mathbf{x}).$$

По сравнению с исходной алгеброй (24.71)–(24.72), изменения касаются скобок Пуассона (24.90) и (24.91). Скобка Пуассона (24.90) является результатом прямого счета. Вторая скобка Пуассона (24.91) носит кинематический характер и определяется тем, что новая связь является не функцией, а скалярной плотностью.

Получим явные выражения для геометрических объектов в новых переменных. Сначала вычислим символы Кристоффеля:

$$\hat{\Gamma}_{\mu\nu}{}^\rho = \frac{1}{(n-2)\varrho} (\partial_\mu \varrho \delta_\nu^\rho + \partial_\nu \varrho \delta_\mu^\rho - k_{\mu\nu} k^{\rho\sigma} \partial_\sigma \varrho) + \Gamma_{\mu\nu}^{(k)\rho},$$

где “символы Кристоффеля” $\Gamma_{\mu\nu}^{(k)\rho}$ выражаются через тензорную плотность $k_{\mu\nu}$ по тем же формулам, что и символы Кристоффеля через метрику. Конечно, “символы Кристоффеля” $\Gamma_{\mu\nu}^{(k)\rho}$ никакой связности не определяют. Отсюда следует выражение для следа символов Кристоффеля, который определяет дополнительные слагаемые в ковариантных производных тензорных плотностей,

$$\hat{\Gamma}_\mu = \frac{n-1}{n-2} \frac{\partial_\mu \varrho}{\varrho},$$

т.к.

$$\Gamma_\mu^{(k)} := \Gamma_{\nu\mu}^{(k)\nu} = \frac{1}{2} k^{\nu\rho} \partial_\mu k_{\nu\rho} = 0$$

в силу условия $|\det k_{\mu\nu}| = 1$. Нетрудно проверить ковариантное постоянство новых переменных:

$$\hat{\nabla}_\mu k_{\nu\rho} = 0, \quad \hat{\nabla}_\mu \varrho = 0.$$

Прямые вычисления приводят к следующему тензору Риччи

$$\hat{R}_{\mu\nu} = R_{\mu\nu}^{(k)} + \frac{n-3}{n-2} \frac{\partial_{\mu\nu}^2 \varrho}{\varrho} - \frac{(n-1)(n-3)}{(n-2)^2} \frac{\partial_\mu \varrho \partial_\nu \varrho}{\varrho^2} - \frac{n-3}{n-2} \Gamma_{\mu\nu}^{(k)\rho} \frac{\partial_\rho \varrho}{\varrho} - \frac{1}{(n-2)^2} k_{\mu\nu} \frac{\partial \varrho^2}{\varrho^2} + \frac{1}{n-2} k_{\mu\nu} \frac{\partial_\rho (k^{\rho\sigma} \partial_\sigma \varrho)}{\varrho}, \quad (24.93)$$

где введен “тензор Риччи” $R_{\mu\nu}^{(k)}$, построенный по “метрике” $k_{\mu\nu}$. Отсюда следует выражение для тензорной плотности скалярной кривизны, которая возникает после свертки тензора Риччи с $k^{\mu\nu}$:

$$\check{R} := k^{\mu\nu} \hat{R}_{\mu\nu} = \varrho^{\frac{2}{n-2}} \hat{R} = R^{(k)} + 2 \frac{\partial_\mu (k^{\mu\nu} \partial_\nu \varrho)}{\varrho} - \frac{n-1}{n-2} \frac{(\partial \varrho)^2}{\varrho^2}.$$

После переопределения динамической связи (24.84) гамильтониан также равен линейной комбинации связей:

$$\mathcal{H} = \int d\mathbf{x} (\tilde{N} K_\perp + N^\mu H_\mu),$$

только с новым множителем Лагранжа $\tilde{N} := \varrho^{-\frac{n-1}{n-2}} N$. Соответствующее действие содержит дополнительные слагаемые:

$$S_{\text{HE}} = \int dx (P^{\mu\nu} \dot{k}_{\mu\nu} + P \dot{\varrho} - \tilde{N} K_\perp - N^\mu H_\mu - \lambda (|\det k_{\mu\nu}| - 1) - \mu P^{\mu\nu} k_{\mu\nu}),$$

где мы учли связи (24.77) с помощью множителей Лагранжа λ и μ . Уравнения движения для новых канонических переменных примут вид:

$$\dot{\varrho} = -\frac{n-2}{2(n-1)} \tilde{N} \varrho^2 P + \frac{n-2}{n-1} \hat{\nabla}_\mu N^\mu \varrho, \quad (24.94)$$

$$\dot{P} = \frac{n-2}{2(n-1)} \tilde{N} \varrho P^2 + 2 \tilde{N} \varrho \check{R} + 2 k^{\mu\nu} \hat{\nabla}_\mu \hat{\nabla}_\nu \tilde{N} \varrho + \frac{1}{n-1} \hat{\nabla}_\mu N^\mu P + N^\mu \hat{\nabla}_\mu P, \quad (24.95)$$

$$\dot{k}_{\mu\nu} = 2 \tilde{N} P_{\mu\nu} + \hat{\nabla}_\mu N^\rho k_{\nu\rho} + \hat{\nabla}_\nu N^\rho k_{\mu\rho} - \frac{2}{n-1} \hat{\nabla}_\rho N^\rho k_{\mu\nu}, \quad (24.96)$$

$$\begin{aligned} \dot{P}^{\mu\nu} = & -2 \tilde{N} P^{\mu\rho} P^\nu{}_\rho + \hat{\nabla}_\rho (N^\rho P^{\mu\nu}) + \frac{2}{n-1} \hat{\nabla}_\rho N^\rho P^{\mu\nu} - P^{\mu\rho} \hat{\nabla}_\rho N^\nu - P^{\nu\rho} \hat{\nabla}_\rho N^\mu - \\ & - \varrho^2 k^{\mu\rho} k^{\nu\sigma} \left(\tilde{N} \hat{R}_{\rho\sigma} + \hat{\nabla}_\rho \hat{\nabla}_\sigma \tilde{N} \right) + \frac{\varrho^2 k^{\mu\nu}}{n-1} \left(\tilde{N} \check{R} + k^{\rho\sigma} \hat{\nabla}_\rho \hat{\nabla}_\sigma \tilde{N} \right). \end{aligned} \quad (24.97)$$

Отметим, что множители Лагранжа λ и μ в уравнения движения вообще не входят, т.к. связи являются функциями Казимира. Эта система уравнений движения записана для тензорных плотностей:

$$\begin{aligned} \deg k_{\mu\nu} &= \frac{2}{n-1}, & \deg \varrho &= -\frac{n-2}{n-1}, \\ \deg P^{\mu\nu} &= -\frac{n+1}{n-1}, & \deg P &= -\frac{1}{n-1}. \end{aligned}$$

Новый множитель Лагранжа также является тензорной плотностью:

$$\deg \tilde{N} = 1, \quad \deg N^\mu = 0.$$

Поскольку степени тензорных плотностей при умножении складываются, то

$$\deg(\tilde{N}\varrho) = \frac{1}{n-1}, \quad \deg(\tilde{N}\varrho k^{\mu\nu}) = -\frac{1}{n-1}.$$

Ковариантные производные от тензорных плотностей в системе уравнений (24.94)–(24.97) имеют следующий вид:

$$\begin{aligned}\hat{\nabla}_\mu \tilde{N} &= \partial_\mu \tilde{N} + \frac{n-1}{n-2} \frac{\partial_\mu \varrho}{\varrho} \tilde{N}, \\ \hat{\nabla}_\mu P &= \partial_\mu P - \frac{1}{n-2} \frac{\partial_\mu \varrho}{\varrho} P.\end{aligned}$$

Полученная система уравнений движения (24.94)–(24.97) является полиномиальной.

24.8 Проблема энергии в теории гравитации

Определение энергии гравитационного поля является одной из главных нерешенных проблем, которая привлекает большое внимание с момента создания общей теории относительности. Взгляд на эту проблему существенно меняется с течением времени, поэтому мы опишем несколько подходов.

Сначала сформулируем в чем именно заключается проблема. В механике точечных частиц, а также в теории поля в пространстве Минковского под энергией понимают численное значение гамильтониана системы. Если гамильтониан не зависит от времени явно, то энергия сохраняется. Закон сохранения энергии можно получить также из теоремы Нетер. Если действие инвариантно относительно трансляций во времени $t \mapsto t + \text{const}$, то из первой теоремы Нетер следует закон сохранения энергии (см. раздел 20.2.1), при этом компонента T_0^0 энергии-импульса в теории поля совпадает с определением гамильтоновой плотности системы полей. Этот подход к определению энергии в теории гравитации приводит к ответу, мало удовлетворительному с физической точки зрения. В теории гравитации действие инвариантно относительно общих преобразований координат и, тем более, относительно трансляций во времени. С другой стороны, мы уже показали в разделе 22.7, что канонический гамильтониан для любой модели, инвариантной относительно произвольного невырожденного преобразования временной координаты пропорционален связи и равен нулю на уравнениях движения. Это значит, что формальный подход к определению энергии в теории гравитации дает нуль для любого решения уравнений движения. Этот результат малосодержателен и плохо согласуется с нашим интуитивным представлением об энергии, т.к. наличие материи ассоциируется с наличием энергии. Заметим также, что в пространстве-времени Минковского действие инвариантно относительно трансляций только в декартовой системе координат. В общековариантных моделях гравитации действие инвариантно относительно трансляций в произвольной криволинейной системе координат.

Замечание. В настоящем разделе мы рассмотрим определение энергии в общей теории относительности, т.е. при нулевых тензорах кручения и неметричности. Обобщение определения на более общий случай аффинной геометрии проблемы не составляет, т.к. основная трудность, связанная с общей ковариантностью уравнений движения, присутствует во всех моделях. $\square$

24.8.1 Тензор энергии-импульса полей материи

Первые попытки определить энергию, как сохраняющуюся величину для системы полей материи и гравитационного поля, были предприняты на заре исследований по общей теории относительности. В этой модели действие представляет собой сумму действия Гильберта–Эйнштейна (23.7) и действия для полей материи S_M :

$$S = S_{\text{HE}} + S_M. \quad (24.98)$$

Предположим, что действие полей материи зависит только от полей материи и метрики. Тогда вариация действия по метрике приводит к уравнениям Эйнштейна (23.2), где

$$T_{\alpha\beta} := \frac{2}{\sqrt{|g|}} \frac{\delta S_M}{\delta g^{\alpha\beta}} \quad (24.99)$$

– тензор энергии-импульса полей материи. В общем случае мы предполагаем, что действие для полей материи в моделях гравитации получено из действия, записанного в плоском пространстве-времени Минковского, путем минимальной подстановки, т.е. замены обычных производных на ковариантные $\partial_\alpha \mapsto \nabla_\alpha$ и метрики Минковского на нетривиальную метрику пространства-времени $\eta_{ab} \mapsto g_{\alpha\beta} = e_\alpha^a e_\beta^b \eta_{ab}$. Возможно также введение инвариантных слагаемых неминимального взаимодействия, которые обращаются в нуль в пространстве Минковского. В таких случаях действие зависит только от самих полей материи и репера. Соответствующее выражение для тензора энергии-импульса можно записать через вариационную производную по реперу:

$$T_{\alpha\beta} = \frac{2}{\sqrt{|g|}} \frac{\delta S_M}{\delta e_a^\alpha} e_{\beta a}. \quad (24.100)$$

Если действие для полей материи может быть записано через метрику, то данное определение совпадает с (24.99) и приводит к симметричному тензору энергии-импульса. В общем случае это не так. Например, лагранжиан спинорного поля не может быть записан через метрику, т.к. лоренцева связность выражается через репер, и ее нельзя выразить через компоненты метрики. Определение тензора энергии-импульса через репер (24.100) не всегда приводит к симметричному тензору энергии-импульса.

В разделе (23.5) было показано, что инвариантное действие приводит к ковариантному закону сохранения тензора энергии-импульса:

$$\nabla_\beta T_{\alpha}{}^\beta = 0. \quad (24.101)$$

Однако это ковариантное равенство не является законом сохранения. Покажем это. Прямые вычисления приводят к равенству

$$\nabla_\beta T_{\alpha}{}^\beta = \frac{1}{\sqrt{|g|}} \partial_\beta (\sqrt{|g|} T_{\alpha}{}^\beta) - \frac{1}{2} \partial_\alpha g_{\beta\gamma} T_{\alpha}{}^{\beta\gamma} = 0. \quad (24.102)$$

При интегрировании этого равенства по объему первое слагаемое для каждого значения индекса α преобразуется с помощью формулы Стокса (3.87) в интеграл по граничной поверхности:

$$\int_{\mathbb{M}} dx \partial_\beta (\sqrt{|g|} T_{\alpha}{}^\beta) = \int_{\partial\mathbb{M}} ds_\beta \sqrt{|g|} T_{\alpha}{}^\beta,$$

где ds_α – ориентированный элемент граничной гиперповерхности $\partial\mathbb{M}$. Этот интеграл привел бы к закону сохранения в выбранной системе координат, если бы не наличие второго слагаемого в равенстве (24.102).

24.8.2 Псевдотензор энергии-импульса для гравитации

Для решения проблемы сохранения энергии к тензору энергии-импульса материи, который мы будем записывать с одним верхним и одним нижним индексом $T_{\text{м}\alpha}{}^\beta$, было предложено добавить некоторый объект с компонентами $t_\alpha{}^\beta$, зависящими только от метрики и ее первых производных таким образом, чтобы было выполнено равенство

$$\partial_\beta [\sqrt{|g|}(T_{\text{м}\alpha}{}^\beta + t_\alpha{}^\beta)] = 0. \quad (24.103)$$

Объект с компонентами $t_\alpha{}^\beta$ называется *псевдотензором энергии-импульса* гравитационного поля. При этом равенство (24.103) принято рассматривать, как закон сохранения энергии-импульса полей материи и гравитационного поля. Приставка “псевдо” в данном случае означает, что для выполнения равенства (24.103) компоненты $t_\alpha{}^\beta$ не могут образовывать тензор.

Задача о нахождении явного вида псевдотензора энергии-импульса гравитационного поля $t_\alpha{}^\beta$ может быть решена следующим образом. Закон сохранения (24.103) имеет тот же вид, что и закон сохранения энергии-импульса, вытекающий из первой теоремы Нетер (20.44). Из инвариантности полного действия (24.98) относительно трансляций в пространстве-времени следует закон сохранения

$$\partial_\beta (\sqrt{|g|}T_\alpha{}^\beta) = 0, \quad (24.104)$$

где суммарный тензор энергии-импульса материи и гравитационного поля, умноженный на определитель репера, имеет вид

$$\begin{aligned} \sqrt{|g|}T_\alpha{}^\beta &= \partial_\alpha g^{\gamma\delta} \frac{\partial L_{\text{HE}}}{\partial(\partial_\beta g^{\gamma\delta})} + \partial_\alpha \varphi^a \frac{\partial L_{\text{M}}}{\partial(\partial_\beta \varphi^a)} - \delta_\alpha^\beta (L_{\text{HE}} + L_{\text{M}}) = \\ &= \sqrt{|g|}T_{\text{м}\alpha}^{(\text{c})\beta} + \partial_\alpha g^{\gamma\delta} \frac{\partial L_{\text{HE}}}{\partial(\partial_\beta g^{\gamma\delta})} - \delta_\alpha^\beta L_{\text{HE}}, \end{aligned} \quad (24.105)$$

где φ^a , $a = 1, 2, \dots$, – совокупность всех полей материи. В этом выражении $T_{\text{м}\alpha}^{(\text{c})\beta}$ – канонический тензор энергии-импульса полей материи, определенный соотношением

$$\sqrt{|g|}T_{\text{м}\alpha}^{(\text{c})\beta} := \partial_\alpha \varphi^a \frac{\partial L_{\text{M}}}{\partial(\partial_\beta \varphi^a)} - \delta_\alpha^\beta L_{\text{M}}.$$

В том случае, когда определение тензора энергии-импульса (24.99) является ковариантным обобщением канонического тензора энергии-импульса материи в пространстве Минковского, имеет место формула

$$T_{\text{м}\alpha}{}^\beta = T_{\text{м}\alpha}^{(\text{c})\beta}.$$

Тогда последние два слагаемых в выражении (24.105), зависящие только от метрики и ее первых производных, можно принять за определение псевдотензора энергии-импульса гравитационного поля

$$\sqrt{|g|}t_\alpha{}^\beta := \partial_\alpha g^{\gamma\delta} \frac{\partial L_{\text{HE}}}{\partial(\partial_\beta g^{\gamma\delta})} - \delta_\alpha^\beta L_{\text{HE}}. \quad (24.106)$$

Напомним, что под лагранжианом L_{HE} понимается функция, полученная из $\kappa\sqrt{|g|}R(g)$ добавлением полной производной, которая приводит к сокращению вторых производных от метрики (см. раздел 24.1). Полученное выражение для псевдотензора

энергии-импульса широко использовалось, в том числе классиками науки: Г. Вейлем [177], П. Дираком [178], В. Паули [179], Э. Шредингером [180] и др. После несложных вычислений можно получить явное выражение для псевдотензора энергии-импульса

$$\sqrt{|g|}t_{\alpha}^{\beta} = \kappa(\Gamma_{\gamma\delta}^{\beta} - \delta_{\gamma}^{\beta}\Gamma_{\delta\epsilon}^{\epsilon})\partial_{\alpha}(\sqrt{|g|}g^{\gamma\delta}) - \delta_{\alpha}^{\beta}L_{\text{не}}. \quad (24.107)$$

Отметим недостатки такого определения псевдотензора энергии-импульса гравитационного поля. Из полученного выражения следует, что в нормальной системе координат, где все символы Кристоффеля обращаются в нуль в некоторой точке, все компоненты псевдотензора энергии-импульса гравитационного поля равны нулю. То есть, независимо от кривизны пространства, в любой заданной точке пространства-времени можно обратить в нуль все компоненты псевдотензора энергии-импульса. В то же время в плоском пространстве-времени (отсутствие гравитационного поля) в криволинейной системе координат символы Кристоффеля и, следовательно, компоненты псевдотензора энергии-импульса в общем случае отличны от нуля. Отметим также, что компонента тензора энергии-импульса T_0^0 в законе сохранения (24.104) совпадает с гамильтонианом системы и обращается в нуль на уравнениях движения. Эти замечания ставят под сомнение возможность физической интерпретации псевдотензора энергии-импульса гравитационного поля.

Если равенство (24.103) принято за определение псевдотензора энергии-импульса гравитационного поля, то последний определен неоднозначно. Очевидно, что псевдотензор

$$t'_{\alpha}{}^{\beta} := t_{\alpha}{}^{\beta} + \partial_{\gamma}B_{\alpha}{}^{\gamma\beta}, \quad (24.108)$$

где $B_{\alpha}{}^{\gamma\beta} = -B_{\alpha}{}^{\beta\gamma}$ – произвольный “тензор” третьего ранга, антисимметричный по верхним индексам, также удовлетворяет закону сохранения. В процессе исследования общей теории относительности было предложено несколько явных выражений для псевдотензора энергии-импульса, которые обладают своими достоинствами и недостатками и отличаются между собой на некоторый “тензор” $B_{\alpha}{}^{\gamma\beta}$. Мы не будем останавливаться на обсуждении различных подходов, а отметим лишь общий недостаток. Псевдотензор энергии-импульса $t_{\alpha}{}^{\beta}$ не является тензором. Это значит, что энергия в различных системах координат может быть положительна, равна нулю или отрицательна, что является неудовлетворительным с физической точки зрения.

Аналогичное построение псевдотензора энергии-импульса гравитационного поля можно провести в реперном формализме. При этом все недостатки псевдотензора сохраняются.

Проведенное обсуждение показывает, что в теории гравитации невозможно дать локальное определение энергии-импульса, удовлетворив одновременно двум условиям: 1) компоненты энергии-импульса должны образовывать тензорное поле второго ранга и 2) должен быть выполнен закон сохранения (24.103).

24.8.3 Законы сохранения и векторы Киллинга

Другой подход к законам сохранения связан с симметриями пространства-времени. Если метрика пространства-времени допускает группу движений, то появляется возможность определить законы сохранения для тензора энергии-импульса полей материи без привлечения понятия псевдотензора энергии-импульса гравитационного поля. Пусть тензор энергии-импульса полей материи – это симметричный тензор второго ранга $T_{\text{м}}^{\alpha\beta} = T_{\text{м}}^{\beta\alpha}$, удовлетворяющий условию (24.101). Предположим, что

метрика $g_{\alpha\beta}$ имеет N векторов Киллинга $K_A = K_A^\alpha \partial_\alpha$, $A = 1, \dots, N$. Рассмотрим N векторов

$$P_A^\alpha := K_A^\beta T_{\beta}{}^\alpha.$$

Тогда справедливо равенство

$$\nabla_\alpha P_A^\alpha = \nabla_\alpha K_A^\beta T_{\beta}{}^\alpha + K_A^\beta \nabla_\alpha T_{\beta}{}^\alpha = 0.$$

Здесь первое слагаемое равно нулю как следствие симметричности тензора энергии-импульса и уравнения Киллинга (15.4). Второе слагаемое обращается в нуль в силу уравнения (24.101). С другой стороны, справедливо тождество (6.63)

$$\nabla_\alpha P_A^\alpha = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} P_A^\alpha).$$

Поэтому, если M – компактная ориентируемая область пространства-времени с краем ∂M , то объемный интеграл можно преобразовать в поверхностный по формуле Стокса:

$$\int_M dx \sqrt{|g|} \nabla_\alpha P_A^\alpha = \int_{\partial M} ds_\alpha P_A^\alpha = 0, \quad (24.109)$$

где ds_α – ориентированный элемент площади края. Таким образом, каждому вектору Киллинга соответствует закон сохранения. В частности, если существует вектор Киллинга $K_0 = \partial_0$ и поля материи исчезают на пространственной бесконечности, то ему соответствует закон сохранения энергии $\partial_0 E = 0$, где

$$E = \int d\mathbf{x} P^0 = \int d\mathbf{x} T_{00}.$$

Отметим, что в этом выражении присутствует гамильтонова плотность только для полей материи.

В четырехмерном пространстве-времени Минковского метрика имеет десять векторов Киллинга: четыре вектора соответствуют трансляциям и шесть – (псевдо-) вращениям. Этим векторам Киллинга соответствуют законы сохранения энергии-импульса и момента количества движения.

24.8.4 Полная гравитационная энергия асимптотически плоского пространства-времени

Дальнейшее развитие общей теории относительности изменило подход к определению полной энергии гравитационного поля и материи. Это определение основано на сферически симметричном решении Шварцшильда и учете граничных вкладов в действие Гильберта–Эйнштейна.

Запишем решение Шварцшильда в асимптотически декартовой системе координат, где координаты x, y, z связаны с координатами r, θ, φ обычными формулами трехмерного евклидова пространства. С этой целью воспользуемся формулами (7.10) для перехода от сферических координат к декартовым:

$$\begin{aligned} ds^2 &= \left(1 - \frac{2M}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{2M}{r}} - r^2(d\theta^2 + \sin^2\theta d\varphi^2) = \\ &= \left(1 - \frac{2M}{r}\right) dt^2 - (dx^2 + dy^2 + dz^2) - \frac{2M}{r^2(r - 2M)}(xdx + ydy + zdz)^2. \end{aligned} \quad (24.110)$$

Метрика Шварцшильда на больших расстояниях в первом порядке по $M/r \rightarrow 0$ принимает вид

$$ds^2 \simeq \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 + \frac{2M}{r}\right) (dx^2 + dy^2 + dz^2) - \frac{4M}{r} (xydx dy + xzdx dz + yzdy dz). \quad (24.111)$$

Очевидно, что в нулевом порядке эта метрика совпадает с метрикой Лоренца. Последнее выражение используется для определения асимптотически плоского пространства-времени.

Определение. Топологически тривиальное пространство-время $\mathbb{M} \approx \mathbb{R}^{1,3}$ называется *асимптотически плоским*, если существует такая система координат t, x, y, z , что его метрика при $r \rightarrow \infty$, где $r := \sqrt{x^2 + y^2 + z^2}$ и всех моментов времени $t \in \mathbb{R}$, достаточно быстро стремится к метрике Шварцшильда (24.111), записанной в декартовой системе координат.

Максимально продолженное пространство-время Шварцшильда описывает черные дыры и не является топологически тривиальным. Понятие асимптотически плоского пространства-времени просто обобщается на многообразия с нетривиальной топологией.

Определение. Пусть пространство-время имеет вид прямого произведения $\mathbb{M} = \mathbb{R} \times \mathbb{U}$, где явно выделено время $x^0 = t \in \mathbb{R}$. Допустим, что существует компактное подмножество $\mathbb{K} \subset \mathbb{U}$ такое, что разность $\mathbb{U} \setminus \mathbb{K}$ представляет собой несвязное объединение открытых множеств $\mathbb{S}_A$, $A = 1, \dots, N$. Если в каждом произведении $\mathbb{R} \times \mathbb{S}_A$ существует такая система координат t, x, y, z , что метрика при $r \rightarrow \infty$ и всех $t \in \mathbb{R}$ достаточно быстро стремится к метрике Шварцшильда (24.111) в декартовой системе координат, то пространство-время $\mathbb{M}$ называется *асимптотически плоским*. $\square$

Рассмотрим действие Гильберта–Эйнштейна в гамильтоновой форме, которое вытекает из выражения для гамильтониана (24.65),

$$S_{\text{HE}} = \int dx (p^{\mu\nu} \dot{g}_{\mu\nu} - NH_{\perp} - N^{\mu} H_{\mu}),$$

где функции хода и сдвига рассматриваются в качестве множителей Лагранжа. Динамическая связь $H_{\perp}$ (24.64) содержит трехмерную скалярную кривизну и, следовательно, вторые производные от пространственных компонент метрики $g_{\mu\nu}$. При вариации действия L_{HE} по компонентам метрики приходится интегрировать по частям, отбрасывая граничные члены. Ниже будет показано, что в асимптотически плоском пространстве-времени (для решения Шварцшильда) граничный вклад, возникающий при интегрировании по частям слагаемого со вторыми производными, отличен от нуля. В настоящее время это граничное слагаемое принято в качестве определения полной энергии асимптотически плоского распределения масс. Рассмотрим данный вопрос подробно.

Воспользуемся формулами (24.2)–(24.4) и запишем трехмерную скалярную кривизну, умноженную на определитель репера, в виде

$$\hat{e}\hat{R} = \partial_{\mu} \left[\hat{e}(\hat{g}^{\mu\nu} \hat{\Gamma}_{\rho\nu}{}^{\rho} - \hat{g}^{\nu\rho} \hat{\Gamma}_{\nu\rho}{}^{\mu}) \right] + \hat{e}\hat{L}_{\text{HE}},$$

где последнее слагаемое $\hat{L}_{\text{HE}}$ квадратично по символам Кристоффеля (его вид в настоящий момент не важен). Первое слагаемое в этом представлении имеет вид дивергенции, которую можно выразить через компоненты метрики:

$$\partial_\mu [\hat{e} \hat{g}^{\mu\nu} \hat{g}^{\rho\sigma} (\partial_\nu g_{\rho\sigma} - \partial_\rho g_{\nu\sigma})],$$

где мы воспользовались выражением символов Кристоффеля через метрику (6.24). При интегрировании этого слагаемого по частям возникает граничное слагаемое, которое, как мы увидим ниже, отлично от нуля. Поэтому для того, чтобы компенсировать его вклад, добавим к исходному действию граничный член с самого начала:

$$S_{\text{HE}} \mapsto S_{\text{HE}} + \kappa \int dx \partial_\mu B^\mu,$$

где

$$B^\mu := N \hat{e} \hat{g}^{\mu\nu} \hat{g}^{\rho\sigma} (\partial_\nu g_{\rho\sigma} - \partial_\rho g_{\nu\sigma}) \quad (24.112)$$

и мы восстановили гравитационную константу связи κ .

Определение. Полной энергией асимптотически плоского распределения масс называется интеграл

$$E := \kappa \sum_A \int_A d\mathbf{x} \partial_\mu B^\mu, \quad (24.113)$$

где компоненты B^μ определены в (24.112), и интегрирование проводится по пространственноподобному сечению $x^0 = \text{const}$. $\square$

Компоненты B^μ образуют вектор относительно глобальных преобразований координат в пространстве $\mathbb{R}^3$. В целом, определение полной энергии неинвариантно и зависит от выбора системы координат.

Для трехмерного евклидова пространства $\mathbb{R}^3$ в декартовой системе координат метрика имеет вид $g_{\mu\nu} = -\delta_{\mu\nu}$. Поэтому $B^\mu = 0$ и, следовательно, полная энергия плоского пространства равна нулю, $E = 0$. Уже на этом этапе видна важность выбора системы координат в определении энергии (24.113). Действительно, если система координат не является асимптотически декартовой, то полная энергия трехмерного евклидова пространства может быть отлична от нуля.

Вычислим полную энергию для решения Шварцшильда и, тем самым, для произвольного асимптотически плоского пространства-времени в первом порядке по $M/r \rightarrow \infty$. Для этого запишем решение Шварцшильда в декартовой системе координат (24.110). Пространственная часть метрики отличается от евклидовой метрики: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, где

$$h_{\mu\nu} = -\frac{2M}{r^2(r-2M)} \begin{pmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{pmatrix}$$

На больших расстояниях, $r \gg M$, эта поправка к метрике стремится к нулю. В определении вектора B^μ (24.112) выражение в скобках имеет первый порядок малости. Поэтому остальные множители достаточно учать в нулевом порядке:

$$N = 1, \quad \hat{e} = 1, \quad \hat{g}^{\mu\nu} = \eta^{\mu\nu}.$$

Простые вычисления дают следующие выражения для компонент вектора B^μ в первом порядке по M/r :

$$B^x \simeq \frac{4Mx}{r^3}, \quad B^y \simeq \frac{4My}{r^3}, \quad B^z \simeq \frac{4Mz}{r^3}.$$

Нормальный единичный ковектор к сфере с центром в начале координат евклидова пространства имеет вид

$$\{n_\mu\} = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right), \quad n^2 = -1.$$

Поэтому

$$B^\mu n_\mu \simeq \frac{4M}{r^2}.$$

Теперь можно вычислить полную энергию. Для каждой компоненты связности $\mathbb{S}_A$ в первом порядке получаем равенство

$$E = \kappa \int d\mathbf{x} \partial_\mu B^\mu = \int_0^\pi d\theta \int_0^{2\pi} d\varphi r^2 \sin \theta (B^\mu n_\mu) \simeq 16\pi\kappa M. \quad (24.114)$$

Следовательно

Чтобы придать полученному выражению более привлекательный вид, восстановим размерные константы. Это полезно делать хотя бы изредка. Компоненты метрики, по определению, являются безразмерными $[g_{\alpha\beta}] = 1$. Поскольку

$$\left[\frac{M}{r}\right] = \frac{\Gamma}{\text{см}},$$

то это отношение необходимо обезразмерить, умножив на некоторую комбинацию гравитационной постоянной и скорости света, которые имеют следующие размерности:

$$[G] = \frac{\text{см}^3}{\Gamma \cdot \text{сек}^2}, \quad [c] = \frac{\text{см}}{\text{сек}}.$$

Это можно сделать единственным образом путем замены

$$M \mapsto \frac{GM}{c^2}. \quad (24.115)$$

Если воспользоваться выражением (23.113) для константы связи κ через гравитационную постоянную G , которое было получено при рассмотрении ньютонова предела в общей теории относительности, то выражение для полной энергии примет вид

$$E = Mc^2. \quad (24.116)$$

То, что гравитационная энергия асимптотически плоского пространства-времени совпадает с энергией покоя для массы в решении Шварцшильда, привлекательно с физической точки зрения и оправдывает определение полной энергии через поверхностный интеграл (24.113). Численное значение полной гравитационной энергии зависит от выбора системы координат, т.к. данное определение не инвариантно относительно преобразований координат.

Поскольку M – постоянная интегрирования уравнений Эйнштейна, то полная энергия в асимптотически плоском пространстве-времени сохраняется.

В заключение запишем действие Гильберта–Эйнштейна с размерными постоянными

$$S_{\text{HE}} = \frac{c^4}{16\pi G} \int dx R. \quad (24.117)$$

Именно в таком виде (обычно при $c = 1$) его часто можно встретить в литературе.

Глава 25

Скалярные и калибровочные поля

Скалярные и электромагнитное поля, вместе со спинорными полями, являются важнейшими объектами квантовой теории поля. В настоящей главе мы напомним основные свойства скалярных и калибровочных полей в пространстве Минковского, рассмотрим эти поля на произвольных многообразиях с заданной аффинной геометрией и обсудим их уравнения в общей теории относительности.

25.1 Действительное скалярное поле

25.1.1 Скалярное поле в пространстве Минковского

Рассмотрим действительное скалярное поле (т.е. функцию) $\varphi(x) \in \mathcal{C}^2(\mathbb{R}^{1,n-1})$ в пространстве Минковского $\mathbb{R}^{1,n-1}$ произвольного числа измерений. Обозначим декартовы координаты через x^α , $\alpha = 0, 1, \dots, n-1$. Свободное скалярное поле описывается квадратичным лагранжианом

$$L = \frac{1}{2}\eta^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - \frac{1}{2}m^2\varphi^2, \quad (25.1)$$

где $\eta^{\alpha\beta} := \text{diag}(+ - \dots -)$ – обратная метрика Лоренца и $m = \text{const} > 0$ – масса скалярного поля. Соответствующее действие имеет вид

$$S = \int dx L.$$

Знаки слагаемых в лагранжиане выбран таким образом, чтобы канонический гамильтониан, рассмотренный ниже, был положительно определен.

Посчитаем размерности. Действие и компоненты метрики, по определению, безразмерны. Координаты имеют размерность длины, $[x^\alpha] = l$, $\alpha = 0, 1, \dots, n-1$ (скорость света мы положили равной единице). Поэтому скалярное поле и масса имеют следующие размерности:

$$[\varphi] = l^{\frac{2-n}{2}}, \quad [m] = l^{-1}. \quad (25.2)$$

Уравнение движения для свободного скалярного поля линейно:

$$S_{,\varphi} := \frac{\delta S}{\delta \varphi} = -(\square + m^2)\varphi = 0, \quad (25.3)$$

где

$$\square := \eta^{\alpha\beta}\partial_\alpha\partial_\beta = \partial_t^2 - \partial_1^2 - \dots - \partial_{n-1}^2$$

– волновой оператор (оператор Даламбера). Это уравнение называется *уравнением Клейна–Гордона–Фока*, которое при $m = 0$ сводится к уравнению Даламбера. Уравнение (25.3) было предложено независимо в работах [181, 182, 183, 184].

Скалярное поле называется свободным, потому что описывается линейным уравнением движения.

Поскольку свободное скалярное поле удовлетворяет волновому гиперболическому уравнению, то для постановки задачи Коши в полупространстве $x^0 \geq 0$ для однозначного определения решения необходимо задать начальные данные на гиперповерхности $x^0 = 0$ для поля φ и его производной по времени $\dot{\varphi} := \partial_0 \varphi$. Говорят, что скалярное поле описывает одну динамическую степень свободы. В квантовой теории поля оно описывает нейтральные скалярные частицы.

Для того, чтобы описать скалярное поле с самодействием, к лагранжиану (25.1) добавляется потенциал взаимодействия $V(\varphi)$,

$$L \mapsto L = \frac{1}{2} \eta^{\alpha\beta} \partial_\alpha \varphi \partial_\beta \varphi - V(\varphi),$$

где $V(\varphi)$ – некоторая достаточно гладкая функция одного аргумента. Для упрощения обозначений мы включили массовый член в определение потенциала $V(\varphi)$. В общем случае функция $V(\varphi)$ зависит от некоторого набора размерных или безразмерных констант связи. Для самодействующего скалярного поля уравнение движения становится нелинейным:

$$S, \varphi := \frac{\delta S}{\delta \varphi} = -\square \varphi - V'(\varphi) = 0, \quad (25.4)$$

где $V' := dV/d\varphi$.

Действие для скалярного поля инвариантно относительно глобального действия группы Пуанкаре $\mathbb{IO}(1, n-1)$. Согласно первой теореме Нетер эта инвариантность приводит к законам сохранения энергии-импульса и момента количества движения (см. разделы 20.2.1, 20.2.2). Выражение для сохраняющегося канонического тензора энергии-импульса (23.170) скалярного поля имеет вид

$$T_{\alpha\beta}^{(c)} = \partial_\alpha \varphi \partial_\beta \varphi - \eta_{\alpha\beta} \left(\frac{1}{2} \partial \varphi^2 - V \right), \quad (25.5)$$

где использовано сокращенное обозначение для кинетического слагаемого

$$\partial \varphi^2 := \eta^{\alpha\beta} \partial_\alpha \varphi \partial_\beta \varphi.$$

Тензор энергии-импульса, очевидно, симметричен по своим индексам.

Спиновый момент скалярного поля (20.53) равен нулю, $S_{\alpha\beta}{}^\gamma = 0$, т.к. при вращениях равна нулю вариация формы скалярного поля. Отсюда вытекает, что в квантовой теории поля скалярное поле описывает частицы с нулевым спином. Момент количества движения скалярного поля полностью определяется орбитальным моментом по формуле (20.51)

$$J_{\alpha\beta}{}^\gamma = M_{\alpha\beta}{}^\gamma := x_\beta T_{\alpha}^{(c)\gamma} - x_\alpha T_{\beta}^{(c)\gamma}. \quad (25.6)$$

Функция $V(\varphi)$ может быть положительно определена и при “неправильном” знаке квадрата массы, $m^2 < 0$. Например, положим (см. рис.25.1a)

$$V(\varphi) = \frac{1}{4} \lambda (\varphi^2 - a^2)^2 = \frac{1}{4} \lambda (\varphi^4 - 2\varphi^2 a^2 + a^4), \quad a > 0. \quad (25.7)$$

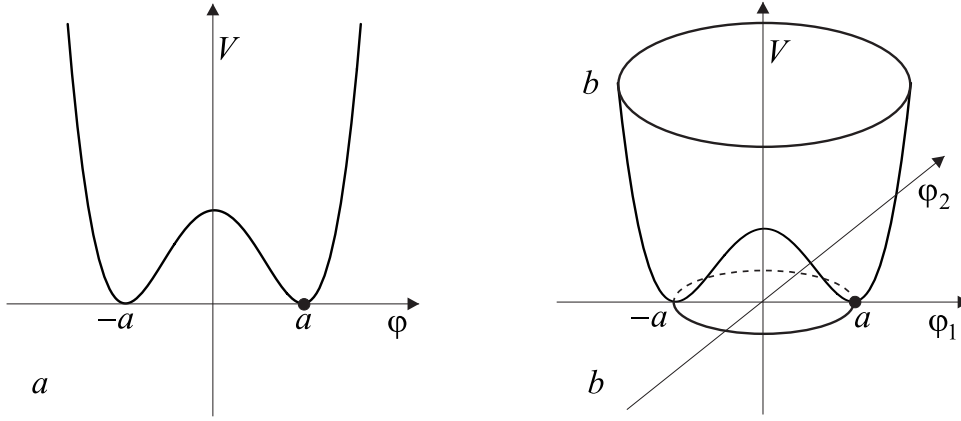


Рис. 25.1: Минимум потенциала действительного (a) и комплексного (b) скалярного поля.

При $\lambda > 0$ потенциал положительно определен, однако “квадрат массы”, $m^2 = -\lambda a^2$, отрицателен. Здесь возникает вопрос: “Что называется массой частицы, соответствующей полю φ ?” В квантовой теории поля масса частиц определяется квадратичным приближением лагранжиана, вокруг которого строится теория возмущений. Для этого необходимо выбрать вакуумное решение уравнений движения, вблизи которого будет происходить разложение полей. Простейшее решение уравнения движения (25.4) – это постоянное поле $\varphi_0 = \text{const}$, которое удовлетворяет условию

$$V'(\varphi_0) = 0.$$

То есть мы предполагаем, что вакуум однороден и статичен. Для потенциала (25.7) уравнение равновесия принимает вид

$$\varphi(\varphi^2 - a^2) = 0.$$

Оно имеет три решения: $\varphi_0 = 0$, $\varphi_0 = \pm a$. Решение $\varphi_0 = 0$ соответствует неустойчивому положению равновесия. Именно поэтому вблизи данного решения квадрат массы имеет неправильный знак. Оба решения $\varphi_0 = \pm a$ соответствуют устойчивым положениям равновесия и имеют одинаковую нулевую плотность энергии. Любое из этих решений можно выбрать в качестве вакуума, вокруг которого строится теория возмущений. Положим

$$\varphi = a + \tilde{\varphi}.$$

Тогда потенциал примет вид

$$V = \lambda \left(a^2 \tilde{\varphi}^2 + a \tilde{\varphi}^3 + \frac{1}{4} \tilde{\varphi}^4 \right).$$

Из квадратичного приближения следует, что масса скалярного поля положительна, $m^2 = 2\lambda a^2$. То есть масса частиц положительна, что соответствует устойчивому положению равновесия.

Перейдем к канонической формулировке. Импульс, канонически сопряженный скалярному полю, равен производной по времени от скалярного поля

$$p := \frac{\partial L}{\partial(\dot{\varphi})} = \dot{\varphi}. \quad (25.8)$$

Одновременные скобки Пуассона для канонически сопряженных переменных имеют вид

$$[\varphi, p'] := [\varphi(t, \mathbf{x}), p(t, \mathbf{x}')] = \delta(\mathbf{x} - \mathbf{x}'),$$

где введена пространственная δ -функция

$$\delta(\mathbf{x} - \mathbf{x}') := \delta(x^1 - x'^1) \dots \delta(x^{n-1} - x'^{n-1}).$$

Соответствующая плотность гамильтониана равна

$$H = T_{00}^{(c)} = \frac{1}{2}p^2 - \frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi + V(\varphi), \quad (25.9)$$

где греческие буквы из середины алфавита пробегают, как и раньше, только пространственные значения: $\mu, \nu = 1, \dots, n-1$. Если функция $V(\varphi)$ положительно определена, то гамильтонова плотность H также положительно определена, т.к. $\eta^{\mu\nu} := -\delta^{\mu\nu}$.

Гамильтониан скалярного поля получается интегрированием гамильтоновой плотности по всему пространству:

$$\mathcal{H} = \int d\mathbf{x} H. \quad (25.10)$$

Если этот интеграл сходится, то полная энергия скалярного поля сохраняется.

Уравнение движения второго порядка (25.3) в гамильтоновой форме эквивалентно системе двух уравнений движения первого порядка относительно производных по времени:

$$\begin{aligned} \dot{\varphi} &= [\varphi, \mathcal{H}] = p, \\ \dot{p} &= [p, \mathcal{H}] = \Delta\varphi - m^2\varphi - V', \end{aligned} \quad (25.11)$$

где Δ – оператор Лапласа.

Компонента $T_{00}^{(c)}$ тензора энергии-импульса скалярного поля (25.5) совпадает с плотностью гамильтониана (25.9). Остальные компоненты канонического тензора энергии-импульса также могут быть выражены через канонические переменные:

$$\begin{aligned} H_\mu &:= T_{0\mu}^{(c)} = p\partial_\mu\varphi, \\ T_{\mu\nu} &= \partial_\mu\varphi\partial_\nu\varphi - \left[\frac{1}{2}(p^2 + \partial^\rho\varphi\partial_\rho\varphi - V(\varphi)) \right] \eta_{\mu\nu}. \end{aligned}$$

Ковектор энергии-импульса $\{P_\alpha\} = \{P_0 := \mathcal{H}, P_\mu\}$ для скалярного поля получается из компонент канонического тензора энергии-импульса интегрированием по пространству:

$$P_0 := \int d\mathbf{x} H, \quad P_\mu := \int d\mathbf{x} H_\mu, \quad (25.12)$$

где интегрирование проводится по сечениям $x^0 = \text{const}$. Численное значение P_0 и P_μ дают полную энергию и полный импульс скалярного поля. Если поле достаточно быстро убывает на бесконечности, что соответствует, в частности, отсутствию излучения, то полные энергия и импульс сохраняются:

$$\dot{P}_0 = 0, \quad \dot{P}_\mu = 0.$$

Отсюда следует, что компоненты канонического тензора энергии-импульса $T_{0\mu}^{(c)}$ имеют физический смысл плотности импульса скалярного поля.

Полный момент импульса скалярного поля получается интегрированием нулевых компонент тензора момента количества движения (25.6):

$$\mathcal{J}_{\alpha\beta} := \int d\mathbf{x} J_{\alpha\beta}^0.$$

Если скалярное поле достаточно быстро убывает на пространственной бесконечности, то он также сохраняется,

$$\dot{\mathcal{J}}_{\alpha\beta} = 0.$$

Строго говоря, полным моментом импульса являются только пространственные компоненты $\mathcal{J}_{\mu\nu} = -\mathcal{J}_{\nu\mu}$.

25.1.2 Скалярное поле в аффинной геометрии

Пусть на многообразии (пространстве-времени) $\mathbb{M}$ произвольной размерности n задана аффинная геометрия, т.е. задана метрика g лоренцевой сигнатуры и связность Γ . Лагранжиан скалярного поля, минимальным образом взаимодействующего с гравитацией, выбирается в виде

$$L = \sqrt{|g|} \left(\frac{1}{2} \partial\varphi^2 - V(\varphi) \right), \quad (25.13)$$

где введено сокращенное обозначение,

$$\partial\varphi^2 := g^{\alpha\beta} \partial_\alpha \varphi \partial_\beta \varphi,$$

для кинетической части лагранжиана. Лагранжиан (25.13) зависит только от метрики, а аффинная связность в него не входит. Это означает, что лагранжиан скалярного поля при минимальной подстановке имеет один и тот же вид как в римановой, так и в аффинной геометрии.

Размерности полей в аффинной геометрии те же, что и в пространстве Минковского, т.к. метрика, по определению, безразмерна.

Вычислим вариационные производные действия:

$$S,_{\varphi} := \frac{\delta S}{\delta \varphi} = -\sqrt{|g|} \left(\tilde{\square} \varphi + V'(\varphi) \right) = 0, \quad (25.14)$$

$$S,_{\alpha\beta} := \frac{\delta S}{\delta g^{\alpha\beta}} = \frac{1}{2} \sqrt{|g|} T_{\alpha\beta}, \quad (25.15)$$

где

$$\tilde{\square} := g^{\alpha\beta} \tilde{\nabla}_\alpha \tilde{\nabla}_\beta.$$

– инвариантный волновой оператор, построенный по псевдоримановой метрике $g_{\alpha\beta}$ и

$$T_{\alpha\beta} = \partial_\alpha \varphi \partial_\beta \varphi - g_{\alpha\beta} \left(\frac{1}{2} \partial\varphi^2 - V \right) \quad (25.16)$$

– ковариантное обобщение тензора энергии-импульса (25.5) для пространства Минковского. Уравнение (25.14) является инвариантным уравнением движения для скалярного поля в аффинной геометрии. Однако в него входит не аффинная связность,

которая может быть задана на многообразии, а только символы Кристоффеля. Это происходит потому что скалярное поле не чувствует кручения и неметричности при минимальной подстановке. Взаимодействие скалярного поля с кручением и неметричностью можно ввести неминимальным образом, добавив к лагранжиану соответствующие слагаемые. При этом существует много возможностей, поэтому мы их обсуждать не будем. Вариационная производная (25.16) определяет тензор энергии-импульса скалярного поля, который служит источником для гравитационного поля (компонент метрики) в уравнениях Эйнштейна.

Замечание. То, что вариационная производная действия по метрике (25.16) пропорциональна ковариантному обобщению канонического тензора энергии-импульса не является общим свойством. Дальнейшие примеры покажут, что вариация действия по метрике в общем случае не всегда пропорциональна ковариантному обобщению тензора энергии-импульса. $\square$

Предложение 25.1.1. Если метрика на многообразии M имеет лоренцеву сигнатуру и координата x^0 является временем, то временная компонента тензора энергии-импульса действительного скалярного поля,

$$T_{00} = \partial_0 \varphi \partial_0 \varphi - g_{00} \left(\frac{1}{2} g^{\alpha\beta} \partial_\alpha \varphi \partial_\beta \varphi - V \right), \quad (25.17)$$

при $V \geq 0$ положительно определена и, следовательно, удовлетворяет слабому энергетическому условию (23.138).

Доказательство. Обозначим $u^\alpha := g^{\alpha\beta} \partial_\beta \varphi$. Тогда

$$\partial_0 \varphi = g_{00} u^0 + g_{0\mu} u^\mu.$$

Легко проверить равенство

$$(\partial_0 \varphi)^2 - \frac{1}{2} g_{00} g^{\alpha\beta} \partial_\alpha \varphi \partial_\beta \varphi = \frac{1}{2} (g_{00} u^0 - g_{0\mu} u^\mu)^2 - \frac{1}{2} g_{00} \left(g_{\mu\nu} - \frac{g_{0\mu} g_{0\nu}}{g_{00}} \right) u^\mu u^\nu. \quad (25.18)$$

Если метрика имеет лоренцеву сигнатуру и координаты выбраны так, что x^0 — это время, то из теоремы 4.2.2 следует, что $g_{00} > 0$ и матрица

$$g_{\mu\nu} - \frac{g_{0\mu} g_{0\nu}}{g_{00}}$$

отрицательно определена. Это означает, что временная компонента (25.17) является положительно определенной как сумма положительно определенных слагаемых в выражении (25.18). $\square$

Если действие инвариантно относительно общих преобразований координат, которые параметризуются n функциями, то согласно второй теореме Нетер уравнения движения удовлетворяют n тождествам. Допустим, что действие зависит только от метрики и скалярного поля. Тогда инвариантность действия означает равенство нулю вариации

$$\delta S = \int dx \sqrt{|g|} (S,^{\alpha\beta} \delta g_{\alpha\beta} + S,_{\varphi} \delta \varphi) = 0.$$

Отсюда с учетом явного вида вариации компонент метрики (2.104) и скалярного поля (2.99) получаем, что уравнения движения (25.14) и (25.15) удовлетворяют n тождествам:

$$2\tilde{\nabla}_\alpha S,^{\alpha\beta} - S,_{\varphi} \partial_\beta \varphi = 0, \quad (25.19)$$

где $\tilde{\nabla}_\alpha$ – ковариантная производная с символами Кристоффеля.

С формальной точки зрения сдвиги на постоянный вектор, $x^\alpha \mapsto x^\alpha + a^\alpha$, $a^\alpha = \text{const}$, образуют подгруппу группы общих преобразований координат. Поэтому, так же, как и в пространстве Минковского, можно построить полную “энергию” и “импульс” скалярного поля. Эти величины будут сохраняться на уравнениях движения, однако им не всегда можно придать физический смысл, т.к. понятие декартовой системы координат в общем случае отсутствует. Это построение имеет смысл в асимптотически плоском пространстве-времени, когда на больших расстояниях пространство-время приближается к пространству Минковского. Это же верно и для момента количества движения.

25.1.3 Скалярное поле в общей теории относительности

Рассмотрим действительное скалярное поле, которое минимальным образом взаимодействует с гравитацией в общей теории относительности. В этом случае действие имеет вид

$$S = \kappa S_{\text{не}} + S_\varphi, \quad (25.20)$$

где $S_{\text{не}}$ – действие Гильберта–Эйнштейна (23.7) и S_φ – действие для скалярного поля с лагранжианом (25.13). Добавление к действию скалярного поля действия Гильберта–Эйнштейна дает кинетический член для метрики. Поэтому вариация этого действия по обратной метрике приводит к уравнениям движения Эйнштейна

$$\frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta g^{\alpha\beta}} : \quad \kappa \left(\tilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \tilde{R} \right) + g_{\alpha\beta} \Lambda + \frac{1}{2} T_{\alpha\beta} = 0, \quad (25.21)$$

где $T_{\alpha\beta}$ – тензор энергии-импульса скалярного поля (25.16). Поскольку действие Гильберта–Эйнштейна не зависит от скалярного поля, то уравнение движения для скалярного поля остается прежним (25.14). Таким образом, полная система уравнений движения для скалярного поля в общей теории относительности состоит из уравнений (25.14) и (25.21).

Действие (25.20) инвариантно относительно общих преобразований координат. Поэтому, согласно второй теореме Нетер, между уравнениями движения существует линейная зависимость (25.19).

Перепишем уравнения Эйнштейна в другом виде. След уравнений Эйнштейна (25.21),

$$\kappa \left(1 - \frac{n}{2} \right) \tilde{R} + n\Lambda + \frac{1}{2} \left(1 - \frac{n}{2} \right) \partial\varphi^2 + \frac{n}{2} V = 0,$$

позволяет исключить скалярную кривизну. В результате уравнения движения (25.21) можно записать в эквивалентной форме

$$\kappa \tilde{R}_{\alpha\beta} = -\frac{1}{2} \partial_\alpha \varphi \partial_\beta \varphi + \frac{1}{n-2} g_{\alpha\beta} (2\Lambda + V). \quad (25.22)$$

При нулевой космологической постоянной $\Lambda = 0$ и потенциале $V = 0$ уравнения (25.22) существенно упрощаются,

$$\kappa \tilde{R}_{\alpha\beta} = -\frac{1}{2} \partial_\alpha \varphi \partial_\beta \varphi. \quad (25.23)$$

Эта модель безмассового скалярного поля в общей теории относительности привлекает в последнее время большое внимание из-за своей относительной простоты.

Покажем, что уравнение движения для скалярного поля (25.14) являются следствием уравнений Эйнштейна (25.21).

Предложение 25.1.2. Если в некоторой области пространства-времени $\mathbb{U} \subset \mathbb{M}$ градиент скалярного поля отличен от нуля, $\{\partial_\alpha \varphi \neq 0\}$, то в этой области уравнение для скалярного поля (25.14) является следствием уравнений Эйнштейна (25.21).

Доказательство. Подействуем оператором ковариантного дифференцирования $\tilde{\nabla}^\beta$ на уравнение (25.21). В силу свернутых тождеств Бианки (6.121),

$$\tilde{\nabla}^\beta \left(\tilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \tilde{R} \right) = 0,$$

получим равенство

$$\tilde{\nabla}^\beta T_{\alpha\beta} = \partial_\alpha \varphi (\tilde{\square} \varphi + V') = 0.$$

Отсюда вытекает сделанное утверждение. $\square$

Доказанное предложение позволяет вместо решения полной системы уравнений для скалярного поля и метрики ограничиться решением только уравнений Эйнштейна.

В рассматриваемом случае ситуация аналогична точечным частицам в общей теории относительности, предложение 23.9.2.

До сих пор мы рассматривали минимальное взаимодействие скалярного с метрикой. Представляет также интерес неминимальное взаимодействие, поскольку в этом случае возможно появление дополнительной важной локальной инвариантности.

Теорема 25.1.1. Действие

$$S = \int_{\mathbb{M}} dx \sqrt{|g|} \left(\frac{1}{2} \partial \varphi^2 - \frac{n-2}{8(n-1)} \varphi^2 \tilde{R} - \lambda \varphi^{\frac{2n}{n-2}} \right), \quad \lambda = \text{const}, \quad (25.24)$$

с точностью до граничных слагаемых инвариантно относительно преобразований Вейля:

$$\begin{aligned} g_{\alpha\beta} &\mapsto \bar{g}_{\alpha\beta} = e^{2\phi} g_{\alpha\beta}, \\ \varphi &\mapsto \bar{\varphi} = e^{-\frac{n-2}{2}\phi} \varphi, \end{aligned} \quad (25.25)$$

где $\phi(x) \in \mathcal{C}^2(\mathbb{M})$ – произвольная функция.

Доказательство. Прямая проверка. При этом удобно использовать формулу преобразования скалярной кривизны (17.40) и правило интегрирования по частям (6.72). $\square$

Учитывая размерность скалярного поля (25.2), получаем, что константа связи λ самодействия скалярного поля в действии (25.24) безразмерна.

Коэффициенты перед первыми двумя слагаемыми в действии (25.24) фиксированы требованием инвариантности относительно преобразований Вейля. Их можно умножить только на общую отличную от нуля постоянную. Фиксирован также показатель степени в преобразовании Вейля для скалярного поля (25.25). Нетрудно также проверить, что добавление к действию (25.24) действия Гильберта–Эйнштейна и космологической постоянной нарушает вейлевскую инвариантность.

Уравнения движения для действия (25.24) имеют следующий вид

$$\begin{aligned} \frac{1}{\sqrt{|g|}} S,^{\alpha\beta} : \quad & \frac{n-2}{8(n-1)} \left[\varphi^2 \left(\tilde{R}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} \tilde{R} \right) - g^{\alpha\beta} \square \varphi^2 + \tilde{\nabla}^\alpha \tilde{\nabla}^\beta \varphi^2 \right] + \\ & + \frac{1}{4} g^{\alpha\beta} \partial \varphi^2 - \frac{1}{2} \partial^\alpha \varphi \partial^\beta \varphi - \frac{1}{2} \lambda g^{\alpha\beta} \varphi^{\frac{2n}{n-2}} = 0, \end{aligned} \quad (25.26)$$

$$\frac{1}{\sqrt{|g|}} S, \varphi : \quad -\tilde{\square} \varphi - \frac{n-2}{4(n-1)} \tilde{R} \varphi - \lambda \frac{2n}{n-2} \varphi^{\frac{n+2}{n-2}} = 0. \quad (25.27)$$

Эти уравнения движения не инвариантны, а ковариантны относительно преобразований Вейля, т.к. при преобразовании умножаются на некоторую отличную от нуля функцию.

Поскольку действие (25.24) инвариантно относительно преобразований Вейля, то в силу второй теоремы Нетер между уравнениями движения существует линейная зависимость. Чтобы ее найти, выпишем вариации формы для метрики и скалярного поля при вейлевских преобразованиях (25.25):

$$\delta g_{\alpha\beta} = 2\epsilon g_{\alpha\beta}, \quad \delta \varphi = -\frac{n-2}{2} \epsilon \varphi,$$

где $\epsilon(x)$ – бесконечно малый параметр преобразования. Поскольку действие инвариантно, то выполнено условие

$$\delta S = \int dx \left(\frac{\delta S}{\delta g_{\alpha\beta}} \delta g_{\alpha\beta} + \frac{\delta S}{\delta \varphi} \delta \varphi \right) = 0.$$

Подставив сюда выражения для вариации полей, получим зависимость уравнений движения:

$$2S,^{\alpha\beta} g_{\alpha\beta} - \frac{n-2}{2} \varphi S, \varphi = 0. \quad (25.28)$$

Нетрудно проверить, что уравнения движения (25.26), (25.27) действительно удовлетворяют этому тождеству.

В важном случае четырехмерного пространства-времени, действие принимает хорошо известную форму

$$S = \int_{\mathbb{M}} dx \sqrt{|g|} \left(\frac{1}{2} \partial \varphi^2 - \frac{1}{12} \varphi^2 \tilde{R} - \lambda \varphi^4 \right). \quad (25.29)$$

Оно инвариантно относительно преобразований Вейля

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = e^{2\phi} g_{\alpha\beta}, \quad \varphi \mapsto \bar{\varphi} = e^{-\phi} \varphi. \quad (25.30)$$

Соответствующие уравнения движения,

$$\begin{aligned} \frac{1}{\sqrt{|g|}} S,^{\alpha\beta} : \quad & \frac{1}{12} \left[\varphi^2 \left(\tilde{R}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} \tilde{R} \right) - g^{\alpha\beta} \square \varphi^2 + \tilde{\nabla}^\alpha \tilde{\nabla}^\beta \varphi^2 \right] - \\ & - \frac{1}{2} \partial^\alpha \varphi \partial^\beta \varphi + \frac{1}{4} g^{\alpha\beta} \partial \varphi^2 - \frac{1}{2} \lambda g^{\alpha\beta} \varphi^4 = 0, \end{aligned} \quad (25.31)$$

$$\frac{1}{\sqrt{|g|}} S, \varphi : \quad -\tilde{\square} \varphi - \frac{1}{6} \tilde{R} \varphi - 4\lambda \varphi^3 = 0, \quad (25.32)$$

ковариантны.

Действие (25.29) является выделенным случаем скалярно-тензорных моделей гравитации, рассмотренных в разделе 23.7. Чтобы привести лагранжиан (23.33) к рассматриваемому виду, необходимо положить

$$\phi = -\frac{\varphi^2}{12}, \quad \omega(\phi) = -\frac{3}{2}, \quad V(\phi) = \lambda\varphi^4, \quad \kappa = 1.$$

Построение действия для других полей, которое было бы инвариантно относительно преобразований Вейля, представляет собой отдельную задачу, на которой мы останавливаться не будем.

25.2 Комплексное скалярное поле

25.2.1 Комплексное скалярное поле в пространстве Минковского

Пусть $\varphi(x) = \varphi_1(x) + i\varphi_2(x)$ – комплексное скалярное поле (функция) в пространстве Минковского $\mathbb{R}^{1,n-1}$, где $\varphi_{1,2}(x) \in \mathcal{C}^2(\mathbb{R}^{1,n-1})$ – два действительных скалярных поля (действительная и мнимая части). Комплексно сопряженное поле будем обозначать крестом $\varphi^\dagger := \varphi_1 - i\varphi_2$, как это принято в квантовой теории поля. Точнее, после вторичного квантования скалярное поле становится операторнозначной функцией, для которой крест означает не комплексное, а эрмитово сопряжение.

Лагранжиан свободного поля

$$\begin{aligned} L &= \eta^{\alpha\beta} \partial_\alpha \varphi^\dagger \partial_\beta \varphi - m^2 \varphi^\dagger \varphi = \\ &= \eta^{\alpha\beta} (\partial_\alpha \varphi_1 \partial_\beta \varphi_1 + \partial_\alpha \varphi_2 \partial_\beta \varphi_2) - m^2 (\varphi_1^2 + \varphi_2^2) \end{aligned} \quad (25.33)$$

с точностью до множителя $1/2$ является суммой лагранжианов (25.1) для действительной и мнимой части. Он является действительной функцией.

Размерность комплексного скалярного поля такая же как и действительного (25.2).

Комплексное скалярное поле задается двумя действительными скалярными полями φ_1 и φ_2 . Для получения уравнений движения действие можно варьировать либо по действительной и мнимой частям φ_1 и φ_2 , либо по полям φ и $\varphi^\dagger$, рассматривая их в качестве независимых переменных. Мы остановимся на второй возможности, как более удобной. Тогда вариация действия по $\varphi^\dagger$ приводит к волновому уравнению для φ (уравнению Клейна–Гордона–Фока):

$$\frac{\delta S}{\delta \varphi^\dagger} = -(\square + m^2)\varphi = 0, \quad (25.34)$$

которое эквивалентно двум волновым уравнениям для действительной и мнимой части. Вариация действия по φ дает уравнение, сопряженное к (25.34).

При рассмотрении задачи Коши для волнового уравнения (25.34) необходимо задать начальные условия независимо для полей φ_1 и φ_2 . Поэтому комплексное скалярное поле описывает две степени свободы и в квантовой теории поля описывает заряженные скалярные частицы.

Для описания самодействия комплексного скалярного поля вводится потенциал $V(\varphi^\dagger, \varphi)$, который в общем случае является достаточно гладкой вещественнозначной и положительно определенной функцией двух аргументов $\varphi^\dagger$ и φ . Мы предположим,

что потенциал зависит только от одного аргумента – произведения $\varphi^\dagger\varphi$ – и обозначим $V' := dV/d(\varphi^\dagger\varphi)$. Тогда лагранжиан примет вид

$$L = \eta^{\alpha\beta} \partial_\alpha \varphi^\dagger \partial_\beta \varphi - V(\varphi^\dagger\varphi), \quad (25.35)$$

где, как и в случае действительного скалярного поля, мы включили массовый член в определение потенциала V . Уравнение движения для скалярного поля с самодействием становится нелинейным:

$$\frac{\delta S}{\delta \varphi^\dagger} = -(\square + V')\varphi = 0, \quad (25.36)$$

Как и в случае действительного скалярного поля, действие для комплексного скалярного поля инвариантно относительно группы Пуанкаре $\mathbb{O}(1, n-1)$. При этом трансляциям соответствует сохраняющийся канонический тензор энергии-импульса

$$T_{\alpha\beta}^{(c)} = \partial_\alpha \varphi^\dagger \partial_\beta \varphi + \partial_\beta \varphi^\dagger \partial_\alpha \varphi - \eta_{\alpha\beta} L. \quad (25.37)$$

Его компоненты, очевидно, вещественны. Спиновый момент комплексного скалярного поля равен нулю и момент количества движения равен орбитальному моменту (25.6).

Действие для комплексного скалярного поля инвариантно также относительно глобальных $\mathbb{U}(1)$ преобразований, меняющих фазу скалярного поля сразу во всем пространстве

$$\varphi' = e^{i\epsilon} \varphi, \quad \varphi'^\dagger = e^{-i\epsilon} \varphi^\dagger, \quad \epsilon = \text{const.} \quad (25.38)$$

Согласно первой теореме Нетер, этой инвариантности соответствует сохраняющийся ток (20.37)

$$\partial_\alpha J^\alpha = 0.$$

где

$$J_\alpha = i(\varphi^\dagger \partial_\alpha \varphi - \partial_\alpha \varphi^\dagger \varphi). \quad (25.39)$$

Позже мы увидим, что $\mathbb{U}(1)$ инвариантность модели комплексного скалярного поля соответствует закону сохранения заряда.

Рассмотрим гамильтонову формулировку теории. Из лагранжиана (25.35) следуют выражения для канонических импульсов, сопряженных полям φ и $\varphi^\dagger$,

$$p := \frac{\partial L}{\partial(\partial_0 \varphi)} = \partial_0 \varphi^\dagger, \quad (25.40)$$

$$p^\dagger := \frac{\partial L}{\partial(\partial_0 \varphi^\dagger)} = \partial_0 \varphi. \quad (25.41)$$

Ненулевые одновременные скобки Пуассона для координат и импульсов имеют вид

$$[\varphi(t, \mathbf{x}), p(t, \mathbf{x}')] = [\varphi^\dagger(t, \mathbf{x}), p^\dagger(t, \mathbf{x}')] = \delta(\mathbf{x} - \mathbf{x}').$$

Для положительно определенного потенциала V плотность гамильтониана комплексного скалярного поля,

$$H = p^\dagger p - \eta^{\mu\nu} \partial_\mu \varphi^\dagger \partial_\nu \varphi + V(\varphi^\dagger \varphi), \quad (25.42)$$

положительно определена. Канонические уравнения движения принимают вид

$$\begin{aligned}\dot{\varphi} &= p^\dagger, \\ \dot{\varphi}^\dagger &= p, \\ \dot{p} &= \Delta\varphi^\dagger - V'\varphi^\dagger, \\ \dot{p}^\dagger &= \Delta\varphi - V'\varphi.\end{aligned}\tag{25.43}$$

Теперь рассмотрим обобщение модели (25.7) для комплексного скалярного поля. Потенциал самодействия выберем в виде (рис. 25.1b)

$$V(\varphi^\dagger\varphi) = \frac{1}{2}\lambda[(\varphi^\dagger\varphi)^2 - a^2]^2 = \frac{1}{2}\lambda[(\varphi^\dagger\varphi)^2 - 2\varphi^\dagger\varphi a^2 + a^4], \quad a > 0.\tag{25.44}$$

При $\lambda > 0$ этот потенциал положительно определен. Однородное и статичное решение φ_0 для вакуума должно удовлетворять уравнению

$$\varphi_0(\varphi_0^\dagger\varphi_0 - a^2) = 0.$$

Это уравнение имеет бесконечно много решений:

$$\varphi_0 = 0 \quad \text{и} \quad \varphi_0 = a e^{i\alpha}, \quad \alpha = \text{const},$$

которые зависят от параметра α . Это связано с наличием унитарной $U(1)$ глобальной симметрии (25.38). Решение $\varphi_0 = 0$ соответствует неустойчивому положению равновесия. Поэтому выберем вакуумное решение в виде $\varphi_0 = a$ и разложим поля вблизи данного решения:

$$\varphi_1 = a + \tilde{\varphi}_1, \quad \varphi_2 = \varphi_2.\tag{25.45}$$

В новых переменных потенциал примет вид

$$V = \frac{1}{2}\lambda[4a^2\tilde{\varphi}_1^2 + 4a\tilde{\varphi}_1^3 + 4a\tilde{\varphi}_1\varphi_2^2 + \tilde{\varphi}_1^4 + \varphi_2^4 + 2\tilde{\varphi}_1^2\varphi_2^2].$$

Его квадратичная часть равна

$$V \simeq 2\lambda a^2 \tilde{\varphi}_1^2.$$

Таким образом, вещественная часть скалярного поля имеет положительную массу, в то время как мнимая часть вблизи вакуумного решения является безмассовой. Это ясно из рис.25.1b. Точка $\varphi = a$ соответствует минимуму потенциальной энергии, однако в направлении φ_2 этот минимум является безразличным.

Описанная выше ситуация возникновения безмассовых полей в моделях с симметриями является типичной. Допустим, что модель зависит от некоторого набора скалярных полей $\varphi = \{\varphi^a\}$, $a = 1, \dots, N$, которые можно рассматривать, как компоненты некоторого вектора. Предположим также, что потенциал взаимодействия инвариантен относительно некоторой группы Ли преобразований $\mathbb{G}$. Если группа Ли K -мерна и вблизи единицы параметризуется набором параметров ϵ^A , $A = 1, \dots, K$, то инфинитезимальные преобразования полей имеют вид

$$\varphi(x) \mapsto \varphi(x) + \delta\varphi(x) = \varphi(x) + \epsilon^A R_A^a(\varphi),$$

где $R_A^a(\varphi)$ – некоторые функции, возможно, нелинейные, полей φ^a . Из инвариантности потенциала вытекает равенство нулю его вариации:

$$\delta V(\varphi) = \epsilon^A R_A^a \frac{\partial V}{\partial \varphi^a} = 0.\tag{25.46}$$

Это условие должно выполняться при всех значениях φ . Допустим, что существует вакуумное решение $\varphi_0 = \text{const}$, которое соответствует минимуму потенциала $V(\varphi)$. Тогда выполнено уравнение

$$\left. \frac{\partial V}{\partial \varphi^a} \right|_{\varphi_0} = 0.$$

Само вакуумное решение может нарушать симметрию полностью или частично. Если все векторы $R_A(\varphi_0) \neq 0$ линейно независимы, то симметрия полностью нарушена. Если существуют некоторые наборы параметров ϵ^A , для которых $\epsilon^A R_A(\varphi_0) = 0$, то при соответствующих преобразованиях вакуум не меняется, и соответствующая симметрия является симметрией вакуумного решения. Очевидно, что множество преобразований, сохраняющих вакуум, образуют некоторую подгруппу $\mathbb{H} \subset \mathbb{G}$. Описанное явление называется *спонтанным нарушением симметрии*, и было описано в [185]. Слово “спонтанный” в данном случае означает, что симметрия нарушена не руками в исходном действии, а выбором вакуумного решения уравнений движения.

Теорема 25.2.1 (Голдстоун). *Если потенциал самодействия скалярных полей инвариантен относительно некоторой группы Ли $\mathbb{G}$ преобразований (25.46) и вакуумное решение нарушает симметрию до некоторой собственной подгруппы $\mathbb{H} \subset \mathbb{G}$, то соответствующая модель содержит безмассовые скалярные поля, соответствующие нарушенным симметриям. Число безмассовых полей больше или равно разности $\dim \mathbb{G} - \dim \mathbb{H}$.*

Доказательство. Продифференцируем равенство (25.46) по φ^b :

$$\epsilon^A \left(\frac{\partial R_A^a}{\partial \varphi^b} \frac{\partial V}{\partial \varphi^a} + R_A^a \frac{\partial^2 V}{\partial \varphi^a \partial \varphi^b} \right) = 0.$$

Если φ_0 – вакуумное решение уравнений движения, то первое слагаемое обращается в нуль и, следовательно, должно быть выполнено равенство

$$\epsilon^A R_A^a \left. \frac{\partial^2 V}{\partial \varphi^a \partial \varphi^b} \right|_{\varphi_0} = 0. \quad (25.47)$$

Массы полей определяются квадратичным приближением потенциала вблизи вакуумного решения. Поэтому массовая матрица имеет вид

$$m_{ab}^2 := \left. \frac{\partial^2 V}{\partial \varphi^a \partial \varphi^b} \right|_{\varphi_0}.$$

Собственные значения этой матрицы не могут быть отрицательны, т.к. φ_0 соответствует минимуму потенциала. Она симметрична и может быть приведена к диагональному виду с помощью ортогональной матрицы (теорема 37.1.8). Система линейных уравнений (25.47) для $\epsilon^A R_A$ имеет нетривиальные решения тогда и только тогда, когда массовая матрица вырождена. Каждой нарушенной симметрии соответствует некоторый нетривиальный вектор $\epsilon^A R_A$. Таких линейно независимых векторов ровно $\dim \mathbb{G} - \dim \mathbb{H}$. Это означает, что массовая матрица должна иметь по крайней мере $\dim \mathbb{G} - \dim \mathbb{H}$ нулевых собственных значений. В общем случае возможно появление дополнительных нулевых собственных значений, вызванных не нарушением симметрии, а спецификой модели. Нулевые собственные числа массовой матрицы означают наличие безмассовых скалярных полей. $\square$

Замечание. В доказанной теореме возможная зависимость параметров преобразования ϵ^A от точки пространства-времени никак не использовалась. Поэтому теорема Голдстоуна справедлива как для глобальных, так и для локальных (калибровочных) преобразований. $\square$

Пример 25.2.1. Рассмотрим комплексное скалярное поле с положительно определенным потенциалом

$$V = \frac{1}{4}\nu(\varphi^\dagger\varphi - a^2)^4, \quad \nu, a > 0.$$

Вакуумное решение должно удовлетворять уравнению

$$\varphi_0(\varphi_0^\dagger\varphi_0 - a^2)^3 = 0.$$

Как и для потенциала (25.44) минимуму энергии соответствует бесконечно много решений. Выберем вакуумное решение в виде $\varphi_0 = a$ и произведем вблизи него разложение (25.45). Тогда потенциал примет вид

$$V = \frac{1}{4}\nu(2a\tilde{\varphi}_1 + \tilde{\varphi}_1^2 + \varphi_2^2)^4.$$

Этот потенциал вообще не имеет квадратичных слагаемых, и, значит, оба поля $\tilde{\varphi}_1$ и φ_2 являются безмассовыми. В то же время, вакуумное решение $\varphi_0 = a$ нарушает $\mathbb{U}(1)$ -симметрию. Таким образом, число безмассовых полей превышает разность $\dim \mathbb{G} - \dim \mathbb{H} = 1$. $\square$

В рассмотренном примере появление дополнительного безмассового поля связано не с нарушением симметрии, а со спецификой модели. В квантовой теории поля безмассовые частицы, соответствующие безмассовым скалярным полям в спонтанно нарушенных моделях теории поля, которые связаны именно с нарушением симметрии, называются *голдстоуновскими бозонами*. Их число равно разности $\dim \mathbb{G} - \dim \mathbb{H}$.

25.2.2 Комплексное скалярное поле в аффинной геометрии

В аффинной геометрии минимальная подстановка для комплексного скалярного поля приводит к следующему лагранжиану

$$L = \sqrt{|g|}[g^{\alpha\beta}\partial_\alpha\varphi^\dagger\partial_\beta\varphi - V(\varphi^\dagger\varphi)], \quad (25.48)$$

и аффинная связность в него не входит. Вариация соответствующего действия по полям приводит к следующим уравнениям движения и тензору энергии-импульса

$$S_{,\varphi^\dagger} := \frac{\delta S}{\delta\varphi^\dagger} = -\sqrt{|g|}(\tilde{\square} + V')\varphi = 0, \quad (25.49)$$

$$S_{, \varphi} := \frac{\delta S}{\delta\varphi} = -\sqrt{|g|}(\tilde{\square} + V')\varphi^\dagger = 0, \quad (25.50)$$

$$\frac{\delta S}{\delta g^{\alpha\beta}} = \frac{1}{2}\sqrt{|g|}T_{\alpha\beta}, \quad (25.51)$$

где

$$T_{\alpha\beta} = \partial_\alpha\varphi^\dagger\partial_\beta\varphi + \partial_\beta\varphi^\dagger\partial_\alpha\varphi - g_{\alpha\beta}(g^{\gamma\delta}\partial_\gamma\varphi^\dagger\partial_\delta\varphi - V) \quad (25.52)$$

– ковариантное обобщение канонического тензора энергии-импульса (25.37).

Перепишем тензор энергии-импульса через действительную и мнимую части,

$$T_{\alpha\beta} = 2\partial_\alpha\varphi_1\partial_\beta\varphi_1 + 2\partial_\alpha\varphi_2\partial_\beta\varphi_2 - g_{\alpha\beta}(\partial\varphi_1^2 + \partial\varphi_2^2 - V).$$

Слагаемые с производными имеют тот же вид, что и в случае двух действительных скалярных полей. Поэтому из доказательства предложения 25.1.1 вытекает положительная определенность их временных компонент. Следовательно, справедливо

Предложение 25.2.1. *Если метрика на многообразии $\mathbb{M}$ имеет лоренцеву сигнатуру и координата x^0 является временем, то временная компонента тензора энергии-импульса комплексного скалярного поля,*

$$T_{00} = \partial_0\varphi^\dagger\partial_0\varphi + \partial_0\varphi^\dagger\partial_0\varphi - g_{00}(g^{\gamma\delta}\partial_\gamma\varphi^\dagger\partial_\delta\varphi - V), \quad (25.53)$$

при $V \geq 0$ положительно определена и, следовательно, удовлетворяет слабому энергетическому условию.

Если инвариантное действие зависит только от метрики и скалярного поля, то из инвариантности действия относительно общих преобразований координат следует зависимость уравнений движения:

$$2\tilde{\nabla}_\alpha S,^\alpha_\beta - S,_{\varphi}\partial_\beta\varphi - S,_{\varphi^\dagger}\partial_\beta\varphi^\dagger = 0. \quad (25.54)$$

Эта зависимость уравнений движения будет иметь место, например, в общей теории относительности, когда к действию Гильберта–Эйнштейна добавляется действие для комплексного скалярного поля.

Действие для комплексного скалярного поля инвариантно также относительно $\mathbb{U}(1)$ преобразований в пространстве-мишени (25.38). Эта инвариантность сохраняется и в аффинной геометрии, если компоненты метрики и связности не преобразуются. Таким образом, закон сохранения заряда для комплексного скалярного поля

$$\partial_\alpha J^\alpha = 0,$$

где

$$J^\alpha = i\sqrt{|g|}g^{\alpha\beta}(\varphi^\dagger\partial_\beta\varphi - \varphi\partial_\beta\varphi^\dagger), \quad (25.55)$$

имеет место также на многообразиях с нетривиальной геометрией.

25.2.3 Комплексное скалярное поле в общей теории относительности

В общей теории относительности действие для комплексного скалярного поля при минимальной подстановке имеет вид

$$S = S_{\text{HE}} + S_\varphi,$$

где

$$S_\varphi = \int dx \sqrt{|g|} [g^{\alpha\beta} \partial_\alpha \varphi^\dagger \partial_\beta \varphi - V(\varphi^\dagger \varphi)].$$

Уравнения Эйнштейна имеют прежний вид (25.21), где тензор энергии-импульса определен равенством (25.52). Их необходимо дополнить двумя уравнениями для скалярного поля (25.49) и (25.50).

Как и в случае действительного скалярного поля действие оператора ковариантного дифференцирования на уравнения Эйнштейна приводит к ковариантному закону сохранения тензора энергии-импульса

$$\tilde{\nabla}^\beta T_{\alpha\beta} = \partial_\alpha \varphi^\dagger (\tilde{\square} \varphi + V' \varphi) + \partial_\alpha \varphi (\tilde{\square} \varphi^\dagger + V' \varphi^\dagger) = 0.$$

Ясно, что если выполнены уравнения движения (25.49), (25.50), то тензор энергии-импульса сохраняется. Однако обратное утверждение теперь неверно, и аналога предложения 25.1.2 в комплексном случае нет.

Как и в случае действительного скалярного поля (25.22) из уравнений Эйнштейна можно исключить скалярную кривизну и переписать их в эквивалентном виде

$$\kappa \tilde{R}_{\alpha\beta} = -\partial_\alpha \varphi^\dagger \partial_\beta \varphi + \frac{2}{n-2} g_{\alpha\beta} (\Lambda + V). \quad (25.56)$$

При неминимальной подстановке также возможно появление инвариантности относительно преобразований Вейля.

Предложение 25.2.2. *Действие*

$$S = \int_{\mathbb{M}} dx \sqrt{|g|} \left(g^{\alpha\beta} \partial_\alpha \varphi^\dagger \partial_\beta \varphi - \frac{n-2}{4(n-1)} \varphi^\dagger \varphi \tilde{R} - \lambda (\varphi^\dagger \varphi)^{\frac{n}{n-2}} \right), \quad \lambda = \text{const}, \quad (25.57)$$

с точностью до граничных слагаемых инвариантно относительно преобразований Вейля:

$$\begin{aligned} g_{\alpha\beta} &\mapsto \bar{g}_{\alpha\beta} = e^{2\phi} g_{\alpha\beta}, \\ \varphi &\mapsto \bar{\varphi} = e^{-\frac{n-2}{2}\phi} \varphi, \\ \varphi^\dagger &\mapsto \bar{\varphi}^\dagger = e^{-\frac{n-2}{2}\phi} \varphi^\dagger, \end{aligned} \quad (25.58)$$

где $\phi(x) \in C^2(\mathbb{M})$ – произвольная вещественнозначная функция.

Доказательство. Перепишем действие (25.57) через вещественную и мнимую части:

$$S = \int_{\mathbb{M}} dx \sqrt{|g|} \left(\partial \varphi_1^2 + \partial \varphi_2^2 - \frac{n-2}{4(n-1)} (\varphi_1^2 + \varphi_2^2) \tilde{R} - \lambda (\varphi_1^2 + \varphi_2^2)^{\frac{n}{n-2}} \right).$$

Первые три слагаемых инвариантны как сумма двух инвариантных слагаемых для φ_1 и φ_2 . Инвариантность последнего слагаемого просто проверяется. $\square$

25.3 Электромагнитное поле

Электромагнитное поле является важнейшим геометрическим объектом в математической физике. Оно наблюдается в экспериментах и широко используется на протяжении многих лет.

С геометрической точки зрения в природе существует следующая конструкция. Пространство-время, в котором мы живем, представляет собой четырехмерное многообразие $\mathbb{M}$. Строится главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(1))$ (см. главу 12), базой которого является пространство-время, а структурной группой – мультипликативная группа Ли комплексных чисел $\mathbb{U}(1)$, равных по модулю единице, которая изоморфна группе двумерных вращений $\mathbb{U}(1) \simeq \mathbb{SO}(2)$, и, как многообразие, диффеоморфна

окружности, $\mathbb{U}(1) \approx \mathbb{S}^1$. На этом главном расслоении задается связность (см. главу 13). Связность, в свою очередь, определяет форму связности и, после проектирования на базу с какого-либо сечения, локальную форму связности:

$$\omega = dx^\alpha A_\alpha.$$

Группа $\mathbb{U}(1)$ одномерна, и поэтому форма ω принимает значения в $\mathbb{R}$, которое рассматривается как одномерное векторное пространство. Компоненты локальной формы связности $A_\alpha(x)$, после добавления уравнений Максвелла, называются в физике *потенциалом электромагнитного поля*, а соответствующие компоненты

$$F_{\alpha\beta} := \partial_\alpha A_\beta - \partial_\beta A_\alpha, \quad (25.59)$$

локальной формы кривизны,

$$F = \frac{1}{2} dx^\alpha \wedge dx^\beta F_{\alpha\beta}, \quad (25.60)$$

— называются *напряженностью электромагнитного поля* или просто *электромагнитным полем*.

Как и любая другая форма кривизны, напряженность электромагнитного поля удовлетворяет тождествам Бианки (13.46), которые в рассматриваемом случае имеют вид

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0. \quad (25.61)$$

Обозначим два локальных сечения через

$$e^{ia(x)}, e^{ia'(x)} \in \mathbb{U}(1), \quad a, a' \in [0, 2\pi].$$

Тогда существует некоторая функция $\phi(x)$, связывающая эти сечения:

$$e^{ia'(x)} = e^{i\phi(x)} e^{ia(x)}.$$

При переходе от одного локального сечения к другому компоненты локальной формы связности меняются по правилу (13.16), которое в данном случае принимает простой вид

$$A'_\alpha := A_\alpha + \partial_\alpha \phi. \quad (25.62)$$

Нетрудно видеть, что при калибровочном преобразовании (25.62) компоненты напряженности электромагнитного поля (25.59) не меняются.

Подчеркнем, что преобразование компонент локальной формы связности возникает в дифференциальной геометрии до того, как вводится понятие ковариантной производной (см. раздел 13.1).

Для построения полевых моделей математической физики одного электромагнитного поля недостаточно и поэтому выбирается некоторый набор дополнительных полей $\psi := \{\psi^A(x)\}$, $A = 1, \dots, N$. Если поля вещественнозначные, то группа $\mathbb{U}(1)$ действует на них тривиально, и они описывают в квантовой теории поля нейтральные частицы. Поэтому рассмотрим комплекснозначные поля. Мы предполагаем, что эти поля являются локальными сечениями некоторого векторного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{U}(1), \mathbb{P})$, которое ассоциировано с главным расслоением $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(1))$. Оно имеет ту же базу $\mathbb{M}$ и структурную группу $\mathbb{U}(1)$. Типичным слоем ассоциированного расслоения является комплексное векторное пространство $\mathbb{V}$, комплексной размерности $\dim_{\mathbb{C}} \mathbb{V} = N$, в котором принимает значения поле $\psi \in \mathbb{V}$. Структурная

группа $\mathbb{U}(1)$ действует в $\mathbb{V}$ в соответствии с некоторым представлением. Все неприводимые комплексные представления группы $\mathbb{U}(1)$ хорошо известны (см., например, [186]). Они одномерны и параметризуются произвольным целым числом k :

$$\mathbb{U}(1) \ni e^{i\phi} \mapsto e^{ik\phi} \in \mathbb{C}, \quad k \in \mathbb{Z}. \quad (25.63)$$

Число k обязано быть целым, поскольку для группы $\mathbb{U}(1)$ мы отождествляем точки $\phi = 0$ и $\phi = 2\pi$.

Далее, мы предполагаем, что каждая компонента векторного поля ψ^A при действии структурной группы преобразуется по некоторому неприводимому представлению:

$$\mathbb{U}(1) \ni e^{i\phi} : \quad \psi^A \mapsto e^{ik_A\phi} \psi^A, \quad k_A \in \mathbb{Z}. \quad (25.64)$$

При переходе от одного локального сечения к другому, функция $\phi = \phi(x)$ является некоторой достаточно гладкой функцией на пространстве-времени, $x \in \mathbb{M}$, которая принимает значения в интервале $[0, 2\pi]$, концы которого отождествлены.

Преобразование (25.62), (25.64) является калибровочным, т.к. параметр преобразования $\phi(x)$ зависит от точки пространства-времени. Это преобразование называется также *градиентным*.

Связность в главном расслоении $\mathbb{P}$ определяет связности во всех ассоциированных расслоениях $\mathbb{E}$ и, следовательно, понятие ковариантной производной (13.25) для локальных сечений. Для определения ковариантной производной, необходимо знать представление генератора группы $\mathbb{U}(1)$ в $\mathbb{C}$. Разлагая “матрицу” представления (25.64) вблизи единицы, получим

$$e^{ik_A\phi} = 1 + ik_A\phi + o(\phi).$$

Отсюда вытекает, что числа ik_A задают представления генератора группы $\mathbb{U}(1)$, и, следовательно, ковариантная производная (13.25) каждой компоненты поля ψ^A имеет вид

$$\nabla_\alpha \psi^A := (\partial_\alpha - ik_A A_\alpha) \psi^A. \quad (25.65)$$

Замечательным свойством ковариантной производной является, как легко проверить, ее тензорный характер при калибровочных преобразованиях (25.62) и (25.64):

$$\nabla_\alpha \psi^A \mapsto \nabla'_\alpha \psi'^A = e^{ik_A\phi(x)} \nabla_\alpha \psi^A,$$

где

$$\nabla'_\alpha \psi'^A := (\partial_\alpha - ik_A A'_\alpha) \psi'^A. \quad (25.66)$$

Все приведенные формулы уже встречались нам в главе 13 в общем виде. Калибровочное преобразование (25.64), (25.66) можно интерпретировать либо как переход между двумя сечениями главного и ассоциированного расслоений, которые определяют локальные формы связности и кривизны, либо как вертикальный автоморфизм (пример 12.3.1). В первом случае калибровочное преобразование интерпретируется как пассивное, когда точки расслоения остаются на месте, а меняется только локальное сечение. Во втором случае калибровочное преобразование рассматривается как активное, когда точки расслоения сдвигаются под действием структурной группы.

Замечание. Описанная выше конструкция не зависит от того, задана ли или нет на многообразии $\mathbb{M}$ аффинная геометрия, т.е. метрика и аффинная связность. Ни одна из приведенных выше формул, включая тождества Бианки (25.61), не содержит метрики и аффинной связности. $\square$

При построении моделей математической физики для полей ψ необходимо написать некоторую систему уравнений. Мы предполагаем, что уравнения движения для полей, не взаимодействующих с электромагнитным полем, следуют из принципа наименьшего действия для некоторого инвариантного интеграла

$$S = \int_{\mathbb{M}} dx L(\psi, \partial\psi),$$

где лагранжиан L зависит от полей $\psi = \{\psi^A\}$ и их производных $\partial\psi = \{\partial_\alpha \psi^A\}$. Как правило, лагранжиан также зависит от метрики и аффинной связности, которые используются для построения инвариантов. Тогда взаимодействие с электромагнитным полем вводится путем *минимальной подстановки*, которая заключается в замене всех частных производных на ковариантные:

$$\partial_\alpha \psi^A \mapsto \nabla_\alpha \psi^A := (\partial_\alpha - ik_A A_\alpha) \psi^A, \quad (25.67)$$

и добавлению к исходному лагранжиану лагранжиана для свободного электромагнитного поля:

$$L(\psi, \partial\psi) \mapsto -\frac{1}{4e^2} F^{\alpha\beta} F_{\alpha\beta} + L(\psi, \nabla\psi),$$

где подъем индексов напряженности электромагнитного поля осуществляется с помощью пространственно-временной метрики и e – универсальная постоянная электромагнитного взаимодействия.

Теперь сделаем переобозначение $A_\alpha \mapsto eA_\alpha$ для того, чтобы уравнения движения свободного электромагнитного поля не содержали константу связи e и можно было бы переходить к пределу $e \rightarrow 0$. Тогда лагранжиан примет вид

$$-\frac{1}{4} F^{\alpha\beta} F_{\alpha\beta} + L(\psi, \nabla\psi),$$

где изменится выражение для ковариантной производной

$$\nabla_\alpha \psi^A := \partial_\alpha \psi^A - ik_A A_\alpha \psi^A. \quad (25.68)$$

Коэффициент $q_A := ek_A$ описывает “силу” взаимодействия компоненты поля ψ^A с электромагнитным полем и называется *электрическим зарядом* данного поля. Для слабых электромагнитных полей именно величина q_A входит в закон Кулона (см., например, [169]).

Мы видим, что при минимальной подстановке заряды всех полей квантуются: они должны быть пропорциональны некоторому универсальному заряду e . Причина этого в том, что все неприводимые представления группы $\mathbb{U}(1)$ параметризуются целыми числами (25.63). И это соответствует действительности, т.к. заряды всех известных в настоящее время элементарных частиц кратны заряду электрона. Поэтому естественно считать, что e – это заряд электрона. Таким образом, дифференциальная геометрия диктует нам *квантование электрического заряда*.

Замечание. В моделях, объединяющих электрослабые и сильные взаимодействия, существенную роль играют кварки, электрический заряд которых пропорционален одной трети заряда электрона. В этом нет ничего страшного. Если кварки будут обнаружены экспериментально, то под e надо понимать $1/3$ заряда электрона. $\square$

Замечание. Электроны в квантовой электродинамике описываются спинорным полем, которое имеет четыре комплексные компоненты. В этом случае для каждой из этих компонент следует положить $k_A = -1$. $\square$

В релятивистских теориях поля пространство-время отождествляется с пространством Минковского $\mathbb{R}^{1,n-1}$ некоторого числа измерений. Поскольку пространство Минковского $\mathbb{R}^{1,n-1}$ как многообразие диффеоморфно евклидову пространству $\mathbb{R}^n$, то согласно теореме 12.1.1 все расслоения тривиальны:

$$\mathbb{P} \simeq \mathbb{R}^{1,n-1} \times \mathbb{U}(1), \quad \mathbb{E} \simeq \mathbb{R}^{1,n-1} \times \mathbb{V}.$$

Это означает, что топология пространства-времени в релятивистских моделях при наличии электромагнитного поля всегда тривиальна, и все поля определены глобально в соответствующей системе координат, покрывающей все пространство Минковского $\mathbb{R}^{1,n-1}$. Нетривиальной может быть только геометрия, если кривизна $\mathbb{U}(1)$ связности отлична от нуля.

Однако, в моделях математической физики, как правило, компоненты связности ищутся путем решения некоторых уравнений. Если окажется так, что соответствующие решения определены не во всем пространстве Минковского, а лишь на некотором подмногообразии $\mathbb{U} \subset \mathbb{R}^{1,n-1}$, то возникает главное расслоение $\mathbb{P}(\mathbb{U}, \pi, \mathbb{U}(1))$ с другой базой $\mathbb{U}$. В этом случае топология $\mathbb{U}$ может быть нетривиальна, и можно говорить о топологических эффектах. Такова, например, ситуация с монополями Дирака.

Точечные частицы в (псевдо)римановом пространстве движутся вдоль мировых линий $q_i(t) = \{q_i^\alpha(t)\} \in \mathbb{M}$, $i = 1, 2, \dots$. Если каждой частице приписать заряд e_i , то минимальная подстановка состоит в добавлении к лагранжиану свободной частицы (см. 22.8) лагранжиана взаимодействия:

$$-m\sqrt{g_{\alpha\beta}\dot{q}_i^\alpha\dot{q}_i^\beta} \mapsto -m\sqrt{g_{\alpha\beta}\dot{q}_i^\alpha\dot{q}_i^\beta} + e_i A_\alpha \dot{q}_i^\alpha,$$

где $\dot{q}_i := dq_i/dt$ – скорость частицы.

Мы считаем, что при калибровочном $\mathbb{U}(1)$ преобразовании (25.62) координаты частиц не меняются. Тогда лагранжиан минимального взаимодействия с частицей сам по себе не инвариантен:

$$e_i A_\alpha \dot{q}_i^\alpha \mapsto e_i A_\alpha \dot{q}_i^\alpha + e_i \partial_\alpha \phi \dot{q}_i^\alpha = e_i A_\alpha \dot{q}_i^\alpha + e_i \frac{d\phi}{dt}.$$

Однако при интегрировании вдоль мировой линии частицы,

$$e_i \int_{q_1}^{q_2} dt \partial_\alpha \phi \dot{q}_i^\alpha = e_i [\phi(q_2) - \phi(q_1)],$$

результат зависит только от положения начальной и конечной точек q_1, q_2 , но не от траектории частицы. Отсюда вытекает инвариантность действия при фиксированных граничных условиях и, следовательно, ковариантность уравнений движения для точечной частицы, минимальным образом взаимодействующей с электромагнитным полем.

Замечание. Если рассматривать только замкнутые траектории с началом и концом в точке q_1 , то интеграл

$$\oint dt \dot{q}^\alpha A_\alpha, \quad (25.69)$$

определяет элемент группы голономии $\mathbb{U}(1)$ связности в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(1))$ в точке $q_1 \in \mathbb{M}$ (см. разделы 13.4 и 13.5). $\square$

В моделях математической физики считается, что в экспериментах можно наблюдать только калибровочно инвариантные объекты. Таковыми являются компоненты напряженности электромагнитного поля, которые, как мы увидим ниже, отождествляются с напряженностью электрического и магнитного полей. При этом сам потенциал электромагнитного поля A_α считается ненаблюдаемым, т.к. зависит от выбора калибровки. Однако, наблюдаемым также является элемент группы голономии (25.69), который инвариантен относительно калибровочных преобразований. Это свойство называется эффектом Ааронова–Бома (см. раздел 14.3).

Для точечных частиц заряд e_1 может быть произволен, и его квантование не вытекает из описанного выше подхода.

В настоящей главе мы не будем рассматривать взаимодействие точечных частиц с электромагнитным полем.

Теперь перейдем к описанию свободного электромагнитного поля в пространстве Минковского.

25.3.1 Лагранжева формулировка

Пусть задано пространство Минковского $\mathbb{R}^{1,n-1}$ произвольного числа измерений и декартова система координат x^α , $\alpha = 0, 1, \dots, n-1$. Электромагнитное поле описывается компонентами $A_\alpha(x)$ локальной формы $U(1)$ связности на пространстве Минковского. Оно является ковекторным полем по отношению к преобразованиям координат – в данном случае – по отношению к преобразованиям Лоренца. Временную и пространственные компоненты электромагнитного поля будем обозначать следующим образом:

$$A_\alpha = \{A_0, A_\mu\}, \quad \mu = 1, \dots, n-1. \quad (25.70)$$

То есть греческие буквы из середины алфавита будут пробегать только пространственные значения.

Действие для свободного электромагнитного поля, по определению, имеет вид

$$S_{\text{EM}} = \int dx L_{\text{EM}}, \quad L_{\text{EM}} = -\frac{1}{4} F^{\alpha\beta} F_{\alpha\beta}, \quad (25.71)$$

где подъем и опускание индексов производится с помощью метрики Минковского $\eta_{\alpha\beta} = \text{diag}(+ \dots -)$ и $F_{\alpha\beta}$ – напряженность электромагнитного поля (25.59). Кроме этого мы положили заряд электрона (константу связи) равным единице. Поле $F_{\alpha\beta} = -F_{\beta\alpha}$ является антисимметричным ковариантным тензором второго ранга типа $(0, 2)$ и представляет собой компоненты локальной 2-формы кривизны (25.60) $U(1)$ связности.

Из безразмерности действия (25.71) следует, что размерность электромагнитного потенциала равна

$$[A_\alpha] = l^{1-\frac{n}{2}}. \quad (25.72)$$

В четырехмерном пространстве-времени размерность потенциала равна обратной длине, $[A_\alpha] = l^{-1}$.

Уравнения движения для действия (25.71) линейны:

$$S_{\text{EM}, \alpha} := \frac{\delta S_{\text{EM}}}{\delta A_\alpha} = \partial_\beta F^{\beta\alpha} = 0. \quad (25.73)$$

Поэтому мы говорим, что действие (25.71) описывает свободное электромагнитное поле.

Лагранжиан (25.71), напряженность (25.59) и уравнения движения (25.73) инвариантны относительно калибровочных преобразований (25.62) действующих только на поля и не затрагивающие координаты. Из калибровочной инвариантности, согласно второй теореме Нетер, следует, что между уравнениями движения имеется линейная зависимость:

$$\partial_\alpha S_{\text{ЕМ}},^\alpha = 0. \quad (25.74)$$

Замечание. Легко видеть, что добавление массового члена $-\frac{1}{2}m^2 A^\alpha A_\alpha$ к лагранжиану электромагнитного поля (25.71) нарушает калибровочную инвариантность электродинамики. $\square$

Пример 25.3.1. Напомним, что в *четырёхмерном* пространстве Минковского $\mathbb{R}^{1,3}$ напряженности электрического и магнитного полей связаны с потенциалом $A_\alpha = \{A_0, \mathbf{A}\}$ известными соотношениями (см., например, [169]):

$$\mathbf{E} = \text{grad } A_0 - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{H} = \text{rot } \mathbf{A}, \quad (25.75)$$

где жирным шрифтом обозначены трехмерные векторы. Это значит, что временная компонента электромагнитного поля A_0 является электростатическим потенциалом, а трехмерный ковектор $\mathbf{A}$ – векторным потенциалом магнитного поля. Отсюда вытекает, что все компоненты напряженности электромагнитного поля определяются напряженностями электрического и магнитного полей:

$$F_{\alpha\beta} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & H_3 & -H_2 \\ E_2 & -H_3 & 0 & H_1 \\ E_3 & H_2 & -H_1 & 0 \end{pmatrix} \quad (25.76)$$

Лагранжиан электромагнитного поля (25.71) также можно выразить через напряженности электрического и магнитного полей:

$$L_{\text{ЕМ}} = \frac{1}{2}(\mathbf{E}^2 - \mathbf{H}^2). \quad \square$$

Проанализируем уравнения Эйлера–Лагранжа (25.73) в пространстве Минковского произвольного числа измерений подробнее. Для временной и пространственных компонент электромагнитного поля уравнения движения имеют вид

$$-\Delta A_0 - \partial_0 \partial_\mu A^\mu = 0, \quad (25.77)$$

$$\square A_\mu - \partial_\mu \partial_\nu A^\nu - \partial_\mu \partial_0 A^0 = 0, \quad (25.78)$$

где введены операторы Лапласа и Даламбера:

$$\Delta := -\partial^\mu \partial_\mu = \partial_1^2 + \partial_2^2 + \dots + \partial_{n-1}^2, \quad \square := \partial^\alpha \partial_\alpha = \partial_0^2 - \Delta.$$

Уравнение для временной составляющей (25.77) для каждого момента времени представляет собой уравнение Пуассона в евклидовом пространстве $\mathbb{R}^{n-1}$ с правой частью, определяемой пространственными компонентами. Сформулируем две теоремы о существовании и единственности решения этого уравнения во всем пространстве и в ограниченной области, доказательство которых приведено, например, в [3].

Теорема 25.3.1. *Если правая часть уравнения Пуассона дифференцируема и убывает на бесконечности, то решение уравнения Пуассона во всем пространстве $\mathbb{R}^{n-1}$, $n \geq 2$, существует и единственно в классе обобщенных функций, обращающихся в нуль на бесконечности. При этом решением уравнения Пуассона является дважды непрерывно дифференцируемая функция во всем пространстве и убывающая на бесконечности.*

Теорема 25.3.2. *Если правая часть уравнения Пуассона дифференцируема в ограниченной области $U \subset \mathbb{R}^{n-1}$, $n \geq 2$, с кусочно гладкой границей ∂U и непрерывна в замыкании $\bar{U}$, то решение внутренней задачи Дирихле для уравнения Пуассона существует, единственно и непрерывно зависит от граничных условий на ∂U .*

Если для внешней задачи Дирихле предположить, что правая часть уравнения Пуассона и решение стремятся к нулю на бесконечности, то справедлива аналогичная теорема для внешней задачи Дирихле. Аналогичные теоремы справедливы также для задачи Неймана, когда на границе задается значение нормальной производной искомой функции. В последнем случае решение определяется с точностью до аддитивной постоянной.

Приведенные теоремы показывают, что при достаточно общих предположениях оператор Лапласа в евклидовом пространстве, а также в ограниченной области с заданными граничными условиями обратим, и решение уравнения Пуассона (25.77) можно формально записать в виде

$$A_0 = -\Delta^{-1} \partial_0 \partial_\mu A^\mu, \quad (25.79)$$

где Δ^{-1} – “обратный” оператор Лапласа. Конечно, в строгом смысле оператор Лапласа необратим, т.к. решение уравнения Пуассона в общем случае определяется с точностью до произвольной гармонической функции. Тем не менее такая запись удобна и часто употребляется в физической литературе, предполагая, что решение соответствующего уравнения Пуассона однозначно выделено, например, с помощью граничных условий. Запись (25.79) означает, что в каждый момент времени временная компонента электромагнитного поля однозначно определяется пространственными компонентами.

Теперь рассмотрим уравнение для пространственных компонент (25.78). Пространственные компоненты электромагнитного поля A_μ образуют ковектор в евклидовом пространстве $\mathbb{R}^{n-1}$. Если поле A_μ непрерывно дифференцируемо, то при достаточно общих граничных условиях его можно взаимно однозначно разложить на поперечную (соленоидальную) и продольную (потенциальную) составляющие:

$$A_\mu = A_\mu^T + A_\mu^L, \quad (25.80)$$

которые удовлетворяют следующим условиям:

$$\partial_\mu A^{\mu T} = 0, \quad (25.81)$$

$$\partial_\mu A_\nu^L - \partial_\nu A_\mu^L = 0. \quad (25.82)$$

Формула (25.80) однозначно определяет ковекторное поле A_μ при заданных полях A_μ^T и A_μ^L . Покажем, что пространственная часть электромагнитного поля A_μ также однозначно определяет свою продольную и поперечную составляющую. В силу леммы Пуанкаре (3.3.3) из условия (25.82) следует, что продольная составляющая, по крайней мере локально, является градиентом от некоторой функции $\varphi(x)$, определенной

с точностью до постоянной, которая фиксируется граничными условиями. Для односвязных многообразий, как в нашем случае, это утверждение является глобальным. Запишем данное свойство в виде равенства

$$A_\mu = A_\mu^T + \partial_\mu \varphi. \quad (25.83)$$

Дифференцирование данного соотношения с учетом условия (25.81) приводит к тому, что скалярное поле, соответствующее продольной составляющей, является решением уравнения Пуассона:

$$\Delta \varphi = -\partial_\mu A^\mu, \quad \Delta := -\partial^\mu \partial_\mu, \quad (25.84)$$

которое существует и единственно при достаточно общих граничных условиях (см. теоремы 25.3.1, 25.3.2). После этого поперечная составляющая однозначно определена:

$$A_\mu^T = A_\mu - \partial_\mu \varphi. \quad (25.85)$$

Нетрудно проверить, что ее дивергенция действительно равна нулю. Таким образом, вместо $n-1$ компоненты ковекторного поля A_μ можно рассматривать скалярное поле φ , определяемое уравнением (25.84), и соленоидальное ковекторное поле A_μ^T (25.85). При таком разложении число независимых компонент сохранено, т.к. поперечное ковекторное поле имеет $n-2$ независимые компоненты в силу дополнительного условия (25.81).

В квантовой теории поля теоремы 25.3.1 и 25.3.2 часто не выполняются. В основе этой теории лежит метод вторичного квантования, который существенно опирается на преобразование Фурье. Плоские волны $A_\mu = A_{0\mu}(k) e^{ikx}$, $\mathbf{k} = \{k^\mu\} \in \mathbb{R}^3$ не убывают на бесконечности и, следовательно, условия теоремы 25.3.1 нарушаются. Если $\mathbf{k} \neq 0$, то разложение на продольную и поперечную части проводится с помощью проекционных операторов:

$$A_\mu^T := \left(\delta_\mu^\nu - \frac{k_\mu k^\nu}{\mathbf{k}^2} \right) A_\nu, \quad A_\mu^L := \frac{k_\mu k^\nu}{\mathbf{k}^2} A_\nu, \quad (25.86)$$

где $\mathbf{k}^2 := -k^\mu k^\nu \eta_{\mu\nu}$. Это значит, что ковекторы A^T и A^L , соответственно, перпендикулярны и параллельны направлению распространения волны. Ясно, что разложение плоской волны (25.86) взаимно однозначно. Кроме этого, разложение (25.86) согласуется с данным ранее определением (25.81), (25.82). Отсюда, кстати, и название продольные и поперечные волны.

Замечание. При $n = 2$ поперечная составляющая у электромагнитного поля отсутствует. $\square$

Перепишем уравнения движения (25.77), (25.78) для продольной и поперечной части:

$$-\Delta A_0 + \Delta \dot{\varphi} = 0, \quad (25.87)$$

$$\partial_\mu \square \varphi + \partial_\mu \Delta \varphi - \partial_\mu \ddot{\varphi} = 0, \quad (25.88)$$

$$\square A_\mu^T = 0, \quad (25.89)$$

где уравнение (25.78) также расщеплено на продольную и поперечную части. Для полей, убывающих на бесконечности, уравнение (25.87) для временной компоненты электромагнитного поля легко решается: $A_0 = \dot{\varphi}$, поскольку решение уравнения Лапласа в данном случае единственно. Уравнение (25.88) тождественно удовлетворяется. Поскольку все проделанные операции обратимы, то мы доказали следующее

Предложение 25.3.1. Для полей, убывающих на бесконечности, лагранжевы уравнения движения для электромагнитного поля (25.73) эквивалентны следующей системе уравнений:

$$A_0 = \dot{\varphi}, \quad (25.90)$$

$$\square A_\mu^\tau = 0, \quad (25.91)$$

дополненной определением поля φ (25.83), (25.84).

Тот факт, что n лагранжевых уравнений оказались эквивалентными $n - 1$ уравнению (25.90), (25.91) не является удивительным. Как уже отмечалось, из-за калибровочной симметрии, в силу второй теоремы Нетер, между исходными уравнениями движения имеется одно линейное соотношение (25.74).

Если рассматривать задачу Коши для электромагнитного поля, то мы видим, что она может быть поставлена только для поперечных компонент электромагнитного потенциала. Следовательно, электромагнитное поле описывает только $n - 2$ физических (распространяющихся) степени свободы. При этом скалярное поле φ является произвольной функцией (непрерывно дифференцируемой и удовлетворяющей граничным условиям для рассматриваемой физической задачи), т.к. для нее нет ни одного уравнения. Временная компонента электромагнитного потенциала выражается через φ . Конечно, можно считать наоборот, что компонента A_0 произвольна, а функция φ находится из уравнения (25.90).

Поперечные степени свободы являются распространяющимися и соответствуют электромагнитным волнам. В двух измерениях электромагнитное поле не описывает ни одной степени свободы и электромагнитное излучение отсутствует. В пространстве Минковского трех и четырех измерений электромагнитное поле имеет, соответственно, одну и две степени свободы.

То обстоятельство, что мы не получили ни одного уравнения на продольную составляющую является следствием калибровочной инвариантности электродинамики относительно преобразований (25.62). Нетрудно проверить, что напряженность электромагнитного поля при выполнении уравнения движения (25.90) зависит только от поперечных компонент:

$$F_{0\mu} = \partial_0 A_\mu^\tau, \quad F_{\mu\nu} = \partial_\mu A_\nu^\tau - \partial_\nu A_\mu^\tau. \quad (25.92)$$

Калибровочные преобразования (25.62) для компонент электромагнитного поля выглядят следующим образом:

$$A'_0 = A_0 + \partial_0 \phi, \quad (25.93)$$

$$\varphi' = \varphi + \phi, \quad (25.94)$$

$$A_\mu'^\tau = A_\mu^\tau. \quad (25.95)$$

Используя произвольную функцию $\phi(x)$ можно зафиксировать одну из компонент электромагнитного поля A_0 или φ , т.е. придать им определенные наперед заданные значения, которые, конечно, должны быть согласованы с уравнением (25.90). Это называется *фиксированием калибровки* и упрощает решение уравнений Эйлера–Лагранжа, т.к. устраняет функциональный произвол в общем решении и уменьшает число неизвестных функций. Наложение калибровочных условий необходимо также для построения квантовой теории. Например, можно положить $A_0 = 0$ и $\varphi = 0$. Для полей, убывающих на бесконечности, последнее условие эквивалентно *кулоновской*

калибровке $\partial_\mu A^\mu = 0$. Среди калибровок, наиболее часто использующихся в электродинамике, отметим следующие¹:

$$\begin{aligned}\partial_\mu A^\mu &= 0 & - \text{кулоновская калибровка,} \\ \partial_\alpha A^\alpha &= 0 & - \text{лоренцева калибровка,} \\ A_0 &= 0 & - \text{временная калибровка,} \\ A_3 &= 0 & - \text{аксиальная калибровка.}\end{aligned}\tag{25.96}$$

Из этих калибровок только кулоновская и аксиальная являются каноническими (см. раздел 22.3). Лоренцева калибровка содержит производную по времени: $\partial_0 A^0 + \partial_\mu A^\mu = 0$. Временная калибровка с гамильтоновой точки зрения фиксирует множитель Лагранжа и оставляет произвол в определении пространственных компонент электромагнитного потенциала A_μ .

Мы называем поля A_0 и A_μ^L нефизическими только потому, что для них нет задачи Коши. Однако это не означает, что они всегда равны нулю и не наблюдаются. Напротив, электростатический потенциал A_0 очень даже наблюдается, например, при взаимодействии точечных зарядов в законе Кулона.

25.3.2 Законы сохранения

Действие для свободного электромагнитного поля в пространстве Минковского инвариантно относительно глобального действия группы Пуанкаре и локальных калибровочных преобразований (25.62). Из теоремы Нетер отсюда следуют законы сохранения тензора энергии-импульса, момента количества движения и вектора электромагнитного тока. Поскольку калибровочные преобразования локальны, то вторая теорема Нетер приводит также к зависимости уравнений движения (25.74), которая была отмечена раньше.

Явное выражение для тензора энергии-импульса (23.170) в случае электромагнитного поля принимает вид

$$\hat{T}_{\alpha\beta} := \hat{T}_\alpha{}^\gamma \eta_{\gamma\beta} = -\partial_\alpha A_\gamma F_\beta{}^\gamma + \frac{1}{4} \eta_{\alpha\beta} F^2.\tag{25.97}$$

Первое слагаемое в этой формуле явно несимметрично относительно перестановки индексов. Воспользуемся произволом (20.39) в определении тензора энергии-импульса и определим новый симметричный тензор энергии-импульса

$$T_{\alpha\beta}^s = -F_{\alpha\gamma} F_\beta{}^\gamma + \frac{1}{4} \eta_{\alpha\beta} F^2 = \hat{T}_{\alpha\beta} + \partial_\gamma A_\alpha F_\beta{}^\gamma.\tag{25.98}$$

С учетом уравнений движения для электромагнитного поля (25.73) разность между новым и старым выражением принимает вид

$$T_\alpha{}^\beta{}^s - \hat{T}_\alpha{}^\beta = \partial_\gamma (A_\alpha F^\beta{}^\gamma),\tag{25.99}$$

что согласуется с видом добавочного члена (20.39). Важным обстоятельством является то, что симметричный тензор энергии-импульса (25.98) калибровочно инвариантен. Закон сохранения энергии-импульса имеет вид (20.44),

$$\partial_\beta T_\alpha{}^\beta{}^s = \partial_\beta \hat{T}_\alpha{}^\beta = 0,$$

¹Лоренцеву калибровку ввел датский физик Ludwig Lorenz в 1867 году, а не голландский физик Н. А. Lorentz (на русский язык обе фамилии в настоящее время принято переводить одинаково, хотя ранее фамилия Lorentz переводилась, как Лорентц), в честь которого названы преобразования Лоренца. По этому поводу см. [187]

для любого решения уравнений движения.

Лоренцевы вращения для электромагнитного поля записываются в виде

$$\delta x^\alpha = -x^\beta \omega_\beta{}^\alpha, \quad \delta A_\alpha = \omega_\alpha{}^\beta A_\beta. \quad (25.100)$$

Поэтому для спинного момента (20.53) получаем следующее выражение

$$S_{\alpha\beta}{}^\gamma = A_\alpha F_\beta{}^\gamma - A_\beta F_\alpha{}^\gamma. \quad (25.101)$$

Полный тензор момента количества движения (20.51) состоит из орбитального и спинного моментов:

$$\hat{J}_{\alpha\beta}{}^\gamma = x_\beta \hat{T}_\alpha{}^\gamma - x_\alpha \hat{T}_\beta{}^\gamma + S_{\alpha\beta}{}^\gamma. \quad (25.102)$$

Используя выражения для симметричного тензора энергии-импульса (25.98) и спинного момента (25.101) его можно переписать в виде

$$\hat{J}_{\alpha\beta}{}^\gamma = M_{\alpha\beta}{}^\gamma + \partial_\delta (x_\alpha A_\beta F^{\gamma\delta} - x_\beta A_\alpha F^{\gamma\delta}), \quad (25.103)$$

где

$$M_{\alpha\beta}{}^\gamma := x_\beta T_\alpha{}^{\text{s}\gamma} - x_\alpha T_\beta{}^{\text{s}\gamma}$$

– орбитальный момент, построенный по симметричному тензору энергии-импульса (25.99). Поскольку второе слагаемое в выражении (25.103) имеет вид (20.39), то тензор момента количества движения можно переопределить, положив

$$J_{\alpha\beta}{}^\gamma = M_{\alpha\beta}{}^\gamma = x_\beta T_\alpha{}^{\text{s}\gamma} - x_\alpha T_\beta{}^{\text{s}\gamma}, \quad (25.104)$$

который также сохраняется

$$\partial_\gamma J_{\alpha\beta}{}^\gamma = 0.$$

Отметим полное отсутствие спинного момента в этом выражении.

С калибровочной инвариантностью электродинамики связан закон сохранения заряда

$$\partial_\alpha J^\alpha = 0, \quad (25.105)$$

где

$$J^\alpha = \partial_\beta F^{\beta\alpha} \quad (25.106)$$

– ток, соответствующий первой теореме Нетер, т.е. когда параметр калибровочных преобразований не зависит от точки пространства-времени. В данном случае закон сохранения заряда (25.105) совпадает с зависимостью уравнений движения, вытекающей из второй теоремы Нетер (25.74). При отсутствии полей материи этот закон тривиален, т.к. ток равен нулю в силу уравнений движения (25.73).

25.3.3 Гамильтонова формулировка

Свободное электромагнитное поле является относительно простым и важным примером калибровочной модели. Ниже мы используем общий гамильтонов анализ для систем со связями, описанный в разделе 22.3.

Электромагнитному потенциалу A_α соответствуют канонически сопряженные импульсы

$$P^\alpha := \frac{\partial L_{\text{ЕМ}}}{\partial(\partial_0 A_\alpha)} = F^{\alpha 0}. \quad (25.107)$$

По определению, скобки Пуассона между координатами и импульсами в каждый момент времени имеют вид

$$[A_\alpha, P'^\beta] := [A_\alpha(t, \mathbf{x}), P^\beta(t, \mathbf{x}')] = \delta_\alpha^\beta \delta(\mathbf{x} - \mathbf{x}'), \quad (25.108)$$

где $(n-1)$ -мерная δ -функция зависит только от пространственных координат. Ввиду антисимметрии напряженности электромагнитного поля в теории возникает первичная связь

$$G_1 := P^0 \approx 0, \quad (25.109)$$

где волнистый знак равенства $\approx$ обозначает, что равенство должно быть выполнено на поверхности всех связей. Других первичных связей электродинамика без полей материи не содержит. Гамильтониан получается простым вычислением

$$\mathcal{H} = \int d\mathbf{x} \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - A_0 \partial_\mu P^\mu + \lambda P^0 \right), \quad (25.110)$$

где мы добавили первичную связь (25.109) с множителем Лагранжа λ . Гамильтоновы уравнения движения для построенного гамильтониана имеют следующий вид:

$$\dot{A}_0 = \lambda, \quad (25.111)$$

$$\dot{A}_\mu = -P_\mu + \partial_\mu A_0, \quad (25.112)$$

$$\dot{P}^0 = \partial_\mu P^\mu, \quad (25.113)$$

$$\dot{P}^\mu = \partial_\nu F^{\nu\mu}, \quad (25.114)$$

которые необходимо дополнить уравнением первичной связи (25.109), возникающей при варьировании соответствующего действия по множителю Лагранжа.

В уравнении (25.111) множитель Лагранжа действительно является произвольной функцией. Чтобы убедиться в этом, необходимо проверить самосогласованность уравнений движения при произвольных λ . Поскольку в теории есть связь, то она должна сохраняться во времени. То есть правая часть уравнения (25.113) должна быть равна нулю для любого решения. Вычислим производную по времени:

$$\partial_0(\partial_\mu P^\mu) = [\partial_\mu P^\mu, \mathcal{H}] = \partial_{\mu\nu}^2 F^{\nu\mu} = 0,$$

где мы использовали антисимметрию компонент напряженности. Поэтому, если связь (25.109) выполнена в начальный момент времени, то она будет удовлетворяться и в последующие моменты времени. Таким образом, у нас имеется только четыре уравнения (25.111)–(25.114) на пять функций. Поэтому функцию λ можно рассматривать как произвольную.

Предложение 25.3.2. *Для полей, убывающих на бесконечности, гамильтоновы уравнения движения для электромагнитного поля (25.109), (25.111)–(25.114) эквивалентны лагранжесвым уравнениям (25.73).*

Доказательство. Пусть выполнены гамильтоновы уравнения. Из уравнения (25.112) находим импульсы

$$P_\mu = -\dot{A}_\mu + \partial_\mu A_0 = F_{\mu 0}.$$

Тогда уравнение (25.114) примет вид лагранжева уравнения

$$\partial_\alpha F^{\alpha\mu} = 0.$$

Поскольку $P^0 = 0$, то уравнение (25.113) примет вид $\partial_\mu P^\mu$ или, с учетом выражения для импульсов

$$\partial_\mu F^{\mu 0} = 0.$$

Таким образом, лагранжевы уравнения являются следствием гамильтоновых. При этом мы никак не использовали уравнение для временной компоненты (25.111).

Обратно. Пусть выполнены лагранжевы уравнения. Обозначим $P^\alpha := F^{\alpha 0}$. Тогда выполнены уравнения (25.112) и связь $P^0 = 0$. Лагранжевы уравнения в новых переменных примут вид

$$\partial_\mu P^\mu = 0, \quad -\dot{P}^\mu + \partial_\nu F^{\nu\mu} = 0,$$

что эквивалентно гамильтоновым уравнениям (25.113), (25.114) с учетом условия $P^0 = 0$. Осталось показать, что уравнение (25.111) является следствием лагранжевых уравнений. Ранее мы доказали, что лагранжевы уравнения эквивалентны системе уравнений (25.90), (25.91). В уравнении (25.90) функция φ , а в уравнении (25.111) функция λ являются совершенно произвольными. Поэтому можно положить $\lambda = \ddot{\varphi}$, и тогда эти уравнения совпадут. $\square$

Замечание. В гамильтониан (25.110) необходимо добавить первичную связь, т.к. в противном случае эквивалентности между гамильтоновыми и лагранжевыми уравнениями не будет. $\square$

Продолжим построение гамильтонова формализма согласно общей схеме анализа систем со связями. Производная по времени от первичной связи (25.109) равна

$$\dot{G}_1 = [P^0, \mathcal{H}] = \partial_\mu P^\mu. \quad (25.115)$$

Это приводит к вторичной связи

$$G_2 := \partial_\mu P^\mu \approx 0. \quad (25.116)$$

Нетрудно проверить, что других связей в теории нет. Связи G_1 и G_2 являются функционально независимыми связями первого рода, т.к. их скобка Пуассона равна нулю:

$$[G_1, G_2] = 0.$$

К полному гамильтониану теории мы добавляем все связи:

$$\mathcal{H}_T = \int d\mathbf{x} \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \lambda P^0 + \mu \partial_\mu P^\mu \right), \quad (25.117)$$

где $\lambda(x)$ и $\mu(x)$ – множители Лагранжа. На поверхности связей $G_1 = 0$ и $G_2 = 0$ гамильтонова плотность, очевидно, положительно определена.

В лагранжевом формализме действие инвариантно относительно калибровочных преобразований (25.62), которые параметризуются одной произвольной функцией $\phi(x)$. Согласно второй теореме Нетер это приводит к зависимости уравнений движения (25.74). Покажем, что в гамильтоновом формализме действие инвариантно относительно калибровочных преобразований, параметризующихся не одной, а двумя произвольными функциями, по числу связей первого рода, и уравнения движения удовлетворяют двум тождествам. Удвоение числа локальных симметрий произошло потому что в гамильтоновом формализме калибровочный параметр $\phi(x)$ и его временная производная $\dot{\phi}(x)$ рассматриваются, как независимые функции и параметризуют, соответственно, различные калибровочные преобразования.

Выразим полное действие для свободного электромагнитного поля через координаты и импульсы:

$$S_{\text{т}} = \int dx (P^\alpha \dot{A}_\alpha + \frac{1}{2} P^\mu P_\mu - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \lambda P^0 - \mu \partial_\mu P^\nu). \quad (25.118)$$

Полное действие совпадает с исходным действием (25.71) в лагранжевой формулировке только при $\lambda = 0$ и $\mu = A_0$ и поэтому является более общим. Уравнения Эйлера–Лагранжа для действия (25.118) принимают вид

$$\frac{\delta S_{\text{т}}}{\delta P^0} = \dot{A}_0 - \lambda = 0, \quad (25.119)$$

$$\frac{\delta S_{\text{т}}}{\delta P^\mu} = \dot{A}_\mu + P_\mu + \partial_\mu \mu = 0, \quad (25.120)$$

$$\frac{\delta S_{\text{т}}}{\delta A_0} = -\dot{P}^0 = 0, \quad (25.121)$$

$$\frac{\delta S_{\text{т}}}{\delta A_\mu} = -\dot{P}^\mu - \partial_\nu F^{\mu\nu} = 0, \quad (25.122)$$

$$\frac{\delta S_{\text{т}}}{\delta \lambda} = -P^0 = 0, \quad (25.123)$$

$$\frac{\delta S_{\text{т}}}{\delta \mu} = -\partial_\mu P^\mu = 0. \quad (25.124)$$

Согласно теореме 22.3.1 каждой связи первого рода соответствует калибровочное преобразование, относительно которого полное действие инвариантно. Первичной связи (25.109) соответствует генератор калибровочных преобразований

$$T_1 = \int d\mathbf{x} \epsilon_1 P^0$$

с некоторым малым параметром $\epsilon_1(x)$. При этом преобразуется только временная компонента электромагнитного поля A_0 и множитель Лагранжа λ :

$$\delta_1 A_0 = [A_0, T_1] = \epsilon_1, \quad \delta_1 \lambda = \dot{\epsilon}_1. \quad (25.125)$$

Все остальные поля остаются без изменения. Согласно второй теореме Нетер из инвариантности действия относительно калибровочного преобразования (25.125) следует зависимость уравнений движения

$$\frac{\delta S_{\text{т}}}{\delta A_0} - \partial_t \frac{\delta S_{\text{т}}}{\delta \lambda} = 0.$$

В справедливости этого тождества нетрудно убедиться прямой проверкой.

Вторичной связи первого рода (25.116) также соответствует генератор преобразований

$$T_2 = \int d\mathbf{x} \epsilon_2 \partial_\mu P^\mu,$$

где $\epsilon_2(x)$ – малый параметр второго калибровочного преобразования. В этом случае преобразуются только пространственные компоненты электромагнитного потенциала A_μ и множитель Лагранжа μ :

$$\delta_2 A_\mu = -\partial_\mu \epsilon_2, \quad \delta_2 \mu = \dot{\epsilon}_2.$$

Из инвариантности действия следует зависимость уравнений Эйлера–Лагранжа

$$\partial_\mu \frac{\delta S_T}{\delta A_\mu} - \partial_t \frac{\delta S_T}{\delta \mu} = 0,$$

что также легко проверить. Таким образом, число параметров калибровочных преобразований полного действия S_T в гамильтоновом формализме по сравнению с лоренцевой формулировкой удвоилось. Эта ситуация типична для калибровочных моделей, когда в лагранжевом формализме локальные преобразования зависят как от самого параметра, так и от его производной по времени.

Генератором калибровочных преобразований в лагранжевом формализме (25.62) является линейная комбинация связей первого рода:

$$T = \int dx (\epsilon P^0 - \epsilon \partial_\mu P^\mu).$$

Действительно,

$$\begin{aligned} \delta A_0 &= [A_0, T] = \epsilon, \\ \delta A_\mu &= [A_\mu, T] = \partial_\mu \epsilon. \end{aligned}$$

Теперь проанализируем гамильтоновы уравнения движения (25.119)–(25.124) и связи (25.109), (25.116) без фиксирования какой-либо калибровки. Для определенности и простоты будем рассматривать задачу в евклидовом пространстве $\mathbb{R}^{n-1}$ для гладких функций достаточно быстро убывающих на бесконечности. По аналогии с разложением электромагнитного поля на поперечную и продольную части (25.83) разложим также импульсы:

$$P_\mu = -P^\mu = P_\mu^T + P_\mu^L,$$

где

$$\partial_\mu P^{T\mu} = 0, \quad P_\mu^L = \partial_\mu \chi.$$

В этом разложении скалярное поле $\chi(x)$ для каждого момента времени является решением уравнения Пуассона

$$\Delta \chi = -\partial_\mu P^\mu = -\partial_\mu P^{L\mu}.$$

Для широкого класса граничных условий в односвязных областях разложение импульсов на продольную и поперечную части является взаимно однозначным. Тогда связь (25.116) примет вид

$$\partial_\mu P^\mu = \partial_\mu P^{L\mu} = 0 \quad \Leftrightarrow \quad \Delta \chi = 0.$$

В рассматриваемом случае мы имеем единственное решение $\chi = 0$, что влечет за собой отсутствие продольной составляющей импульса, $P^{L\mu} = 0$.

Запишем уравнение движения для электромагнитного потенциала (25.120) отдельно для продольной и поперечной компоненты:

$$\dot{A}_\mu^T = -P_\mu^T, \quad \dot{A}_\mu^L = -P_\mu^L - \partial_\mu \mu. \quad (25.126)$$

Последнее уравнение можно переписать в терминах потенциала $A_\mu^L = \partial_\mu \varphi$ и один раз проинтегрировать:

$$\dot{\varphi} = -\chi - \mu = \Delta^{-1} \partial_\mu P^\mu - \mu.$$

Отсюда следует, что эволюция продольной A_μ^L составляющей электромагнитного поля произвольна, т.к. определяется множителем Лагранжа.

С учетом уравнения движения для временной компоненты A_0 , мы видим, что эволюция компонент A_0 и A_μ^L совершенно произвольна, т.к. определяется множителями Лагранжа. При этом уравнения связей позволяют найти сопряженные импульсы $P^0 = 0$ и $P^{L\mu} = 0$. В то же время множители Лагранжа не входят в уравнения движения для поперечных компонент A_μ^T и $P^{T\mu}$. Мы видим, что временная и продольные компоненты соответствуют калибровочным степеням свободы и являются нефизическими.

В теории свободного электромагнитного поля связи можно решить и выписать в явном виде эффективное действие для физических степеней свободы (см. раздел 22.3). Ввиду важности этой процедуры, остановимся на ней подробно.

Электродинамика без полей материи содержит две связи первого рода

$$G_A = \{G_1, G_2\} = 0, \quad A = 1, 2.$$

В соответствии с общим правилом (см. раздел 22.3) наложим два калибровочных условия:

$$F^A = \{F^1, F^2\} = \{A_0, \partial_\mu A^\mu\} = 0.$$

Второе калибровочное условие $\partial_\mu A^\mu = 0$ является уравнением только на продольную составляющую электромагнитного поля:

$$\partial_\mu A^\mu = \partial_\mu A^{L\mu} = 0 \quad \Leftrightarrow \quad \Delta\varphi = 0.$$

При нулевых граничных условиях на бесконечности это уравнение в евклидовом пространстве $\mathbb{R}^{n-1}$ имеет единственное решение $\varphi = 0$, т.е. у пространственных компонент ко вектора A_μ отсутствует продольная составляющая. Эта калибровка называется кулоновской.

Обобщенный гамильтониан электродинамики, который получается после добавления всех связей и калибровочных условий, примет вид

$$\mathcal{H}_E = \int dx \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \lambda P^0 + \mu \partial_\mu P^\mu + \nu A_0 + \xi \partial_\mu A^\mu \right),$$

где λ, μ, ν, ξ – множители Лагранжа. Из условия сохранения калибровочных условий и связей во времени получаем уравнения на множители Лагранжа:

$$\lambda = 0, \quad \Delta\mu = 0, \quad \nu = 0, \quad \Delta\xi = 0.$$

При сделанных предположениях уравнения на μ и ξ имеют единственное решение $\mu = 0, \xi = 0$. Таким образом, кулоновская калибровка является канонической, поскольку однозначно фиксирует множители Лагранжа.

Для точечных частиц отличие от нуля определителя (22.32) является необходимым и достаточным условием однозначного определения множителей Лагранжа в расширенном гамильтониане S_E . Вопрос о критерии однозначного определения множителей Лагранжа в теории поля сложен. Действительно, введем обозначение для полной системы связей и калибровочных условий:

$$\Phi^M = \{G_1, F^1, G_2, F^2\}, \quad M = 1, 2, 3, 4.$$

Тогда матрица, составленная из скобок Пуассона, примет вид

$$[\Phi^M, \Phi'^N] = \begin{pmatrix} 0 & -\delta(\mathbf{x}' - \mathbf{x}) & 0 & 0 \\ \delta(\mathbf{x}' - \mathbf{x}) & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta\delta(\mathbf{x}' - \mathbf{x}) \\ 0 & 0 & -\Delta\delta(\mathbf{x}' - \mathbf{x}) & 0 \end{pmatrix}.$$

Вопрос о том, что такое определитель этой матрицы сложен, и мы на нем останавливаться не будем.

Фазовое пространство электродинамики $\mathbb{N}$ бесконечномерно, $\dim \mathbb{N} = 2n \times \infty$, и задается координатами $A_\alpha(\mathbf{x})$ и сопряженными импульсами $P^\alpha(\mathbf{x})$. Физическое подпространство (поверхность связей и калибровочных условий $\Phi = 0$) $\mathbb{M} \subset \mathbb{N}$ размерности $\dim \mathbb{M} = 2(n - 2) \times \infty$ определяется уравнениями $\Phi^M = 0$. Построим на $\mathbb{N}$ такую систему координат, которая фигурирует в теореме 22.2.2. Сначала заметим, что на поверхности связей и калибровочных условий $\Phi^M = 0$ продольные компоненты координат и импульсов отсутствуют:

$$A_\mu|_{\Phi=0} = A_\mu^T, \quad P^\mu|_{\Phi=0} = P^{T\mu}.$$

Исходной системой координат на $\mathbb{N}$ являются координаты A_α и сопряженные импульсы P^α с канонической скобкой Пуассона (25.108). Введем вместо пространственных компонент A_μ и P^μ их разложение на поперечную и продольную составляющую:

$$\{A_\mu, P^\mu\} \mapsto \{A_\mu^T, A_\mu^L, P^{T\mu}, P^{L\mu}\}.$$

Явные формулы для поперечных и продольных компонент имеют вид:

$$\begin{aligned} A_\mu^T &= A_\mu + \frac{\partial_\mu \partial_\nu}{\Delta} A^\nu, & A_\mu^L &= -\frac{\partial_\mu \partial_\nu}{\Delta} A^\nu, \\ P^{T\mu} &= P^\mu + \frac{\partial^\mu \partial_\nu}{\Delta} P^\nu, & P^{L\mu} &= -\frac{\partial^\mu \partial_\nu}{\Delta} P^\nu, \end{aligned}$$

где $\partial^\mu = -\partial_\mu$. Пуассонова структура в новых координатах уже не будет иметь канонического вида, т.к. на них имеются дополнительные соотношения (25.81), (25.82) и такие же соотношения на импульсы. Простые вычисления показывают, что для пространственных компонент отличны от нуля только две скобки Пуассона:

$$[A_\mu^T, P^{T\nu}] = \left(\delta_\mu^\nu + \frac{\partial_\mu \partial^\nu}{\Delta} \right) \delta(\mathbf{x} - \mathbf{x}'), \quad [A_\mu^L, P^{L\nu}] = -\frac{\partial_\mu \partial^\nu}{\Delta} \delta(\mathbf{x} - \mathbf{x}'). \quad (25.127)$$

В качестве независимых координат на поверхности связей и калибровочных условий $\mathbb{M}$ выберем первые $n - 2$ поперечные компоненты электромагнитного поля

$$q^* = \{A_a^T\}, \quad a = 1, \dots, n - 2.$$

Тогда последняя компонента A_{n-1} находится из уравнения поперечности (25.81)

$$A_{n-1}^T = - \int_{-\infty}^{x^{n-1}} dx' \partial_a A^{Ta}(x^0, x^1, \dots, x^{n-2}, x') := \frac{1}{\partial_{n-1}} \partial_a A^{Ta},$$

где мы ввели обозначение $1/\partial_{n-1}$ для оператора, обратного к частной производной ∂_{n-1} . При этом мы предполагаем, что все компоненты полей достаточно быстро убывают при $x^{n-1} \rightarrow -\infty$ и, следовательно, оператор частной производной ∂_{n-1} обратим. Импульсы, сопряженные к координатам q^* , должны иметь с ними каноническую

скобку Пуассона. Учитывая формулу (25.127), нетрудно проверить, что импульсы, сопряженные к q^* , имеют вид

$$p^* = \left\{ P^{ta} - \frac{1}{\partial_{n-1}} \partial^a P^{t, n-1} \right\}, \quad a = 1, \dots, n-2.$$

Координаты q^* и сопряженные им импульсы p^* образуют систему координат на физическом подмногообразии $\mathbb{M} \subset \mathbb{N}$ фазового пространства. Заметим, что при $x^{n-1} \rightarrow \infty$ компоненты импульсов p^* в общем случае не обязательно убывают, поскольку содержат обратный оператор $1/\partial_{n-1}$. Эта ситуация типична в теории поля, где связи представляют собой дифференциальные уравнения.

Таким образом, мы построили канонически сопряженные координаты на поверхности связей. Это – не единственный выбор. В качестве независимых координат можно было бы выбрать непосредственно компоненты $\{q^*, p^*\} = \{A_a, P^a\}$, $a = 1, \dots, n-2$. В этом случае преобразования координат фазового пространства были бы проще, однако выражение для эффективного гамильтониана для физических степеней свободы – сложнее.

Поверхность связей и калибровочных условий $\Phi^M = 0$ можно задать канонически сопряженными координатами и импульсами:

$$Q_A = \{A_0, \partial_\mu A^\mu\}, \quad P^A = \{P^0, \Delta^{-1} \partial_\mu P^\mu\}, \quad A = 1, 2.$$

Действительно, нетрудно проверить, что скобки Пуассона имеют канонический вид:

$$[Q_A, P'^B] = \delta_A^B \delta(\mathbf{x} - \mathbf{x}'), \quad [Q_A, Q'_B] = 0, \quad [P^A, P'^B] = 0.$$

Эти координаты описывают нефизические (калибровочные) степени свободы.

Таким образом, в соответствии с теоремой 22.2.2 мы построили систему координат на фазовом пространстве $\mathbb{N}$

$$\{q^*, p^*, Q_A, P^A\}$$

такую, что канонически сопряженные координаты и импульсы q^*, p^* являются координатами на физическом подпространстве $\mathbb{M} \subset \mathbb{N}$, а поверхность связей и калибровочных условий задается особенно просто:

$$\Phi^M = 0 \quad \Leftrightarrow \quad Q_A = 0, \quad P^A = 0.$$

Пример 25.3.2. В четырехмерном пространстве-времени на поверхности связей и калибровочных условий физический (эффективный) гамильтониан принимает простой вид

$$\mathcal{H}_{\text{ph}} = \int d\mathbf{x} \left(-\frac{1}{2} P^{\tau\mu} P_\mu^\tau + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right) = \frac{1}{2} \int d\mathbf{x} (\mathbf{E}^{\tau^2} + \mathbf{H}^2),$$

где мы учли выражения для напряженностей электрического и магнитного полей (25.75). Напомним, что из определения напряженности магнитного поля следует, что $\mathbf{H} = \mathbf{H}^T$. Таким образом, мы получили стандартное выражение для плотности энергии свободного электромагнитного поля, которое, очевидно, положительно определено. $\square$

Построенная система координат нелокальна в том смысле, что содержит обратные дифференциальные операторы. Эта ситуация типична для калибровочных теорий поля, где связи являются дифференциальными уравнениями по пространственным координатам, поскольку их решения задаются интегралами по пространству.

Решения дифференциальных уравнений в частных производных существенно зависят от области, в которой ищется решение, и граничных условий. В частности, оператор Лапласа может быть необратим, и это может приводить к физическим следствиям.

Таким образом, электродинамика в пространстве Минковского $\mathbb{R}^{1,n-1}$ описывает $n - 2$ физические (распространяющиеся) степени свободы. В четырехмерном пространстве-времени мы говорим, что фотон имеет две поляризации. При этом временная и продольная компоненты электромагнитного поля являются нефизическими (нераспространяющимися). Это не значит, что они равны нулю. Например, в электростатике временная компонента A_0 нетривиальна и наблюдаема, т.к. представляет собой электрический кулонов потенциал взаимодействия зарядов. Нефизичность в рассматриваемом контексте означает только то, что для соответствующих полей нельзя поставить задачу Коши.

В двумерном пространстве-времени, $n = 2$, пространственная часть электромагнитного поля A_μ имеет только одну продольную составляющую. Все координаты и импульсы находятся из решения связей и калибровочных условий. Поэтому электромагнитное поле в двух измерениях не имеет физических (распространяющихся) степеней свободы. В этом смысле модель тривиальна. Тем не менее она представляет интерес, особенно при включении взаимодействия с другими полями.

Гамильтонова формулировка свободной электродинамики на произвольном многообразии $\mathbb{M}$ с метрикой g проводится, по существу, так же как и в пространстве Минковского. Необходимо произвести замену $\eta_{\alpha\beta} \mapsto g_{\alpha\beta}$ и рассматривать компоненты метрики $g_{\alpha\beta}$ как внешнее поле. Почти все формулы останутся при этом без изменения, только в ковариантных производных появятся символы Кристоффеля, возникающие при интегрировании по частям.

25.3.4 Скалярная электродинамика

Рассмотрим простейший пример нетривиального взаимодействия полей материи с электромагнитным полем в пространстве Минковского $\mathbb{R}^{1,n-1}$ произвольного числа измерений. Эта модель – *скалярная электродинамика* – описывает комплексное скалярное поле φ , минимально взаимодействующее с электромагнитным полем A_α . Соответствующий лагранжиан имеет вид

$$L = -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \eta^{\alpha\beta}\nabla_\alpha\varphi^\dagger\nabla_\beta\varphi - V(\varphi^\dagger\varphi), \quad (25.128)$$

где ковариантные производные определены следующими соотношениями:

$$\begin{aligned} \nabla_\alpha\varphi &:= \partial_\alpha\varphi - iqA_\alpha\varphi, \\ \nabla_\alpha\varphi^\dagger &:= \partial_\alpha\varphi^\dagger + iqA_\alpha\varphi^\dagger, \end{aligned} \quad (25.129)$$

и $q = \text{const}$ – заряд поля (как мы увидим позже).

Уравнения движения для скалярной электродинамики имеют вид

$$\frac{\delta S}{\delta A_\alpha} = \partial_\beta F^{\beta\alpha} + iq\varphi^\dagger\nabla^\alpha\varphi - iq\varphi\nabla^\alpha\varphi^\dagger = 0, \quad (25.130)$$

$$\frac{\delta S}{\delta \varphi^\dagger} = -\nabla^\alpha\nabla_\alpha\varphi - V'\varphi = 0, \quad (25.131)$$

$$\frac{\delta S}{\delta \varphi} = -\nabla^\alpha\nabla_\alpha\varphi^\dagger - V'\varphi^\dagger = 0. \quad (25.132)$$

Сравнение этих уравнений с уравнениями Максвелла в той форме, как они обычно записываются в физической литературе (см., например, [169]), показывает, что константа связи q имеет физическую интерпретацию заряда поля.

Лагранжиан (25.128) инвариантен относительно калибровочных преобразований

$$\begin{aligned}\varphi &\mapsto \varphi' = e^{iq\phi}\varphi, \\ \varphi^\dagger &\mapsto \varphi'^\dagger = e^{-iq\phi}\varphi^\dagger, \\ A_\alpha &\mapsto A'_\alpha = A_\alpha + \partial_\alpha\phi,\end{aligned}\tag{25.133}$$

где $\phi(x)$ – локальный параметр преобразования. Согласно второй теореме Нетер отсюда вытекает линейная зависимость уравнений движения:

$$-\partial_\alpha \frac{\delta S}{\delta A_\alpha} + iq\varphi \frac{\delta S}{\delta \varphi} - iq\varphi^\dagger \frac{\delta S}{\delta \varphi^\dagger} = 0.\tag{25.134}$$

В случае свободного электромагнитного поля зависимость уравнений движения (25.74) выглядит тривиально. Для скалярной электродинамики доказать зависимость уравнений (25.130)–(25.132) без обращения к калибровочной инвариантности намного сложнее.

В пространстве Минковского лагранжиан инвариантен относительно действия группы Пуанкаре $\mathbb{ISO}(1, n-1)$. Этой инвариантности, согласно первой теореме Нетер, соответствует закон сохранения энергии-импульса и момента количества движения. Канонический тензор энергии-импульса равен сумме тензора энергии-импульса электромагнитного (25.97) и скалярного полей (25.37):

$$\widehat{T}_{\alpha\beta}^{(c)} = -\partial_\alpha A_\gamma F_\beta{}^\gamma + \frac{1}{4}\eta_{\alpha\beta} F^2 + \partial_\alpha \varphi^\dagger \nabla_\beta \varphi + \nabla_\beta \varphi^\dagger \partial_\alpha \varphi - \eta_{\alpha\beta} [\nabla^\gamma \varphi^\dagger \nabla_\gamma \varphi - V(\varphi^\dagger \varphi)],\tag{25.135}$$

который, очевидно, несимметричен. Как и в случае свободного электромагнитного поля, добавление дивергенции (25.99) с учетом уравнений движения позволяет определить симметричный и калибровочно инвариантный тензор энергии-импульса

$$T_{\alpha\beta}^{(s)} = -F_{\alpha\gamma} F_\beta{}^\gamma + \frac{1}{4}\eta_{\alpha\beta} F^2 + \nabla_\alpha \varphi^\dagger \nabla_\beta \varphi + \nabla_\beta \varphi^\dagger \nabla_\alpha \varphi - \eta_{\alpha\beta} [\nabla^\gamma \varphi^\dagger \nabla_\gamma \varphi - V(\varphi^\dagger \varphi)].\tag{25.136}$$

Выражение для тензора спина в скалярной электродинамике такое же, как и для свободного электромагнитного поля (25.101). Связь полного момента количества движения с орбитальным моментом, определяемым симметричным тензором энергии-импульса также остается прежней (25.104).

Сохраняющийся вектор электромагнитного тока, соответствующий глобальным преобразованиям (25.133), имеет вид

$$J^\alpha = -\partial_\beta F^{\alpha\beta} + iq(\varphi^\dagger \nabla^\alpha \varphi - \varphi \nabla^\alpha \varphi^\dagger).\tag{25.137}$$

Заметим, что зависимость уравнений движения (25.134), вытекающая из второй теоремы Нетер, и сохранение тока $\partial_\alpha J^\alpha$ – не одно и то же.

Спектр модели, т.е. массы, спины и заряды частиц после вторичного квантования, определяется квадратичным приближением лагранжиана, которое, в свою очередь, зависит от вида потенциала самодействия скалярного поля. Если скалярное поле свободно, т.е.

$$V(\varphi^\dagger \varphi) = m^2 \varphi^\dagger \varphi, \quad m > 0,$$

то модель (25.128) в низкоэнергетическом пределе описывает заряженные скалярные частицы массы m , взаимодействующие с безмассовым электромагнитным полем.

Если потенциал скалярного поля описывает самодействие, то ситуация может сильно измениться. Продемонстрируем это на примере потенциала “ $\lambda\varphi^4$ ” (25.44). Сначала перепишем лагранжиан (25.128) через действительную и мнимую части:

$$L = -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \eta^{\alpha\beta}(\nabla_\alpha\varphi_1\nabla_\beta\varphi_1 + \nabla_\alpha\varphi_2\nabla_\beta\varphi_2) - V(\varphi_1^2 + \varphi_2^2),$$

где

$$\begin{aligned}\nabla_\alpha\varphi_1 &:= \partial_\alpha\varphi_1 + qA_\alpha\varphi_2, \\ \nabla_\alpha\varphi_2 &:= \partial_\alpha\varphi_2 - qA_\alpha\varphi_1.\end{aligned}$$

В таком виде скалярные поля φ_1 и φ_2 преобразуются по неприводимому двумерному вещественному представлению унитарной группы $\mathbb{U}(1)$ (векторному представлению группы вращений $\mathbb{SO}(2)$) Вакуумное решение выберем однородным и изотропным:

$$\varphi_0 = a, \quad A_\alpha = 0$$

и произведем разложение скалярного поля (25.45) вблизи вакуума. В новых переменных лагранжиан (25.128) примет вид

$$\begin{aligned}L = & -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \eta^{\alpha\beta}(\partial_\alpha\tilde{\varphi}_1\partial_\beta\tilde{\varphi}_1 + \partial_\alpha\varphi_2\partial_\beta\varphi_2) + \\ & + 2qA^\alpha(\varphi_2\partial_\alpha\tilde{\varphi}_1 - \tilde{\varphi}_1\partial_\alpha\varphi_2 - a\partial_\alpha\varphi_2) + q^2A^\alpha A_\alpha(a^2 + 2a\tilde{\varphi}_1 + \tilde{\varphi}_1^2 + \varphi_2^2) + \\ & - \frac{1}{2}\lambda[4a^2\tilde{\varphi}_1^2 + 4a\tilde{\varphi}_1^3 + 4a\tilde{\varphi}_1\varphi_2^2 + \tilde{\varphi}_1^4 + \varphi_2^4 + 2\tilde{\varphi}_1^2\varphi_2^2].\end{aligned}$$

Проанализируем квадратичное приближение. Мы видим, что вблизи вакуума векторное поле приобрело массовый член $q^2a^2A^\alpha A_\alpha$, и появилось слагаемое с перемешиванием $-2qaA^\alpha\partial_\alpha\varphi_2$. Заметим, что квадратичные слагаемые по A_α и $\partial_\alpha\varphi_2$ можно записать в виде полного квадрата:

$$\eta^{\alpha\beta}\partial_\alpha\varphi_2\partial_\beta\varphi_2 - 2qaA^\alpha\partial_\alpha\varphi_2 + q^2a^2A^\alpha A_\alpha = q^2a^2\eta^{\alpha\beta}\left(A_\alpha - \frac{1}{qa}\partial_\alpha\varphi_2\right)\left(A_\beta - \frac{1}{qa}\partial_\beta\varphi_2\right).$$

Поскольку замена компонент векторного поля

$$A_\alpha \mapsto B_\alpha := A_\alpha - \frac{1}{qa}\partial_\alpha\varphi_2$$

не изменяет кинетического члена, то это преобразование диагонализует квадратичную часть лагранжиана. В результате возникает лагранжиан с квадратичной частью, которая вообще не содержит голдстоуновское поле φ_2 :

$$-\frac{1}{4}F_{\text{в}}^{\alpha\beta}F_{\text{в}\alpha\beta} + q^2a^2B^\alpha B_\alpha + \partial\tilde{\varphi}_1^2 - 2\lambda a^2\tilde{\varphi}_1^2, \quad (25.138)$$

где

$$F_{\text{в}\alpha\beta} := \partial_\alpha B_\beta - \partial_\beta B_\alpha.$$

Теперь можно описать спектр теории. Модель описывает массивное векторное поле, взаимодействующее с нейтральным массивным скалярным полем. Таким образом,

безмассовое голдстоуновское поле исчезло, а векторное поле приобрело массу. При этом общее число степеней свободы не изменилось, т.к. массивное векторное поле, как мы увидим позже в разделе 25.4, имеет на одну степень свободы больше, чем безмассовое. То есть одна степень свободы комплексного скалярного поля перешла в продольную компоненту массивного векторного поля.

Обратим внимание, что массы полей имеют правильные знаки.

Поле φ_2 при описанном преобразовании остается в высших порядках лагранжиана, однако это не влияет на спектр теории.

Появление массивного векторного поля в физическом секторе модели можно пояснить по-другому. Параметризуем комплексное векторное поле его амплитудой и фазой:

$$\varphi := \rho e^{iq\theta}, \quad \rho = \rho(x), \quad \theta = \theta(x).$$

Поскольку исходная теория калибровочно инвариантна, то физические степени свободы остаются после фиксирования калибровки. Совершим калибровочное преобразование

$$\begin{aligned} \varphi &\mapsto \varphi' = e^{-iq\theta}, \\ A_\alpha &\mapsto A'_\alpha = A_\alpha - \partial_\alpha \theta. \end{aligned}$$

Другими словами, зафиксируем калибровку $\text{im } \varphi = 0$. Тогда лагранжиан примет вид (опустив штрихи)

$$\begin{aligned} L &= -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \eta^{\alpha\beta}(\partial_\alpha \rho + iqA_\alpha \rho)(\partial_\beta \rho - iqA_\beta \rho) - V(\rho^2) = \\ &= -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \partial\rho^2 + q^2 A^\alpha A_\alpha \rho^2 - V(\rho^2). \end{aligned}$$

Теперь разложим поля вблизи вакуумного решения

$$\rho = a + \tilde{\rho}, \quad A_\alpha = A_\alpha.$$

Тогда квадратичная часть лагранжиана примет вид

$$-\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \partial\tilde{\rho}^2 + q^2 a^2 A^\alpha A_\alpha - 2\lambda a^2 \tilde{\rho}^2.$$

Это выражение с точностью до обозначений совпадает с полученным ранее выражением (25.138). То есть мы получили прежний спектр модели. При таком подходе голдстоуновское нефизическое поле, каковым является фаза скалярного поля, вообще исключается из лагранжиана.

Описанная выше конструкция обобщается на теорию полей Янга–Миллса, взаимодействующих со скалярными полями. При этом безмассовые голдстоуновские бозоны исчезают, а калибровочные поля приобретают массу. Это явление приобретения масс первоначально безмассовыми векторными полями за счет спонтанного нарушения симметрии называется *механизмом Хиггса*. Оно было независимо открыто Энглертом, Броутом и Хиггсом [188, 189].

Как видим, спектр теории, который определяется низкоэнергетическим приближением, существенно меняется при спонтанном нарушении симметрии. Современные модели, объединяющие электромагнитные, слабые и сильные взаимодействия, основываются на калибровочной симметрии и соответствующих полях Янга–Миллса. Поэтому теория изначально содержит большое число безмассовых векторных полей.

С другой стороны, нам экспериментально известно только одно безмассовое векторное поле – электромагнитное. Чтобы устранить возникающее противоречие теории и эксперимента вводится взаимодействие со скалярными полями (хиггсовскими бозонами) таким образом, чтобы вакуумное решение уравнений движения нарушало калибровочную симметрию. В результате калибровочные поля приобретают массу за счет механизма Хиггса, и можно сказать, что мы эти поля не наблюдаем, потому что они слишком массивны.

Важно отметить, что перенормируемость квантовополевой модели определяется не низко-, а высокоэнергетическим приближением, которое у спонтанно нарушенной теории такое же, как и до нарушения симметрии. Поэтому свойство перенормируемости сохраняется при спонтанном нарушении симметрии.

25.3.5 Электромагнитное поле в аффинной геометрии

Лагранжиан электромагнитного поля на многообразии $\mathbb{M}$ с заданной аффинной геометрией при минимальной подстановке принимает вид

$$L_{\text{EM}} = -\frac{1}{4}\sqrt{|g|}g^{\alpha\beta}g^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta} := -\frac{1}{4}\sqrt{|g|}F^2, \quad (25.139)$$

где $F_{\alpha\beta}$ – напряженность электромагнитного поля (25.59). Ему соответствует действие

$$S_{\text{EM}} = \int_{\mathbb{M}} dx L_{\text{EM}}. \quad (25.140)$$

Поскольку метрика, по определению, безразмерна, то электромагнитный потенциал имеет ту же размерность (25.72), что и в пространстве Минковского.

Замечание. Как и в случае скалярных полей, лагранжиан (25.139) не зависит от аффинной связности. Это связано с тем, что $F_{\alpha\beta}$ – это компоненты локальной формы кривизны для $U(1)$ связности и являются компонентами тензора независимо от того задана или нет на многообразии $\mathbb{M}$ аффинная геометрия. С геометрической точки зрения в выражении для $F_{\alpha\beta}$ при минимальной подстановке нет никакой необходимости заменять частные производные на ковариантные, т.к. локальная форма кривизны для $U(1)$ связности уже является тензором относительно общих преобразований координат.

Если все же в определении компонент напряженности электромагнитного поля заменить частные производные на ковариантные, определяемые аффинной связностью Γ , то получим добавочное слагаемое с тензором кручения:

$$\nabla_{\alpha}A_{\beta} - \nabla_{\beta}A_{\alpha} = \partial_{\alpha}A_{\beta} - \partial_{\beta}A_{\alpha} - T_{\alpha\beta}^{\gamma}A_{\gamma}.$$

Полученное выражение ковариантно относительно общих преобразований координат, однако добавочный член явно нарушает калибровочную инвариантность. Такая замена соответствует неминимальной подстановке. $\square$

Замечание. В настоящее время считается, что спектр модели: массы, спины и заряды частиц, описываемых данным полем, определяется линейным приближением в уравнениях движения. При этом метрика, как правило, раскладывается вблизи метрики Минковского. Это означает, что электромагнитное поле в аффинной геометрии описывает безмассовые нейтральные частицы со спином единица. Такой подход вызывает много вопросов. Например, если пространство-время имеет нетривиальную

топологию, то говорить про метрику Лоренца не имеет никакого смысла. Мы не будем обсуждать эти важные вопросы, потому что лучших предложений пока не поступало. $\square$

Уравнения движения для электромагнитного поля и тензор энергии-импульса получаются варьированием соответствующего действия

$$\sqrt{|g|}S_{\text{EM},\alpha} := \frac{\delta S_{\text{EM}}}{\delta A_\alpha} = \partial_\beta(\sqrt{|g|}F^{\beta\alpha}) = \sqrt{|g|}\tilde{\nabla}_\beta F^{\beta\alpha} = 0, \quad (25.141)$$

$$\sqrt{|g|}S_{\text{EM},\alpha\beta} := \frac{\delta S_{\text{EM}}}{\delta g^{\alpha\beta}} = \frac{1}{2}\sqrt{|g|}T_{\text{EM}\alpha\beta} = -\frac{1}{2}\sqrt{|g|}F_{\alpha\gamma}F_\beta{}^\gamma + \frac{1}{8}\sqrt{|g|}g_{\alpha\beta}F^2. \quad (25.142)$$

Отсюда следует выражение для тензора энергии-импульса электромагнитного поля

$$T_{\text{EM}\alpha\beta} = -F_{\alpha\gamma}F_\beta{}^\gamma + \frac{1}{4}g_{\alpha\beta}F^2. \quad (25.143)$$

Замечание. След тензора энергии-импульса электромагнитного поля в четырехмерном пространстве-времени равен нулю, $T_{\text{EM}}{}^\alpha{}_\alpha = 0$. Здесь важна четырехмерность пространства-времени, т.к. если размерность не равна четырем, то след тензора энергии-импульса отличен от нуля. $\square$

Замечание. Вариация действия по метрике (25.142) приводит к тензору энергии-импульса электромагнитного поля (25.143), стоящему в правой части уравнений Эйнштейна (23.2). Он является ковариантным обобщением симметричного канонического тензора энергии-импульса (25.98) в пространстве Минковского. В этом отношении ситуация с электромагнитным полем такая же, как и для скалярного поля. В общем случае вариация действия полей материи по метрике совсем необязательно приводит к ковариантному обобщению канонического тензора энергии-импульса, следующему из теоремы Нетер. $\square$

Если присутствуют источники электромагнитного поля, то в правой части уравнения (25.141) появляется электрический ток:

$$\tilde{\nabla}_\beta F^{\beta\alpha} = J^\alpha, \quad (25.144)$$

где $J = \{J^\alpha\}$ – вектор тока. Происхождение электрического тока в настоящий момент не имеет значения.

Предложение 25.3.3. Если пространство-время топологически тривиально (диффеоморфно $\mathbb{R}^n$), то система уравнений второго порядка для потенциала A_α (25.144) эквивалентна системе уравнений первого порядка для компонент напряженности $F_{\alpha\beta} = -F_{\beta\alpha}$ электромагнитного поля:

$$\begin{aligned} \tilde{\nabla}_\beta F^{\beta\alpha} &= J^\alpha, \\ \partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} &= 0 \end{aligned} \quad (25.145)$$

Доказательство. Пусть $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$. Тогда из уравнения (25.144) следует система уравнений (25.145). Обратно. В силу леммы Пуанкаре 3.3.3 из второго уравнения (25.145) для односвязных многообразий следует существование потенциала A_α . Подстановка соответствующего выражения (25.59) в первое уравнение (25.145) приводит к (25.144). $\square$

Определение. Система уравнений первого порядка (25.145) для напряженности электромагнитного поля $F_{\alpha\beta} = -F_{\beta\alpha}$ называется *уравнениями Максвелла*. $\square$

Запишем действие для электромагнитного поля в еще одной форме.

Предложение 25.3.4. Уравнения движения для действия (25.140) эквивалентны уравнениям Эйлера–Лагранжа для действия

$$S = - \int_{\mathbb{M}} dx \sqrt{|g|} \left[\frac{1}{2} F^{\alpha\beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha) - \frac{1}{4} F^{\alpha\beta} F_{\alpha\beta} \right], \quad (25.146)$$

в котором компоненты $F^{\alpha\beta}$ и A_α рассматриваются в качестве независимых переменных, по которым проводится варьирование.

Доказательство. Уравнения движения для компонент напряженности являются алгебраическими:

$$\frac{\delta S}{\delta F^{\alpha\beta}} = \frac{1}{2} (\partial_\alpha A_\beta - \partial_\beta A_\alpha - F_{\alpha\beta}) = 0.$$

Их решение приводит к выражению компонент напряженности электромагнитного поля через потенциал (25.59). Согласно теореме 20.4.1 решение уравнений Эйлера–Лагранжа можно подставлять в действие. После подстановки действие (25.146) совпадает с (25.140). $\square$

Рассмотрение в качестве независимых переменных и компонент потенциала электромагнитного поля A_α , и компонент напряженности $F_{\alpha\beta}$ называется *формализмом первого порядка*.

Действие для лагранжиана (25.139) инвариантно относительно общих преобразований координат. Соответствующие вариации полей имеют вид (см. раздел 2.13)

$$\begin{aligned} \delta A_\alpha &= -\partial_\alpha \epsilon^\beta A_\beta - \epsilon^\beta \partial_\beta A_\alpha, \\ \delta g^{\alpha\beta} &= g^{\alpha\gamma} \partial_\gamma \epsilon^\beta + g^{\beta\gamma} \partial_\gamma \epsilon^\alpha - \epsilon^\gamma \partial_\gamma g^{\alpha\beta}. \end{aligned}$$

Отсюда, согласно второй теореме Нетер (20.63), следует зависимость уравнений движения

$$\partial_\beta S, {}^\beta A_\alpha + S, {}^\beta F_{\beta\alpha} - \partial_\beta S, {}^\beta_\alpha - \partial_\beta S, {}^\beta_\alpha - S, {}_{\beta\gamma} \partial_\alpha g^{\beta\gamma} = 0. \quad (25.147)$$

Или, в ковариантном виде,

$$\tilde{\nabla}_\beta S, {}^\beta A_\alpha + S, {}^\beta F_{\beta\alpha} - 2\tilde{\nabla}^\beta S, {}_{\beta\alpha} = 0. \quad (25.148)$$

Если выполнены уравнения движения для электромагнитного поля (25.141), то с учетом свернутых тождеств Бианки соотношение (25.148) принимает вид

$$\tilde{\nabla}_\beta T_{\text{EM}\alpha}{}^\beta = 0. \quad (25.149)$$

Полученное равенство, выполненное для всех решений уравнений Эйлера–Лагранжа, можно интерпретировать, как ковариантное обобщение закона сохранения тензора энергии-импульса.

Действие для электромагнитного поля инвариантно также относительно калибровочных преобразований (25.62), при которых метрика не меняется. Из калибровочной инвариантности в силу второй теоремы Нетер следует зависимость уравнений движения:

$$\tilde{\nabla}_\alpha S, {}^\alpha = 0. \quad (25.150)$$

С учетом уравнения Максвелла (25.144) это равенство приводит к закону сохранения электрического тока: $\tilde{\nabla}_\alpha J^\alpha = 0$. Если ток возникает при варьировании некоторого калибровочно инвариантного действия для полей заряженной материи, то сохранение тока будет выполняться автоматически в силу второй теоремы Нетер. Если же электрический ток вводится в уравнения Максвелла “руками”, то условие $\tilde{\nabla}_\alpha J^\alpha = 0$ необходимо для самосогласованности уравнений.

25.3.6 Электромагнитное поле в общей теории относительности

В настоящем разделе мы ограничимся четырехмерным пространством-временем $\mathbb{M}$, в котором формулы электродинамики имеют свою специфику.

В свободной электродинамике многие формулы принимают особенно простой вид, если использовать язык дифференциальных форм, описанных в главе 3. Как уже было отмечено, электромагнитный потенциал – это локальная 1-форма $\mathbb{U}(1)$ связности,

$$A := dx^\alpha A_\alpha. \quad (25.151)$$

Слово “локальная” в данном случае означает, что форма связности задана на некоторой координатной окрестности базы $\mathbb{U} \subset \mathbb{M}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{U}(1))$, а не на самом (пятимерном) пространстве расслоения $\mathbb{P}$. Локальная 2-форма кривизны – это внешняя производная от локальной формы связности,

$$F := dA = \frac{1}{2} dx^\alpha \wedge dx^\beta (\partial_\alpha A_\beta - \partial_\beta A_\alpha). \quad (25.152)$$

В дальнейшем слово “локальная” мы, для краткости, будем опускать.

Поскольку форма кривизны является точной, то ее внешняя производная равна нулю (3.33),

$$dF = \frac{1}{2} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \partial_\alpha F_{\beta\gamma} = 0. \quad (25.153)$$

Это равенство, очевидно, эквивалентно тождествам Бианки (25.61).

Приведенные выше формулы справедливы независимо от того задана ли на $\mathbb{M}$ аффинная геометрия или нет.

Если на пространстве-времени задана метрика, то можно определить 2-форму, дуальную к форме кривизны:

$$*F := \frac{1}{2} dx^\alpha \wedge dx^\beta (*F)_{\alpha\beta}, \quad (25.154)$$

где компоненты дуальной формы кривизны заданы равенством

$$*F_{\alpha\beta} := \frac{1}{2} \varepsilon_{\alpha\beta\gamma\delta} F^{\gamma\delta}. \quad (25.155)$$

В приведенной формуле $\varepsilon_{\alpha\beta\gamma\delta}(x)$ – полностью антисимметричный тензор четвертого ранга. В определении дуальной формы кривизны метрика встречается дважды. Во-первых, с ее помощью производится подъем и опускание индексов:

$$F^{\gamma\delta} := g^{\gamma\alpha} g^{\delta\beta} F_{\alpha\beta}. \quad (25.156)$$

Во-вторых, метрика входит в определение полностью антисимметричного тензора:

$$\varepsilon_{\alpha\beta\gamma\delta} := \sqrt{|g|} \hat{\varepsilon}_{\alpha\beta\gamma\delta},$$

где $\hat{\varepsilon}_{\alpha\beta\gamma\delta}$ — полностью антисимметричная тензорная плотность четвертого ранга, чьи компоненты равны по модулю единице (см. раздел 2.8).

Легко проверить, что для метрики лоренцевой сигнатуры справедлива формула $**F = -F$, что находится в согласии с общей формулой (3.53).

Пример 25.3.3. В четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$ компоненты напряженности выражаются через компоненты электрического и магнитного полей (25.76). В этом случае нетрудно получить выражения для компоненты дуальной напряженности:

$$(*F)_{\alpha\beta} = \begin{pmatrix} 0 & H_1 & H_2 & H_3 \\ -H_1 & 0 & -E_3 & E_2 \\ -H_2 & E_3 & 0 & -E_1 \\ -H_3 & -E_2 & E_1 & 0 \end{pmatrix}. \quad (25.157)$$

Это означает, что преобразование дуальности меняет местами электрическое и магнитное поля:

$$\mathbf{E} \leftrightarrow -\mathbf{H}. \quad \square$$

Используя дуальную форму кривизны, действие для свободного электромагнитного поля можно переписать в виде интеграла от 4-формы:

$$S_{\text{EM}} := -\frac{1}{4} \int dx \sqrt{|g|} F^{\alpha\beta} F_{\alpha\beta} = -\frac{1}{2} \int *F \wedge F, \quad (25.158)$$

поскольку, согласно общей формуле для внешнего произведения (3.11),

$$*F \wedge F = \frac{1}{4} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \wedge dx^\delta *F_{\alpha\beta} F_{\gamma\delta} = \frac{1}{8} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \wedge dx^\delta \varepsilon_{\alpha\beta\epsilon\zeta} F^{\epsilon\zeta} F_{\gamma\delta}.$$

Кроме этого мы воспользовались формулой (3.81) для интегрирования форм и сверткой полностью антисимметричных тензоров четвертого ранга (37.72).

Свернув уравнения Максвелла (25.145) с антисимметричным тензором, получим уравнение Максвелла в терминах дуальной формы кривизны:

$$\begin{aligned} \partial_\alpha (*F)_{\beta\gamma} + \partial_\beta (*F)_{\gamma\alpha} + \partial_\gamma (*F)_{\alpha\beta} &= J^\delta \varepsilon_{\delta\alpha\beta\gamma}, \\ \nabla_\beta (*F)^{\beta\gamma} &= 0. \end{aligned} \quad (25.159)$$

Поскольку преобразование дуальности меняет местами электрическое и магнитное поле, то второе из этих уравнений означает отсутствие в природе магнитных зарядов.

Первую пару уравнений Максвелла можно записать, используя оператор кограницы (3.56), который в рассматриваемом случае имеет вид

$$\delta := *d*.$$

Несложные вычисления позволяют переписать первую пару уравнений Максвелла (25.159) в виде равенства

$$\delta F = \frac{1}{2} J,$$

где $J := dx^\alpha J_\alpha$ — 1-форма электрического тока.

Остановимся на уравнениях электромагнитного поля без источников. Тогда первую пару уравнений Максвелла (25.159) можно записать в виде

$$d * F = 0. \quad (25.160)$$

Это уравнение можно получить путем варьирования действия (25.158) непосредственно в терминах форм. Поскольку $*F$ – это 2-форма, то из свойства (3.32) следует формула

$$d(*F \wedge A) = d(*F) \wedge A + *F \wedge dA.$$

Если вариации всех полей исчезают на границе, то уравнения движения в форме (25.160) легко получить, варьируя непосредственно действие (25.158):

$$\bar{\delta} S_{\text{EM}} = - \int *F \wedge \bar{\delta}(dA) = \int d(*F) \wedge \bar{\delta}A,$$

где мы пометили вариацию формы чертой, чтобы отличить ее от оператора кограницы.

В четырехмерном пространстве-времени можно построить еще одну 4-форму

$$F \wedge F.$$

Эта форма является точной, потому что из тождеств Бианки следует равенство

$$F \wedge dA = d(F \wedge A) - dF \wedge A = d(F \wedge A).$$

Замечание. Если метрика имеет евклидову сигнатуру, то 2-форму на четырехмерном многообразии можно разложить на самодуальную и антисамодуальную компоненты. Для метрики лоренцевой сигнатуры этого сделать нельзя. Однако комплексная 2-форма $\tilde{F} := E + iB$ самодуальна с следующим смыслом: $*\tilde{F} = i\tilde{F}$. Нетрудно проверить, что последнее уравнение инвариантно относительно действия оператора Ходжа. $\square$

Теперь изучим свойства тензора энергии-импульса в четырехмерном пространстве-времени. Из определения дуального тензора напряженности (25.154) вытекает равенство

$$*F_{\alpha\gamma} *F_{\beta}{}^{\gamma} = -\frac{1}{2}g_{\alpha\beta}F^{\gamma\delta}F_{\gamma\delta} + F_{\alpha\gamma}F_{\beta}{}^{\gamma},$$

где мы воспользовались формулами свертки антисимметричных тензоров (37.72). Теперь выражение для тензора энергии-импульса можно переписать в эквивалентном виде

$$T_{\text{EM}\alpha\beta} = -\frac{1}{2}[F_{\alpha\gamma}F_{\beta}{}^{\gamma} + *F_{\alpha\gamma} *F_{\beta}{}^{\gamma}]. \quad (25.161)$$

Из приведенного выражения следует, что тензор энергии-импульса электромагнитного поля инвариантен относительно *дуальных* $\text{SO}(2)$ вращений:

$$\begin{aligned} F_{\alpha\beta} &\mapsto F'_{\alpha\beta} = F_{\alpha\beta} \cos \omega - *F_{\alpha\beta} \sin \omega, \\ *F_{\alpha\beta} &\mapsto *F'_{\alpha\beta} = F_{\alpha\beta} \sin \omega + *F_{\alpha\beta} \cos \omega, \end{aligned} \quad (25.162)$$

где $\omega \in [0, 2\pi]$ – угол вращения. Конечно, второе из этих уравнений является следствием первого.

Пример 25.3.4. В пространстве Минковского $\mathbb{R}^{1,3}$ компоненты тензора напряженности и его дуального имеют вид (25.76), (25.157). Поэтому для временной компоненты тензора энергии-импульса (25.161) получаем хорошо известное выражение

$$T_{\text{EM}00} = \frac{1}{2}(\mathbf{E}^2 + \mathbf{H}^2). \quad (25.163)$$

Смешанная компонента контравариантного тензора энергии-импульса электромагнитного поля имеет вид

$$T_{\text{EM}}^{\mu 0} = F^{\mu\nu} F^0_{\nu}.$$

Используя выражение напряженности электромагнитного поля через электрическое и магнитное поля (25.76) нетрудно получить соотношение

$$T_{\text{EM}}^{\mu 0} = \varepsilon^{\mu\nu\rho} E_{\nu} H_{\rho},$$

где $\varepsilon^{\mu\nu\rho}$ – полностью антисимметричный тензор третьего ранга (37.66). Полученное равенство можно записать в виде векторного произведения

$$\mathbf{T} = \mathbf{E} \times \mathbf{H}. \quad (25.164)$$

Этот вектор называется *вектором Пойнтинга* и имеет смысл плотности потока энергии электромагнитного поля. При этом закон сохранения электромагнитной энергии примет вид

$$\partial_0 T_{\text{EM}}^{00} + \partial_{\mu} T_{\text{EM}}^{\mu 0} = 0. \quad \square$$

Предложение 25.3.5. Если метрика на многообразии $\mathbb{M}$ имеет лоренцеву сигнатуру и координата x^0 является временем, то временная компонента тензора энергии-импульса электромагнитного поля

$$T_{\text{EM}00} = -\frac{1}{2} [F_{0\gamma} F_0^{\gamma} + *F_{0\gamma} * F_0^{\gamma}] \quad (25.165)$$

положительно определена и, следовательно, удовлетворяет слабому энергетическому условию.

Доказательство. Ввиду антисимметрии по индексам первое слагаемое в выражении (25.165) имеет вид

$$-\frac{1}{2} F_{0\mu} F_{0\nu} g^{\mu\nu}.$$

При сделанных предположениях матрица $g^{\mu\nu}$ согласно предложению 24.2.1 отрицательно определена. То же верно и для второго слагаемого. $\square$

Следствие. Временная компонента тензора энергии-импульса электромагнитного поля удовлетворяет сильному энергетическому условию (23.139).

Доказательство. Поскольку след тензора энергии-импульса равен нулю, то понятие слабого и сильного энергетического условия эквивалентны. $\square$

Предложение 25.3.6. Действие электромагнитного поля в аффинной геометрии (25.140) инвариантно относительно преобразований Вейля:

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = e^{2\phi} g_{\alpha\beta}, \quad A_{\alpha} \mapsto \bar{A}_{\alpha} = A_{\alpha}, \quad (25.166)$$

где $\phi(x) \in \mathcal{C}^2(\mathbb{M})$ – произвольная вещественнозначная функция.

Доказательство. Утверждение следует из равенства

$$\sqrt{|g|} g^{\alpha\beta} g^{\gamma\delta} = \sqrt{|\bar{g}|} \bar{g}^{\alpha\beta} \bar{g}^{\gamma\delta}. \quad \square$$

Замечание. Вейлевская инвариантность действия электромагнитного поля является спецификой четырехмерности пространства-времени. Если размерность пространства-времени отлична от четырех, то след тензора энергии-импульса электромагнитного поля отличен от нуля. $\square$

Равенство нулю следа тензора энергии-импульса электромагнитного поля связано с наличием вейлевской калибровочной инвариантности. Этот факт является общим.

Предложение 25.3.7. Пусть действие зависит от метрики $g_{\alpha\beta}$ и некоторого набора полей материи φ^A

$$S_M = \int dx L_M(g, \varphi).$$

Если действие инвариантно относительно преобразований Вейля, не затрагивающих поля материи,

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = e^{2\phi} g_{\alpha\beta}, \quad \varphi^A \mapsto \bar{\varphi}^A = \varphi^A,$$

то след тензора энергии-импульса полей материи равен нулю, $T_{M\alpha}{}^\alpha = 0$.

Доказательство. Поскольку параметр преобразования Вейля зависит от точки пространства-времени, то справедлива вторая теорема Нетер. В рассматриваемом случае это означает следующее. Инвариантность действия имеет вид

$$\delta S_M = \int dx \frac{\delta S_M}{\delta g^{\alpha\beta}} \delta g^{\alpha\beta} = 0.$$

Учтем определение тензора энергии-импульса в общей теории относительности (23.11) и выражение для бесконечно малых преобразований Вейля

$$\delta g^{\alpha\beta} = -2\phi g^{\alpha\beta}.$$

Тогда равенство нулю вариации действия полей материи равносильно равенству $g^{\alpha\beta} T_{M\alpha\beta} = 0$. $\square$

Пример 25.3.5. Бозонная струна описывается набором скалярных безмассовых полей $X^a(x)$, $a = 0, 1, \dots, D-1$, заданных на двумерном пространстве-времени (поверхности) $\mathbb{M}$ и принимающих значения в D -мерном пространстве Минковского $\mathbb{R}^{1,D-1}$:

$$X : \quad \mathbb{M} \ni \quad x = \{x^\alpha\} \mapsto X(x) = \{X^a(x)\} \quad \in \mathbb{R}^{1,D-1}.$$

Действие для бозонной струны можно записать в виде

$$S = -\frac{1}{2} \int_{\mathbb{M}} dx \sqrt{|g|} g^{\alpha\beta} \partial_\alpha X^a \partial_\beta X^b \eta_{ab}.$$

Знак этого действия выбран таким образом, чтобы пространственные компоненты X^i , $i = 1, \dots, D-1$, описывались обычным действием (25.13) и давали положительный вклад в гамильтониан системы. Это действие инвариантно относительно преобразований Вейля:

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = e^{2\phi} g_{\alpha\beta}, \quad X^a \mapsto X^a.$$

Тензор энергии-импульса бозонной струны имеет вид

$$T_{\alpha\beta} := \frac{2}{\sqrt{|g|}} \frac{\delta S}{\delta g^{\alpha\beta}} = -\partial_\alpha X^a \partial_\beta X_a + \frac{1}{2} g_{\alpha\beta} g^{\gamma\delta} \partial_\gamma X^a \partial_\delta X_a,$$

где $X_a := X^b \eta_{ba}$. Его след, очевидно, равен нулю. Вейлевская инвариантность действия для безмассового скалярного поля является спецификой двумерия. В пространстве-времени бóльшего числа измерений она отсутствует. $\square$

25.4 Поле Прока́

Поле Прока́ $A = \{A_\alpha\}$, $\alpha = 0, 1, \dots, n-1$, описывает массивные векторные частицы со спином единица [190, 191]. В пространстве Минковского $\mathbb{R}^{1,n-1}$ ему соответствует лагранжиан

$$L = -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \frac{1}{2}m^2 A^\alpha A_\alpha, \quad (25.167)$$

где, как и для электромагнитного поля,

$$F_{\alpha\beta} := \partial_\alpha A_\beta - \partial_\beta A_\alpha$$

и $m > 0$ – масса поля. Он отличается от лагранжиана электромагнитного поля только массовым членом. Знак перед массовым членом выбран так, чтобы вклад пространственных компонент A_μ , $\mu = 1, \dots, n-1$, в плотность энергии был положителен.

Уравнения движения для лагранжиана (25.167) имеют вид

$$S,^\alpha := \frac{\delta S}{\delta A_\alpha} = \partial_\beta F^{\beta\alpha} + m^2 A^\alpha = 0. \quad (25.168)$$

Ввиду антисимметрии относительно перестановки индексов, $F^{\alpha\beta} = -F^{\beta\alpha}$, дивергенция этой системы уравнений приводит к дифференциальному уравнению первого порядка на компоненты поля:

$$\partial_\alpha A^\alpha = 0, \quad (25.169)$$

которое должно быть выполнено для любого решения системы уравнений (25.168). Полученное условие имеет вид калибровки Лоренца (25.96) для электромагнитного поля. С учетом условия (25.169) уравнения движения (25.168) принимают вид

$$\square A^\alpha + m^2 A^\alpha = 0, \quad \square := \partial^\alpha \partial_\alpha. \quad (25.170)$$

Таким образом, для каждой компоненты векторного поля возникает уравнение Клейна–Гордона–Фока. Однако поле Прока́ описывает не n распространяющихся степеней свободы, а только $n-1$, поскольку на компоненты поля наложено дополнительное условие (25.169). Чтобы убедиться в этом проведем более детальный анализ уравнений движения.

Запишем систему уравнений (25.168) отдельно для временной и пространственных компонент:

$$\partial_\mu F^{\mu 0} + m^2 A^0 = 0, \quad (25.171)$$

$$\partial_0 F^{0\mu} + \partial_\nu F^{\nu\mu} + m^2 A^\mu = 0. \quad (25.172)$$

Предположим, что все поля убывают на бесконечности. Тогда согласно теореме 25.3.1 оператор Лапласа обратим и пространственные компоненты поля Прока́, как и в случае электромагнитного поля (25.81), можно взаимно однозначно разложить на поперечную и продольную составляющие:

$$A_\mu = A_\mu^T + \partial_\mu \varphi. \quad (25.173)$$

Тогда лагранжевы уравнения движения (25.171), (25.172) эквивалентны следующей системе уравнений:

$$-\Delta A_0 + \Delta \dot{\varphi} + m^2 A_0 = 0, \quad (25.174)$$

$$\partial_\mu \ddot{\varphi} - \partial_\mu \dot{A}_0 + m^2 \partial_\mu \varphi = 0, \quad (25.175)$$

$$\square A_\mu^T + m^2 A_\mu^T = 0, \quad (25.176)$$

где мы расщепили уравнение (25.172) на продольную и поперечную части, а также ввели обозначения для производной по времени и операторов Лапласа и Даламбера:

$$\dot{\varphi} := \partial_0 \varphi, \quad \Delta := -\partial^\mu \partial_\mu, \quad \square := \partial^\alpha \partial_\alpha = \partial_0^2 - \Delta.$$

Уравнения для поперечных компонент (25.176) – это уравнения Клейна–Гордона–Фока, и они полностью отщепились. Рассмотрим уравнение (25.174) как уравнение на временную компоненту:

$$(\Delta - m^2)A_0 = \Delta \dot{\varphi}. \quad (25.177)$$

Это – уравнение Гельмгольца, решение которого существует и единственно для достаточно широкого класса граничных условий, в частности, для полей, убывающих на бесконечности [192]. Запишем его решение в виде

$$A_0 = \frac{1}{\Delta - m^2} \Delta \dot{\varphi}. \quad (25.178)$$

Для полей, убывающих на бесконечности, уравнение (25.175) можно один раз проинтегрировать:

$$\ddot{\varphi} - \dot{A}_0 + m^2 \varphi = 0.$$

Подставив сюда решение для A_0 (25.178) и умножив на оператор $\Delta - m^2$, получим уравнение Клейна–Гордона–Фока на функцию φ ,

$$\square \varphi + m^2 \varphi = 0. \quad (25.179)$$

Поскольку все сделанные операции обратимы, то мы доказали

Предложение 25.4.1. *Для полей, убывающих на бесконечности, лагранжесвы уравнения движения (25.168) для поля Прокá эквивалентны $n - 1$ уравнениям Клейна–Гордона–Фока (25.176), (25.179), дополненным уравнениями для продольной составляющей (25.84) и выражением для временной компоненты (25.178).*

Для тех, у кого остались сомнения по поводу следствия уравнений движения (25.169), скажем, что подстановка решения (25.178) в это условие, что нетрудно проверить, также приводит к уравнению Клейна–Гордона–Фока (25.179).

Таким образом, чтобы поставить задачу Коши для поля Прокá, необходимо и достаточно поставить задачу Коши для $n - 2$ поперечных компонент A_μ^T и поля φ , определяющего продольную составляющую. Поэтому поле Прокá действительно описывает $n - 1$ физическую (распространяющуюся) степень свободы.

25.4.1 Гамильтонова формулировка

Векторному полю A_α соответствуют сопряженные импульсы

$$P^\alpha := \frac{\partial L}{\partial(\partial_0 A_\alpha)} = F^{\alpha 0}. \quad (25.180)$$

Скобки Пуассона между координатами и импульсами имеют тот же вид, что и для электромагнитного поля (25.108). Как и для электромагнитного поля, из выражения для импульсов (25.180) следует первичная связь

$$G_1 = P^0 \approx 0. \quad (25.181)$$

Других первичных связей рассматриваемая модель не содержит. Несложные вычисления приводят к следующему выражению для гамильтониана

$$\mathcal{H} = \int d\mathbf{x} \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - A_0 \partial_\mu P^\mu - \frac{1}{2} m^2 A^\alpha A_\alpha + \lambda P^0 \right), \quad (25.182)$$

где λ – множитель Лагранжа.

Гамильтоновым уравнениям движения имеют следующий вид:

$$\dot{A}_0 = \lambda, \quad (25.183)$$

$$\dot{A}_\mu = -P_\mu + \partial_\mu A_0, \quad (25.184)$$

$$\dot{P}^0 = \partial_\mu P^\mu + m^2 A_0, \quad (25.185)$$

$$\dot{P}^\mu = \partial_\nu F^{\nu\mu} + m^2 A^\mu. \quad (25.186)$$

Кроме этого, вариация действия по множителю Лагранжа приводит к связи (25.181). Таким образом, полная система гамильтоновых уравнений движения состоит из уравнений (25.183)–(25.186) и связи (25.181). На первый взгляд решения гамильтоновых уравнений содержат функциональный произвол, т.к. множитель Лагранжа λ произволен. Однако этот вывод неверен. Действительно, поскольку связь $P^0 = 0$ для любого решения гамильтоновых уравнений должна выполняться в любой момент времени, то производная по времени от уравнения (25.185) должна быть равна нулю. Это приводит к уравнению

$$m^2 \dot{A}_0 + \partial_\mu \dot{P}^\mu = m^2 (\dot{A}_0 + \partial_\mu A^\mu) = 0, \quad (25.187)$$

где мы воспользовались уравнением (25.186). Сравнивая полученное равенство с уравнением (25.183), заключаем, что для любого решения гамильтоновых уравнений множитель Лагранжа имеет вид

$$\lambda = -\partial_\mu A^\mu. \quad (25.188)$$

Теперь можно доказать эквивалентность лагранжевой и гамильтоновой формулировок.

Предложение 25.4.2. *Полная система гамильтоновых уравнений (25.183)–(25.186) и (25.181) эквивалентна лагранжевым уравнениям движения (25.168).*

Доказательство. Пусть выполнены гамильтоновы уравнения (25.183)–(25.186) и связь (25.181). Из уравнения (25.184) находим импульсы

$$P_\mu = \partial_\mu A_0 - \dot{A}_\mu = F_{\mu 0}.$$

Подстановка этого выражения в формулу (25.186) приводит к уравнению (25.172). Уравнение (25.185) с учетом связи (25.181) дает уравнение (25.171).

Обратно. Пусть выполнены лагранжевы уравнения движения (25.171), (25.172). Введем обозначение $P^\alpha := F^{\alpha 0}$. Отсюда немедленно следует связь (25.181). А для пространственных значений индекса определение $P^\mu = F^{\mu 0}$ эквивалентно уравнению (25.184). Используя новую переменную P^μ , лагранжевы уравнения (25.172) переписываются в виде гамильтонова уравнения (25.186). Соотношение (25.169) является следствием лагранжевых уравнений движения, и оно дает уравнение (25.183) при выполнении условия (25.188). Наконец, лагранжево уравнение (25.171) сводится к уравнению (25.185) при выполнении связи (25.181). $\square$

Замечание. Отметим, что для эквивалентности лагранжевой и гамильтоновой формулировок к гамильтониану необходимо добавить первичную связь (25.181). В противном случае никакой эквивалентности нет. Действительно, равенство $\lambda = 0$ явно противоречит условию (25.188). $\square$

Продолжим построение гамильтонова формализма, согласно общей схеме, рассмотренной в разделе 22.

Производная по времени от первичной связи (25.181) равна

$$\dot{G}_1 = [G_1, \mathcal{H}] = \partial_\mu P^\mu + m^2 A^0.$$

Поскольку первичная связь должна сохраняться во времени, то возникает вторичная связь

$$G_2 = \partial_\mu P^\mu + m^2 A^0 \approx 0. \quad (25.189)$$

Нетрудно вычислить производную по времени от вторичной связи:

$$\dot{G}_2 = [G_2, \mathcal{H}] = m^2(\partial_\mu A^\mu + \lambda).$$

Сохранение вторичной связи во времени определяет множитель Лагранжа:

$$\lambda = -\partial_\mu A^\mu. \quad (25.190)$$

Поэтому в теории не возникает других связей.

Равенство (25.190) уже было получено ранее (25.188) при анализе гамильтоновых уравнений движения.

Скобка связей между собой имеет вид

$$[G_1, G'_2] = -m^2 \delta(\mathbf{x} - \mathbf{x}').$$

Введем обозначение для полного набора связей:

$$\{\Phi^M\} = \{G_1, G_2\}, \quad M = 1, 2.$$

Тогда матрица скобок Пуассона для связей примет вид

$$[\Phi^M, \Phi'^N] = \begin{pmatrix} 0 & -m^2 \\ m^2 & 0 \end{pmatrix} \delta(\mathbf{x} - \mathbf{x}').$$

Она, очевидно, невырождена, и, следовательно, связи G_1 и G_2 являются связями второго рода.

Полный гамильтониан получается после добавления всех связей:

$$\mathcal{H}_T = \int d\mathbf{x} \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - A_0 \partial_\mu P^\mu - \frac{1}{2} m^2 A^\alpha A_\alpha + \lambda P^0 + \mu (\partial_\mu P^\mu + m^2 A^0) \right). \quad (25.191)$$

Поскольку поле Прокá содержит только связи второго рода, то полный гамильтониан совпадает с расширенным, $\mathcal{H}_T = \mathcal{H}_E$.

Один из множителей Лагранжа уже найден (25.190). Второй множитель Лагранжа определяется из условия сохранения первичной связи во времени,

$$\dot{G}_1 = [G_1, \mathcal{H}_T] = \partial_\mu P^\mu + m^2 A^0 - m^2 \mu = 0.$$

Отсюда вытекает равенство

$$\mu = A^0 + \frac{1}{m^2} \partial_\mu P^\mu. \quad (25.192)$$

Предложение 25.4.3. Гамильтоновы уравнения движения для полного гамильтониана (25.191) эквивалентны лагранжевым уравнениям (25.168).

Доказательство. Добавление к исходному гамильтониану (25.182) вторичной связи приводит к возникновению самой связи из вариационного принципа. Кроме того меняются два гамильтоновых уравнения:

$$\begin{aligned}\dot{A}_\mu &= -P_\mu + \partial_\mu A_0 - \partial_\mu \mu, \\ \dot{P}^0 &= \partial_\mu P^\mu + m^2 A_0 - \mu m^2.\end{aligned}$$

Если выполнена вторичная связь, то из соотношения (25.192) следует равенство $\mu = 0$. Поэтому гамильтоновы уравнения для полного гамильтониана эквивалентны исходным гамильтоновым уравнениям и, в силу предложения 25.4.2, лагранжевым уравнениям (25.168). $\square$

Согласно общей схеме, описанной в разделе 22.2, решение связей второго рода можно подставлять в действие. В рассматриваемом случае связи легко решаются:

$$A_0 = -\frac{1}{m^2} \partial_\mu P^\mu, \quad P_0 = 0.$$

После подстановки этого решения в формулу (25.191) получим гамильтониан для физических степеней свободы

$$\mathcal{H}_{\text{ph}} = \int d\mathbf{x} \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{2m^2} (\partial_\mu P^\mu)^2 + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{2} m^2 A^\mu A_\mu \right). \quad (25.193)$$

Таким образом, временная компонента поля Прокá является нефизической и может быть полностью исключена из теории путем решения пары связей второго рода. Пространственные компоненты A_μ и сопряженные импульсы P^μ являются физическими (т.е. распространяющимися). Их динамика полностью определяется гамильтонианом для физических степеней свободы (25.193). Мы видим, что поле Прокá описывает $n - 1$ степень свободы, что на единицу превышает количество степеней свободы электромагнитного поля.

Гамильтонова плотность для физических степеней свободы поля Прокá (25.193), очевидно, положительно определена. В то же время гамильтоновы плотности исходного (25.182) и полного (25.191) гамильтонианов не являются положительно определенными, т.к. содержат множители Лагранжа.

Гамильтоновы уравнения движения для физических степеней свободы, следующие из гамильтониана (25.193) имеют вид

$$\dot{A}_\mu = -P_\mu - \frac{1}{m^2} \partial_\mu \partial_\nu P^\nu, \quad (25.194)$$

$$\dot{P}^\mu = \partial_\nu F^{\nu\mu} + m^2 A^\mu. \quad (25.195)$$

Здесь проявляется важная особенность теории поля: в отличие от гамильтонова описания динамики точечных частиц, уравнение (25.194) для импульсов является не алгебраическим, а дифференциальным. Поэтому выразить импульсы через скорости, чтобы вернуться к лагранжеву описанию, не удастся.

Чтобы сравнить гамильтонову и лагранжеву динамику физических степеней свободы, разложим координаты и импульсы на поперечную и продольную составляющие. Разложение пространственных компонент векторного поля уже было введено формулой (25.173). Для импульсов мы пишем

$$P_\mu = P_\mu^\perp + \partial_\mu \chi.$$

Тогда гамильтоновы уравнения (25.194), (25.195) также раскладываются на поперечные и продольные части:

$$\dot{A}_\mu^\tau = -P_\mu^\tau, \quad (25.196)$$

$$\partial_\mu \dot{\varphi} = -\partial_\mu \chi + \frac{\Delta}{m^2} \partial_\mu \chi, \quad (25.197)$$

$$\dot{P}_\mu^\tau = -\Delta A_\mu^\tau + m^2 A_\mu^\tau, \quad (25.198)$$

$$\partial_\mu \dot{\chi} = m^2 \partial_\mu \varphi. \quad (25.199)$$

Гамильтоновы уравнения для поперечных компонент (25.196), (25.198) отщепились и, как легко видеть, эквивалентны уравнениям Клейна–Гордона–Фока (25.176).

Для полей, убывающих на бесконечности, уравнения (25.197) и (25.199) можно один раз проинтегрировать:

$$\begin{aligned} \dot{\varphi} &= -\chi + \frac{\Delta}{m^2} \chi, \\ \dot{\chi} &= m^2 \varphi. \end{aligned}$$

Первое уравнение можно решить относительно χ ,

$$\chi = \frac{m^2}{\Delta - m^2} \dot{\varphi},$$

и подставить данное решение во второе. В результате получим уравнение Клейна–Гордона–Фока (25.179). Поскольку все сделанные операции были обратимы, то мы доказали

Предложение 25.4.4. *Для полей, убывающих на бесконечности, гамильтоновы уравнения движения для физических степеней свободы (25.194), (25.195) эквивалентны лагранжевым уравнениям для физических степеней свободы (25.176), (25.179).*

Сейчас проявилось очень важное преимущество гамильтонова подхода, как более гибкого. А именно, выделение физических степеней свободы на гамильтоновом языке было проведено явно, и получено локальное выражение для соответствующего гамильтониана. В то же время, выделение физических степеней свободы в лагранжевом формализме нелокально, т.к. для нахождения поля φ необходимо решить дифференциальное уравнение (25.84).

Векторное поле

При рассмотрении поля Прока́ был выбран необычный кинетический член. Это было сделано неспроста. В настоящем подразделе мы продемонстрируем трудность, которая возникает при выборе лагранжиана для векторного поля.

Пусть лагранжиан имеет вид

$$L = \frac{1}{2} \partial^\alpha A^\beta \partial_\alpha A_\beta + \frac{1}{2} m^2 A^\alpha A_\alpha, \quad \partial^\alpha := \eta^{\alpha\beta} \partial_\beta. \quad (25.200)$$

Этот лагранжиан, очевидно, инвариантен относительно глобальных преобразований

Лоренца. Обобщенные импульсы и гамильтониан имеют вид

$$\begin{aligned}
 P^\alpha &= \dot{A}^\alpha, \\
 H &= -\frac{1}{2}(P^0)^2 + \frac{1}{2} \sum_{\mu=1}^{n-1} (P^\mu)^2 - \frac{1}{2} \sum_{\mu=1}^{n-1} (\partial_\mu A_0)^2 + \frac{1}{2} \sum_{\mu,\nu=1}^{n-1} (\partial_\mu A_\nu)^2 - \\
 &\quad - \frac{1}{2} m^2 (A_0)^2 + \frac{1}{2} m^2 \sum_{\mu=1}^{n-1} (A_\mu)^2.
 \end{aligned} \tag{25.201}$$

Мы видим, что в теории отсутствуют какие либо связи, и все компоненты векторного поля являются распространяющимися.

Проблема заключается в том, что временная компонента векторного поля дает отрицательный вклад в гамильтониан, и поэтому модель не приемлема с физической точки зрения. А именно, кинетический член в гамильтониане $-(P^0)^2/2$ отрицателен (гост). Отрицателен также массовый член $-m^2(A_0)^2/2$ (тахин). С точки зрения группы Пуанкаре массивное векторное поле имеет две неприводимые компоненты: спин 0 (компонента A_0) и спин 1 (пространственные компоненты A_μ). Следовательно, лагранжиан (25.200) в квантовой теории поля описывает частицы двух видов. При этом частицы со спином 0 дают отрицательный вклад в канонический гамильтониан, что не приемлемо. Именно по этой причине для поля Прокá был выбран кинетический член того же вида, что и в электродинамике (25.167). В результате в теории появились связи второго рода, которые позволили исключить спин 0 как нефизическое поле.

Аналогичная проблема возникает и для тензорных полей более высокого ранга: компоненты с нечетным и четным числом нулевых индексов входят в гамильтониан с разными знаками, что связано с лоренцевой сигнатурой метрики. Чтобы исключить компоненты одного знака как нефизические, необходимо выбрать соответствующий кинетический член, если это возможно.

25.5 Поля Янга–Миллса

Поля Янга–Миллса (калибровочные поля) играют важную роль в математической физике, потому что лежат в основе современных моделей, объединяющих электромагнитные, слабые и сильные взаимодействия.

С геометрической точки зрения у нас есть следующая конструкция. Во-первых, мы предполагаем, что пространство-время представляет собой многообразие $\mathbb{M}$. Для общности будем рассматривать пространство-время произвольной размерности n . Затем строится главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с некоторой структурной группой Ли $\mathbb{G}$ размерности $\dim \mathbb{G} = N$, о которой мы поговорим несколько позже, и проекцией π . Мы предполагаем, что на главном расслоении задана связность, т.е. инвариантное распределение горизонтальных подпространств (см. раздел 13.1). Каждая связность взаимно однозначно определяется формой связности, заданной на пространстве главного расслоения $\mathbb{P}$. Если задано координатное покрытие базы $\mathbb{M} = \cup_i \mathbb{U}_i$, и семейство локальных сечений $\sigma_i : \mathbb{U}_i \rightarrow \mathbb{P}$, то форма связности взаимно однозначно определяет семейство локальных форм связности, заданных на каждой координатной окрестности $\mathbb{U}_i$, которые определяются формой связности с помощью возвратов отображений σ_i . Локальная форма связности – это 1-форма на координатной окрестности $\mathbb{U}_i$ со

значениями в алгебре Ли $\mathfrak{g}$ группы Ли $\mathbb{G}$:

$$\omega = dx^\alpha A_\alpha^A L_A, \quad (25.202)$$

где L_A , $A = 1, \dots, N$, – базис алгебры Ли (см. раздел 8.1). Компоненты локальной формы связности $A_\alpha^A(x)$ это и есть поля Янга–Миллса (после добавления соответствующих уравнений движения, что будет сделано позже). Базис алгебры Ли удовлетворяет некоторым коммутационным соотношениям

$$[L_A, L_B] = f_{AB}^C L_C,$$

где f_{AB}^C – структурные константы группы Ли $\mathbb{G}$.

При изменении локального сечения $\sigma \mapsto \sigma'$ (или при вертикальном автоморфизме) компоненты локальной формы связности меняются $A_\alpha^A \mapsto A_\alpha'^A$. Правило преобразования компонент связности известно (13.14), однако оно явно содержит функцию композиции для группы Ли $\mathbb{G}$ и поэтому неудобно. Инфинитезимальные калибровочные преобразования с параметрами $\epsilon^A(x)$ имеют вид

$$A_\alpha'^A = A_\alpha^A + \nabla_\alpha \epsilon^A, \quad (25.203)$$

где

$$\nabla_\alpha \epsilon^A := \partial_\alpha \epsilon^A - A_\alpha^B \epsilon^C f_{BC}^A.$$

Чтобы записать конечные калибровочные преобразования в удобном виде, рассмотрим произвольное точное представление группы Ли $\mathbb{G}$. Для определенности рассмотрим присоединенное представление. В этом представлении каждому элементу базиса алгебры Ли $\mathfrak{g}$ соответствует матрица: $L_A \mapsto f_{AB}^C$, где первый индекс A нумерует базисные векторы, а второй и третий рассматриваются, как матричные. Тогда каждой локальной форме связности можно поставить в соответствие матричнозначную 1-форму с компонентами

$$A_{AB}^C := -A_\alpha^A f_{AB}^C. \quad (25.204)$$

(Знак минус не имеет принципиального значения.) Тогда поле Янга–Миллса при изменении локального сечения преобразуется по правилу (13.16)

$$A_\alpha \mapsto A'_\alpha = S A_\alpha S^{-1} + \partial_\alpha S S^{-1}, \quad (25.205)$$

где $S(x) = \{S_A^B(x)\}$ – матрица присоединенного представления, и мы, для краткости, опустили матричные индексы. Это – хорошо известное калибровочное преобразование полей Янга–Миллса.

Каждой локальной форме связности соответствует локальная форма кривизны (см. раздел 13.2). Она представляет собой 2-форму на координатной окрестности $\mathbb{U}_i$ со значениями в алгебре Ли:

$$R = dx^\alpha \wedge dx^\beta F_{\alpha\beta}^A L_A, \quad (25.206)$$

где компоненты имеют вид (13.35)

$$F_{\alpha\beta}^A = \partial_\alpha A_\beta^A - \partial_\beta A_\alpha^A - A_\alpha^B A_\beta^C f_{BC}^A. \quad (25.207)$$

Это есть напряженность поля Янга–Миллса. Ранее было показано, что при калибровочном преобразовании (25.205) компоненты напряженности преобразуются ковариантным образом (13.36)

$$F_{\alpha\beta}^A \mapsto F_{\alpha\beta}'^A = F_{\alpha\beta}^B S^{-1}_B{}^A.$$

Если ввести матричнозначную форму кривизны

$$F_{\alpha\beta}{}^C := -F_{\alpha\beta}{}^A f_{AB}{}^C, \quad (25.208)$$

то ее закон преобразования будет выглядеть немного иначе:

$$F_{\alpha\beta} \mapsto F'_{\alpha\beta} = S F_{\alpha\beta} S^{-1}, \quad (25.209)$$

где мы, для краткости, опустили матричные индексы.

Матричнозначные компоненты напряженности можно выразить через матричнозначные поля Янга–Миллса:

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha - [A_\alpha, A_\beta], \quad (25.210)$$

где $[A_\alpha, A_\beta]$ – коммутатор матриц, и мы опустили матричные индексы. Чтобы получить данное выражение для компонент напряженности из формулы (25.207), использованы тождества Якоби для структурных констант (8.11).

Согласно теореме 13.2.3 каждая форма кривизны удовлетворяет тождествам Бианки. Тождества Бианки для локальной формы кривизны были получены ранее (13.2.6):

$$\nabla_\alpha F_{\beta\gamma}{}^A + \nabla_\beta F_{\gamma\alpha}{}^A + \nabla_\gamma F_{\alpha\beta}{}^A = 0, \quad (25.211)$$

где

$$\nabla_\alpha F_{\beta\gamma}{}^A = \partial_\alpha F_{\beta\gamma}{}^A - A_\alpha{}^B F_{\beta\gamma}{}^C f_{BC}{}^A \quad (25.212)$$

– ковариантная производная (по отношению к калибровочным преобразованиям) от компонент напряженности поля Янга–Миллса.

Предложение 25.5.1. *Для существования такого калибровочного преобразования, чтобы преобразованное поле Янга–Миллса обратилось бы в нуль в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ необходимо и достаточно, чтобы локальная форма кривизны связности была равна нулю в $\mathbb{U}$.*

Доказательство. Если локальная форма связности равна нулю, то локальная форма кривизны (25.207) также обращается в нуль.

Обратно. Рассмотрим правило преобразования полей Янга–Миллса (25.205) как дифференциальное уравнение на матрицу S . Если преобразованное поле Янга–Миллса равно нулю, $A'_\alpha = 0$, то матрица преобразования должна удовлетворять системе дифференциальных уравнений:

$$\partial_\alpha S = -S A_\alpha. \quad (25.213)$$

Эта система локально разрешима тогда и только тогда, когда выполнены условия разрешимости:

$$\partial_{[\beta} \partial_{\alpha]} S = 0,$$

где квадратные скобки обозначают антисимметризацию по индексам. Вторая частная производная от матрицы преобразования имеет вид

$$\partial_\beta \partial_\alpha S = -\partial_\beta S A_\alpha - S \partial_\beta A_\alpha = S(A_\beta A_\alpha - \partial_\beta A_\alpha),$$

где в первом слагаемом мы воспользовались исходным уравнением (25.213). Антисимметризация полученного выражения по индексам α и β приводит к равенству

$$F_{\alpha\beta} = 0.$$

Таким образом, равенство нулю локальной формы кривизны является необходимым и достаточным условием разрешимости системы уравнений (25.213) в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$. $\square$

Определение. Из доказанного утверждения следует, что если локальная форма кривизны равна нулю, то в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ компоненты соответствующей локальной формы связности имеют вид

$$A_\alpha = -S^{-1}\partial_\alpha S = \partial_\alpha S^{-1}S. \quad (25.214)$$

Такое поле Янга–Миллса называется *чистой калибровкой*. $\square$

Легко проверить, что напряженность поля Янга–Миллса для поля (25.214) действительно обращается в нуль.

Описанная выше геометрическая конструкция не зависит от того задана ли на базе $\mathbb{M}$ какая-либо геометрия, т.е. метрика и аффинная связность, или нет. Главное расслоение, связность и кривизна определяются в дифференциальной геометрии самостоятельно. Закон преобразования компонент локальных форм связности и кривизны при изменении сечения следует из определения. Для этого не надо вводить какие-либо дополнительные поля и ковариантные производные. Несмотря на отсутствие аффинной связности в частных производных при определении локальной формы кривизны (25.207) и в тождествах Бианки (25.212), все выписанные соотношения ковариантны относительно общих преобразований координат.

Следующий шаг построения моделей состоит в построении действия. Мы считаем, что действие должно быть инвариантно относительно общих преобразований координат и калибровочных преобразований (25.205). Чтобы построить соответствующий инвариант нам понадобятся две метрики: метрика $g_{\alpha\beta}$ на пространственно-временном многообразии и двусторонне инвариантная метрика η_{AB} на групповом многообразии. Соответствующее действие имеет вид

$$S_{YM} = \int_{\mathbb{M}} dx L_{YM}, \quad L_{YM} := -\frac{1}{4} \sqrt{|g|} g^{\alpha\gamma} g^{\beta\delta} F_{\alpha\beta}{}^A F_{\gamma\delta}{}^B \eta_{AB}. \quad (25.215)$$

В квантовой теории поля в качестве пространства-времени выбирается пространство Минковского $\mathbb{R}^{1,n-1}$, группа общих преобразования координат сужается до группы Пуанкаре $\mathbb{O}(1, n-1)$, и в качестве пространственно-временной метрики выбирается метрика Лоренца, $g_{\alpha\beta} \mapsto \eta_{\alpha\beta}$.

Если калибровочная группа $\mathbb{G}$ является абелевой, как, например, в электродинамике, то в качестве двусторонне инвариантной метрики подойдет любая постоянная матрица. Действительно, при калибровочных преобразованиях (25.205) компоненты напряженности поля Янга–Миллса для абелевой группы вообще не преобразуются.

Для неабелевых групп в качестве двусторонне инвариантной метрики выберем форму Киллинга–Картана (см. раздел 8.47). Согласно теореме 8.5.1 форма Киллинга–Картана невырождена тогда и только тогда, когда группа Ли $\mathbb{G}$ полупроста. Поэтому мы предполагаем, что структурная группа Ли $\mathbb{G}$ полупроста.

Согласно теореме 8.13.6 любая полупростая группа Ли единственным образом представляется в виде прямого произведения простых групп. В этом случае для каждого сомножителя можно написать отдельный инвариант вида (25.215) со своей константой связи. Поэтому в дальнейшем, для простоты, мы ограничимся рассмотрением простых групп Ли. Если мы научимся работать с простыми группами Ли, то построить модель, соответствующую их прямому произведению, не составит труда.

В дальнейшем мы покажем, что канонический гамильтониан для физических степеней свободы, которые описываются действием (25.215), положительно определен тогда и только тогда, когда метрика η_{AB} положительно определена. Согласно теореме 8.5.2 связная полупростая группа Ли компактна тогда и только тогда, когда ее

форма Киллинга–Картана

$$\eta_{AB} := -f_{AC}{}^D f_{BD}{}^C \quad (25.216)$$

положительно определена. Поэтому мы предполагаем, что структурная группа Ли $\mathbb{G}$ связна, проста и компактна.

Базис алгебры Ли L_A определен с точностью до невырожденных линейных преобразований

$$L_A \mapsto S_A{}^B L_B, \quad S \in \mathbb{GL}(N, \mathbb{R}).$$

При этом форма Киллинга–Картана преобразуется по тензорному закону

$$\eta_{AB} \mapsto S_A{}^C S_B{}^D \eta_{CD}.$$

Согласно теореме 37.1.8 произвольную симметричную матрицу можно привести к диагональному виду с помощью ортогональной матрицы. Дальнейшей растяжкой координат можно добиться того, что на диагонали формы Киллинга–Картана будут стоять единицы. Таким образом, не теряя общности, можно считать, что для связной компактной простой группы Ли форма Киллинга–Картана совпадает с единичной матрицей

$$\eta_{AB} = \delta_{AB}. \quad (25.217)$$

Пример 25.5.1. В единых моделях теории поля используются, как правило, унитарные $\mathbb{SU}(N)$ и ортогональные $\mathbb{SO}(N)$ группы. Эти группы связны, компактны и просты. $\square$

Если вместо полей Янга–Миллса $A_\alpha{}^A$ рассматривать их точное матричное представление, то лагранжиан (25.215) можно записать, используя след матриц,

$$L_{\text{YM}} := \frac{1}{4} \sqrt{|g|} g^{\alpha\gamma} g^{\beta\delta} \text{tr} (F_{\alpha\beta} F_{\gamma\delta}).$$

Обратим внимание на изменение общего знака, что связано с определением формы Киллинга–Картана (25.216).

Для построения физически содержательных моделей математической физики, необходимо введение дополнительных полей, например, скаляров и спиноров. Мы предполагаем, что эти поля являются локальными сечениями векторного ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{G}, \mathbb{P})$, типичным слоем которого является векторное пространство $\mathbb{V}$, в котором задано представление структурной группы $\mathbb{G}$. Отображение $\pi_{\mathbb{E}} : \mathbb{E} \rightarrow \mathbb{M}$ – это проекция. Каждое сечение ассоциированного расслоения задается набором компонент $\varphi = \{\varphi^i\}$, $i = 1, \dots, \dim \mathbb{V}$. При калибровочных преобразованиях поля преобразуются по некоторому представлению структурной группы:

$$\varphi \mapsto \varphi' = \varphi S^{-1} = \{\varphi^j S^{-1}{}_j{}^i\}, \quad (25.218)$$

где $S_j{}^i$ – матрица представления. Ковариантная производная имеет вид (13.25)

$$\nabla_\alpha \varphi^i = \partial_\alpha \varphi^i + \varphi^j A_{\alpha j}{}^i, \quad (25.219)$$

в котором

$$A_{\alpha j}{}^i := -A_\alpha{}^A L_{Aj}{}^i,$$

где $L_{Aj}{}^i$ – представление базиса алгебры Ли L_A в векторном пространстве $\mathbb{V}$. Нетрудно убедиться в том, что при калибровочных преобразованиях (25.205), (25.218) ковариантная производная ведет себя ковариантно:

$$\nabla_\alpha \varphi^i \mapsto \nabla'_\alpha \varphi'^i = \nabla_\alpha \varphi^j S^{-1}{}_j{}^i,$$

где

$$\nabla'_\alpha \varphi'^i = \partial_\alpha \varphi'^i + \varphi'^j A'_{\alpha j}{}^i,$$

Определение. В физике векторное пространство $\mathbb{V}$ называется *изотопическим*, а калибровочные преобразования (25.218) – *изотопическими вращениями*. $\square$

При построении моделей математической физики мы предполагаем, что уравнения движения для полей φ , не взаимодействующих с полями Янга–Миллса, следуют из некоторого действия

$$S = \int_{\mathbb{M}} dx L(\varphi, \partial\varphi),$$

где лагранжиан L зависит от полей $\varphi = \{\varphi^i\}$ и их частных производных $\partial\varphi = \{\partial_\alpha \varphi^i\}$. Как и для электромагнитного поля, взаимодействие с полями Янга–Миллса вводится путем *минимальной подстановки*, которая заключается в замене всех частных производных на ковариантные:

$$\partial_\alpha \varphi^i \mapsto \nabla_\alpha \varphi^i := \partial_\alpha \varphi^i + \varphi^j A_{\alpha j}{}^i.$$

Кроме этого, к лагранжиану добавляется лагранжиан поля Янга–Миллса

$$L(\varphi, \partial\varphi) \mapsto -\frac{1}{4e^2} F^{\alpha\beta\Lambda} F_{\alpha\beta\Lambda} + L(\varphi, \nabla\varphi),$$

который приводит к динамическим уравнениям движения для самого поля Янга–Миллса. Выше e – константа связи и

$$F^{\alpha\beta\Lambda} := g^{\alpha\gamma} g^{\beta\delta} \eta^{\Lambda\mathbb{B}} F_{\gamma\delta\mathbb{B}}.$$

Для построения теории возмущений удобно перейти к новым переменным $A_\alpha{}^\Lambda \mapsto eA_\alpha{}^\Lambda$. Тогда лагранжиан примет вид

$$L = -\frac{1}{4} F^{\alpha\beta\Lambda} F_{\alpha\beta\Lambda} + L(\varphi, \nabla\varphi), \quad (25.220)$$

где

$$\begin{aligned} F_{\alpha\beta}{}^\Lambda &= \partial_\alpha A_\beta{}^\Lambda - \partial_\beta A_\alpha{}^\Lambda - e A_\alpha{}^\mathbb{B} A_\beta{}^\mathbb{C} f_{\mathbb{B}\mathbb{C}}{}^\Lambda, \\ \nabla_\alpha \varphi^i &= \partial_\alpha \varphi^i + e \varphi^j A_{\alpha j}{}^i. \end{aligned}$$

В таком виде удобно строить теорию возмущений при малой константе связи e . В дальнейшем, для упрощения обозначений, мы положим $e = 1$.

Если лагранжиан Янга–Миллса записан в виде (25.220), то размерности следующие:

$$[A_\alpha{}^\Lambda] = l^{1-\frac{n}{2}}, \quad [e] = l^{\frac{n}{2}-2}. \quad (25.221)$$

То есть размерность поля Янга–Миллса совпадает с размерностью потенциала электромагнитного поля. Заметим, что в четырехмерном пространстве-времени константа связи e безразмерна.

25.5.1 Лагранжева формулировка

В квантовой теории поля мы предполагаем, что пространством-временем является пространство Минковского, $\mathbb{M} \approx \mathbb{R}^{1,n-1}$, которое топологически тривиально, т.е. диффеоморфно $\mathbb{R}^n$. В этом случае согласно теореме 12.1.1 расслоение тривиально

$$\mathbb{P}(\mathbb{R}^{1,n-1}, \pi, \mathbb{G}) \approx \mathbb{R}^{1,n-1} \times \mathbb{G}.$$

Если выбрать систему координат, например, декартову, покрывающую все пространство-время, то связность на главном расслоении будет находиться во взаимно однозначном соответствии с локальной формой связности. По этой причине в физике часто говорят, что поля Янга–Миллса это и есть связность.

Действие для полей Янга–Миллса в пространстве Минковского $\mathbb{R}^{1,n-1}$ имеет вид

$$S_{\text{YM}} = \int dx L_{\text{YM}}, \quad L_{\text{YM}} := -\frac{1}{4} F^{\alpha\beta A} F_{\alpha\beta A}. \quad (25.222)$$

где подъем и опускание греческих индексов производится с помощью метрики Минковского $\eta_{\alpha\beta}$, а латинских – с помощью формы Киллинга–Картана η_{AB} . Из этого действия следуют нелинейные уравнения движения

$$\nabla_\beta F^{\beta\alpha A} := \partial_\beta F^{\beta\alpha A} - A_\beta{}^B F^{\beta\alpha C} f_{BC}{}^A = 0. \quad (25.223)$$

Добавление к лагранжиану Янга–Миллса массового члена $\frac{1}{2} m^2 A^{\alpha A} A_{\alpha A}$, как легко видеть, нарушает калибровочную инвариантность.

В отличие от электродинамики в пространстве Минковского для неабелевых калибровочных групп $\mathbb{G}$ уравнения (25.223) нелинейны. Поэтому говорить о “свободных” полях Янга–Миллса не имеет смысла. Взаимодействие полей Янга–Миллса есть всегда. В линейном приближении уравнения (25.223) имеют тот же вид, что и набор из N электромагнитных полей. При этом инфинитезимальные калибровочные преобразования (25.205) в линейном приближении по полям принимают вид $\mathbb{U}(1)$ преобразований:

$$A'_\alpha{}^A = A_\alpha{}^A + \partial_\alpha \epsilon^A + \mathfrak{o}_\alpha{}^A(\epsilon, A),$$

где $\epsilon^A(x)$, $A = 1, \dots, N$, – параметры преобразований. Таким образом, в линейном приближении свободное поле Янга–Миллса описывает N безмассовых векторных полей.

По построению, действие калибровочно инвариантно. Согласно второй теореме Нетер в таком случае между уравнениями движения существует N линейных зависимостей по числу параметров преобразования. Для нахождения этих зависимостей достаточно знать вид инфинитезимальных калибровочных преобразований (25.203). Из определения инвариантности следует равенство

$$\delta S_{\text{YM}} = \int dx \frac{\delta S_{\text{YM}}}{\delta A_\alpha{}^A} \nabla_\alpha \epsilon^A = 0.$$

После интегрирования по частям, получим зависимость уравнений движения:

$$\nabla_\alpha \nabla_\beta F^{\beta\alpha}{}_A = 0, \quad A = 1, \dots, N. \quad (25.224)$$

В справедливости полученного равенства можно убедиться с помощью прямых вычислений. Действительно, в силу антисимметрии напряженности, $F^{\alpha\beta}{}_A = -F^{\beta\alpha}{}_A$, равенство (25.224) можно переписать в виде

$$\frac{1}{2} (\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha) F^{\beta\alpha}{}_A = \frac{1}{2} F_{\alpha\beta}{}^C F^{\beta\alpha}{}_A f_{CA}{}^B = 0,$$

где мы воспользовались свойством коммутатора ковариантных производных (13.44), симметрией тензора $F_{\alpha\beta}{}^c F^{\beta\alpha b}$ по индексам c и b и антисимметрией структурных констант (предложение 8.5.1).

Если в правой части уравнений движения (25.223) стоит ток,

$$\nabla_\beta F^{\beta\alpha}{}_A = J^\alpha{}_A, \quad (25.225)$$

То из тождества (25.224) следует закон ковариантного сохранения тока:

$$\nabla_\alpha J^\alpha{}_A = 0, \quad (25.226)$$

выполнение которого необходимо для самосогласованности уравнений движения. В свою очередь, если ток получен в результате варьирования некоторого калибровочно инвариантного действия для других полей, то равенство (25.226) будет автоматически выполнено на уравнениях движения, опять же, в силу второй теоремы Нетер. Если же ток $J^\alpha{}_A$ вводится в уравнения движения “руками”, то на него необходимо наложить условие сохранения (25.226).

Эффективным инструментом исследования свободных уравнений движения в электродинамике явилось разложение пространственных компонент электромагнитного потенциала на поперечную и продольную составляющие. Для каждого значения индекса A поле Янга–Миллса также можно разложить на поперечную и продольные части. Однако это не будет столь же эффективно, т.к. уравнения “свободного” поля Янга–Миллса (25.223) нелинейны. Такое разложение эффективно в квантовой теории поля при построении теории возмущений.

Теперь рассмотрим законы сохранения, вытекающие из первой теоремы Нетер. По построению, действие Янга–Миллса инвариантно относительно глобального действия группы Пуанкаре $\mathbb{O}(1, n-1)$. Из инвариантности относительно трансляций, согласно первой теореме Нетер, следует закон сохранения энергии-импульса (см. раздел 20.2.1),

$$\partial_\beta \hat{T}^\beta{}_\alpha = 0,$$

где

$$\hat{T}_{\alpha\beta} = -\partial_\alpha A_{\gamma A} F_\beta{}^{\gamma A} + \frac{1}{4} \eta_{\alpha\beta} F^2$$

– канонический тензор энергии-импульса поля Янга–Миллса. Это выражение страдает двумя недостатками. Во-первых, оно не инвариантно относительно калибровочных преобразований. Во-вторых, отсутствует симметрия относительно перестановки индексов. Для их устранения используем произвол в выборе канонического тензора энергии-импульса. Заметим, что выполнено равенство

$$\partial_\gamma (A_{\alpha A} F_\beta{}^{\gamma A}) = \partial_\gamma A_{\alpha A} F_\beta{}^{\gamma A} + A_{\alpha A} \nabla_\gamma F_\beta{}^{\gamma A} + A_{\alpha A} A_\gamma{}^c f_{cb}{}^A F_\beta{}^{\gamma B}.$$

Если выполнены уравнения движения, что предполагается в первой теореме Нетер, то второе слагаемое обращается в нуль, и можно переопределить тензор энергии-импульса

$$T^s_{\alpha\beta} := \hat{T}_{\alpha\beta} + \partial_\gamma (A_{\alpha A} F_\beta{}^{\gamma A}) = -F_{\alpha\gamma A} F_\beta{}^{\gamma A} + \frac{1}{4} \eta_{\alpha\beta} F^2. \quad (25.227)$$

Этот тензор энергии-импульса, очевидно, калибровочно инвариантен и симметричен. На уравнениях движения он сохраняется:

$$\partial_\beta T^s{}^\beta{}_\alpha = 0. \quad (25.228)$$

При лоренцевых вращениях поле Янга–Миллса преобразуется, как и электромагнитное поле (25.100). Поэтому для спинового момента получаем следующее выражение

$$S_{\alpha\beta}{}^\gamma = A_{\alpha A} F_\beta{}^{\gamma A} - A_{\beta A} F_\alpha{}^{\gamma A}. \quad (25.229)$$

Полный тензор момента количества движения (20.51) состоит из орбитального и спинового моментов:

$$\widehat{J}_{\alpha\beta}{}^\gamma = x_\beta \widehat{T}_\alpha{}^\gamma - x_\alpha \widehat{T}_\beta{}^\gamma + S_{\alpha\beta}{}^\gamma. \quad (25.230)$$

Используя выражения для симметричного тензора энергии-импульса (25.227) и спинового момента (25.229) его можно переписать в виде

$$\widehat{J}_{\alpha\beta}{}^\gamma = M_{\alpha\beta}{}^\gamma + \partial_\delta (x_\alpha A_{\beta A} F^{\gamma\delta A} - x_\beta A_{\alpha A} F^{\gamma\delta A}), \quad (25.231)$$

где

$$M_{\alpha\beta}{}^\gamma := x_\beta T_\alpha{}^{\text{S}\gamma} - x_\alpha T_\beta{}^{\text{S}\gamma}$$

– орбитальный момент, построенный по симметричному тензору энергии-импульса (25.227). Поскольку второе слагаемое в выражении (25.231) имеет вид (20.39), то тензор момента количества движения можно переопределить, положив

$$J_{\alpha\beta}{}^\gamma = M_{\alpha\beta}{}^\gamma = x_\beta T_\alpha{}^{\text{S}\gamma} - x_\alpha T_\beta{}^{\text{S}\gamma}, \quad (25.232)$$

который также сохраняется

$$\partial_\gamma J_{\alpha\beta}{}^\gamma = 0.$$

Отметим полное отсутствие спинового момента в этом выражении.

Поскольку действие для поля Янга–Миллса инвариантно относительно калибровочных преобразований, то оно также инвариантно относительно глобальных преобразований из калибровочной группы. Поэтому можно воспользоваться первой теоремой Нетер. С калибровочной инвариантностью действия для поля Янга–Миллса связан закон сохранения заряда

$$\partial_\alpha J^\alpha{}_A = 0, \quad (25.233)$$

где

$$J^\alpha{}_A = \nabla_\beta F^{\beta\alpha}{}_A \quad (25.234)$$

– ток, соответствующий первой теореме Нетер, т.е. когда параметр калибровочных преобразований не зависит от точки пространства-времени. При отсутствии полей материи этот закон тривиален, т.к. ток равен нулю в силу уравнений движения (25.223).

25.5.2 Гамильтонова формулировка

Рассмотрим гамильтонову формулировку поля Янга–Миллса в пространстве Минковского $\mathbb{R}^{1,n-1}$ с действием (25.222). По определению, импульсы, сопряженные компонентам $A_\alpha{}^A$, имеют вид

$$P^\alpha{}_A := \frac{\partial L_{\text{YM}}}{\partial \dot{A}_\alpha{}^A} = F^{0\alpha}{}_A, \quad (25.235)$$

где $\dot{A}_\alpha{}^A := \partial_0 A_\alpha{}^A$. Выпишем отличные от нуля одновременные скобки Пуассона:

$$[A_\alpha{}^A, P^\beta{}_B] := [A_\alpha{}^A(x^0, \mathbf{x}), P^\beta{}_B(x^0, \mathbf{x}')] = \delta_\alpha^\beta \delta_B^A \delta(\mathbf{x} - \mathbf{x}').$$

Как и в случае электромагнитного поля в теории возникают первичные связи. Из определения импульсов (25.222) следует N первичных связей:

$$G_{1A} = P^0_A = 0, \quad A = 1, \dots, N. \quad (25.236)$$

Других первичных связей в модели нет. Скобки Пуассона первичных связей между собой, очевидно, равны нулю:

$$[G_{1A}, G'_{1B}] = 0.$$

Несложные вычисления приводят к гамильтониану

$$\mathcal{H} = \int d\mathbf{x} \left(-\frac{1}{2} P^{\mu A} P_{\mu A} + \frac{1}{4} F^{\mu\nu A} F_{\mu\nu A} - A_0^A \nabla_\mu P^\mu_A + \lambda^A P^0_A \right), \quad (25.237)$$

где мы добавили первичные связи с множителями Лагранжа λ^A и ковариантная производная от импульсов равна

$$\nabla_\mu P^\nu_A := \partial_\mu P^\nu_A - A_{\mu A}^B P^\nu_B.$$

Гамильтоновы уравнения движения для построенного гамильтониана имеют вид

$$\dot{A}_0^A = \lambda^A, \quad (25.238)$$

$$\dot{A}_\mu^A = -P_\mu^A + \nabla_\mu A_0^A, \quad (25.239)$$

$$\dot{P}^0_A = \nabla_\mu P^\mu_A, \quad (25.240)$$

$$\dot{P}^\mu_A = \nabla_\nu F^{\nu\mu}_A + A_{0A}^B P^\mu_B, \quad (25.241)$$

где

$$\begin{aligned} \nabla_\mu A_0^A &:= \partial_\mu A_0^A + A_{\mu B}^A A_0^B, \\ \nabla_\nu F^{\rho\mu}_A &:= \partial_\nu F^{\rho\mu}_A - A_{\nu A}^B F^{\rho\mu}_B. \end{aligned}$$

Конечно, эти уравнения необходимо дополнить уравнениями первичных связей (25.236), которые возникают при варьировании соответствующего действия по множителям Лагранжа.

Гамильтоновы уравнения движения (25.238)–(25.241), содержат множители Лагранжа λ^A , которые в исходном действии рассматриваются в качестве произвольных функций. Однако, уравнения движения, в принципе, могут наложить на них некоторые ограничения. Чтобы проверить совместность гамильтоновых уравнений, рассмотрим эволюцию первичных связей во времени. Поскольку $P^0_A = 0$, то из уравнения (25.240) следуют равенства

$$\nabla_\mu P^\mu_A = 0, \quad (25.242)$$

которые должны быть выполнены для любого решения гамильтоновых уравнений. С учетом уравнения (25.241) производная по времени от полученных равенств имеет вид

$$\partial_0(\nabla_\mu P^\mu_A) = \nabla_\mu \nabla_\nu F^{\nu\mu}_A + \nabla_\mu A_{0A}^B P^\mu_B + A_{0A}^B \nabla_\mu P^\mu_B - \dot{A}_{\mu A}^B P^\mu_B. \quad (25.243)$$

Первое слагаемое обращается в нуль:

$$\frac{1}{2}(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) F^{\nu\mu}_A = \frac{1}{2} F_{\mu\nu}^C F^{\nu\mu B} f_{CAB} = 0,$$

где мы воспользовались свойством коммутатора ковариантных производных (13.44), симметрией тензора $F_{\mu\nu}{}^c F^{\nu\mu b}$ по индексам c и b и антисимметрией структурных констант (предложение 8.5.1). Второе и четвертое слагаемые сокращаются с учетом уравнения (25.239). В результате возникает равенство

$$\partial_0(\nabla_\mu P^\mu{}_A) = A_0{}^B \nabla_\mu P^\mu{}_B \approx 0, \quad (25.244)$$

т.к. правая часть равна нулю на уравнениях движения (25.242). На этом анализ заканчивается, и мы видим, что никаких ограничений на множители Лагранжа гамильтоновы уравнения движения не накладывают.

Продолжим общий гамильтонов анализ. Производная по времени от первичных связей равна

$$\dot{G}_1 = [G_1, \mathcal{H}] = \nabla_\mu P^\mu{}_A.$$

Это приводит к вторичным связям:

$$G_{2A} = \nabla_\mu P^\mu{}_A \approx 0. \quad (25.245)$$

Скобка Пуассона первичных и вторичных связей между собой равна нулю,

$$[G_{1A}, G'_{2B}] = 0.$$

Несложно вычислить скобки Пуассона вторичных связей между собой:

$$[G_{2A}, G'_{2B}] = -f_{AB}{}^C G_{2C}. \quad (25.246)$$

Правая часть, очевидно, обращается в нуль на поверхности связей.

Выше мы уже показали, что вторичные связи сохраняются во времени (25.244).

Таким образом, полная система связей

$$\{G_a\} := \{G_{1A}, G_{2A}\}, \quad a = 1, \dots, 2N, \quad (25.247)$$

находится в инволюции, и модель содержит $2N$ связей первого рода.

Полный гамильтониан получается после добавления всех связей:

$$\mathcal{H}_T = \int d\mathbf{x} \left(-\frac{1}{2} P^{\mu A} P_{\mu A} + \frac{1}{4} F^{\mu\nu A} F_{\mu\nu A} + \lambda^A P^0{}_A + \mu^A \nabla_\mu P^\mu{}_A \right),$$

где слагаемые $A_0{}^A \nabla_\mu P^\mu{}_A$ мы абсорбировали, переопределив множители Лагранжа μ^A . Ему соответствует полное действие

$$S_T = \int dx \left(P^\alpha{}_A \dot{A}^\alpha{}_A + \frac{1}{2} P^{\mu A} P_{\mu A} - \frac{1}{4} F^{\mu\nu A} F_{\mu\nu A} - \lambda^A P^0{}_A - \mu^A \nabla_\mu P^\mu{}_A \right). \quad (25.248)$$

Это действие является более общим, чем исходное (25.222), т.к. совпадает с ним только при $\lambda^A = 0$ и $\mu^A = -A_0{}^A$.

Гамильтоновы уравнения движения для действия (25.248) имеют вид:

$$\frac{\delta S_T}{\delta P^0_A} = \dot{A}_0^A - \lambda^A = 0, \quad (25.249)$$

$$\frac{\delta S_T}{\delta P^\mu_A} = \dot{A}_\mu^A + P_\mu^A + \nabla_\mu \mu^A = 0, \quad (25.250)$$

$$\frac{\delta S_T}{\delta A_0^A} = -\dot{P}^0_A = 0, \quad (25.251)$$

$$\frac{\delta S_T}{\delta A_\mu^A} = -\dot{P}^\mu_A + \nabla_\nu F^{\nu\mu}_A - \mu^B f_{AB}^C P^\mu_C = 0, \quad (25.252)$$

$$\frac{\delta S_T}{\delta \lambda^A} = -P^0_A = 0, \quad (25.253)$$

$$\frac{\delta S_T}{\delta \mu^A} = -\nabla_\mu P^\mu_A = 0. \quad (25.254)$$

Согласно теореме 22.3.1 каждой связи первого рода соответствует калибровочное преобразование, относительно которого полное действие инвариантно. Первичным связям (25.236) соответствует генератор калибровочных преобразований

$$T_1 := \int d\mathbf{x} \epsilon_1^A P^0_A$$

с некоторыми малыми параметрами $\epsilon_1^A(x)$. При этом преобразуются только временные компоненты поля Янга–Миллса A_0^A и множители Лагранжа λ^A :

$$\delta_1 A_0^A = [A_0^A, T_1] = \epsilon_1^A, \quad \delta_1 \lambda^A = \dot{\epsilon}_1^A. \quad (25.255)$$

Все остальные поля остаются без изменения. Согласно второй теореме Нетер из инвариантности действия относительно калибровочного преобразования (25.255) следует зависимость уравнений движения:

$$\frac{\delta S_T}{\delta A_0^A} - \partial_t \frac{\delta S_T}{\delta \lambda^A} = 0.$$

В справедливости этого тождества нетрудно убедиться прямой проверкой.

Вторичным связям первого рода (25.245) также соответствует генератор преобразований

$$T_2 := \int d\mathbf{x} \epsilon_2^A \nabla_\mu P^\mu_A,$$

где $\epsilon_2^A(x)$ – малые параметры второго калибровочного преобразования. В этом случае преобразуются только пространственные компоненты поля Янга–Миллса A_μ^A , сопряженные импульсы P^μ_A и множители Лагранжа μ^A :

$$\begin{aligned} \delta_2 A_\mu^A &= [A_\mu^A, T_2] = -\nabla_\mu \epsilon_2^A, \\ \delta_2 P^\mu_A &= [P^\mu_A, T_2] = -\epsilon_2^B f_{AB}^C P^\mu_C, \\ \delta_2 \mu^A &= \dot{\epsilon}_2^A - \epsilon_2^B \mu^C f_{BC}^A. \end{aligned}$$

Согласно второй теореме Нетер, из локальной инвариантности действия следует зависимость уравнений Эйлера–Лагранжа:

$$\nabla_\mu \frac{\delta S_T}{\delta A_\mu^A} - \frac{\delta S_T}{\delta P^\mu_B} f_{AB}^C P^\mu_C - \partial_t \frac{\delta S_T}{\delta \mu^A} - \mu^B f_{AB}^C \frac{\delta S_T}{\delta \mu^C} = 0,$$

что также легко проверить. Таким образом, как и в электродинамике, число параметров калибровочных преобразований полного действия S_T в гамильтоновом формализме по сравнению с лоренцевой формулировкой удвоилось.

Генератором калибровочных преобразований в лагранжевом формализме (25.205) является линейная комбинация связей первого рода:

$$T = \int dx \left[(\dot{\epsilon}^A - A_0^B \epsilon^C f_{BC}^A) P_A^0 - \epsilon^A \nabla_\mu P_A^\mu \right].$$

Действительно,

$$\begin{aligned} \delta A_0^A &= [A_0^A, T] = \dot{\epsilon}^A - A_0^B \epsilon^C f_{BC}^A = \nabla_0 \epsilon^A, \\ \delta A_\mu^A &= [A_\mu^A, T] = \nabla_\mu \epsilon^A, \\ \delta P_A^0 &= \epsilon^B f_{AB}^C P_C^0, \\ \delta P_A^\mu &= \epsilon^B f_{AB}^C P_C^\mu. \end{aligned}$$

По сравнению с электродинамикой возникли нетривиальные преобразования импульсов. Это связано с тем, что в электродинамике компоненты напряженности инвариантны, а для полей Янга–Миллса – ковариантны.

Решения гамильтоновых уравнений движения содержат произвол, соответствующий множителям Лагранжа λ^A и μ^A . Для того, чтобы устранить этот произвол необходимо наложить $2N$ калибровочных условий. В квантовой теории поля обычно используют такие же калибровочные условия (25.96) как и в электродинамике для каждого значения индекса A . Если зафиксировать, например, кулоновскую калибровку $\partial_\mu A^\mu = 0$, то это даст N калибровочных условий. Недостающие N калибровочных условий в гамильтоновом формализме можно получить, рассматривая эволюцию этих калибровочных условий во времени. Из уравнения (25.239) следует

$$\partial_0(\partial_\mu A^\mu_A) = -\partial_\mu P^\mu_A + \partial^\mu \nabla_\mu A_{0A}.$$

Тогда кулоновскую калибровку для потенциала можно дополнить N калибровочными условиями

$$-\partial_\mu P^\mu_A + \partial^\mu \nabla_\mu A_{0A} = 0.$$

В этом случае калибровочные условия будут согласованы с гамильтоновыми уравнениями для исходного гамильтониана (25.237). Вместо этих калибровочных можно выбрать другие, например, $A_{0A} = 0$.

Гамильтонова формулировка поля Янга–Миллса на произвольном многообразии M с метрикой g проводится, по существу, так же как и в пространстве Минковского. Необходимо произвести замену $\eta_{\alpha\beta} \mapsto g_{\alpha\beta}$ и рассматривать компоненты $g_{\alpha\beta}$ как внешнее поле. Почти все формулы останутся при этом без изменения, только в ковариантных производных появятся символы Кристоффеля, возникающие при интегрировании по частям.

25.5.3 Поля Янга–Миллса в аффинной геометрии

Как уже было отмечено, для построения моделей математической физики на многообразии M (пространстве-времени) кроме поля Янга–Миллса желательно иметь метрику и аффинную связность, т.е. аффинную геометрию. Поскольку поле Янга–Миллса – это компоненты локальной формы связности, то действие для “свободного” поля Янга–Миллса (25.215) зависит только от метрики, но не от связности.

Уравнения движения и тензор энергии-импульса поля Янга–Миллса получаются простым варьированием действия (25.215):

$$\sqrt{|g|}S_{\text{YM},\alpha}{}^{\text{A}} := \frac{\delta S_{\text{YM}}}{\delta A_{\alpha}{}^{\text{A}}} = \partial_{\beta}(\sqrt{|g|}F^{\beta\alpha}{}_{\text{A}}) = \sqrt{|g|}\tilde{\nabla}_{\beta}F^{\beta\alpha}{}_{\text{A}} = 0, \quad (25.256)$$

$$\sqrt{|g|}S_{\text{YM},\alpha\beta} := \frac{\delta S_{\text{YM}}}{\delta g^{\alpha\beta}} = \frac{1}{2}\sqrt{|g|}T_{\text{YM}\alpha\beta} = -\frac{1}{2}\sqrt{|g|}F_{\alpha\gamma\text{A}}F_{\beta}{}^{\gamma\text{A}} + \frac{1}{8}\sqrt{|g|}g_{\alpha\beta}F^2, \quad (25.257)$$

где

$$\tilde{\nabla}_{\beta}F^{\beta\alpha}{}_{\text{A}} := \partial_{\beta}F^{\beta\alpha}{}_{\text{A}} + F^{\gamma\alpha}{}_{\text{A}}\tilde{\Gamma}_{\beta\gamma}{}^{\beta} - A_{\beta\text{A}}{}^{\text{B}}F^{\beta\alpha}{}_{\text{B}}.$$

Слагаемое с символами Кристоффеля $\tilde{\Gamma}$ возникло при интегрировании по частям (см. раздел 6.6). Отсюда вытекает выражение для тензора энергии-импульса поля Янга–Миллса

$$T_{\text{YM}\alpha\beta} = -F_{\alpha\gamma\text{A}}F_{\beta}{}^{\gamma\text{A}} + \frac{1}{4}g_{\alpha\beta}F^2. \quad (25.258)$$

Полученное выражение является ковариантным обобщением симметричного канонического тензора энергии-импульса (25.227), полученного в пространстве Минковского из первой теоремы Нетер.

Если поле Янга–Миллса взаимодействует с другими полями, то в правой части уравнений движения (25.256) возникают токи:

$$\tilde{\nabla}_{\beta}F^{\beta\alpha}{}_{\text{A}} = J^{\alpha}{}_{\text{A}}. \quad (25.259)$$

Как и в случае электродинамики, для поля Янга–Миллса можно воспользоваться формализмом первого порядка.

Предложение 25.5.2. Уравнения движения для действия (25.215) эквивалентны уравнениям Эйлера–Лагранжа для действия

$$S = - \int_{\mathbb{M}} dx \sqrt{|g|} \left[\frac{1}{2} F^{\alpha\beta\text{A}} (\partial_{\alpha} A_{\beta\text{A}} - \partial_{\beta} A_{\alpha\text{A}}) - \frac{1}{4} F^{\alpha\beta\text{A}} F_{\alpha\beta\text{A}} \right], \quad (25.260)$$

в котором компоненты $F^{\alpha\beta\text{A}}$ и $A_{\alpha\text{A}}$ рассматриваются в качестве независимых переменных.

Доказательство. Повторяет доказательство предложения 25.3.4 для электродинамики. $\square$

Действие для поля Янга–Миллса (25.215) инвариантно относительно общих преобразований координат. Соответствующие вариации полей имеют вид (см. раздел 2.13)

$$\begin{aligned} \delta A_{\alpha}{}^{\text{A}} &= -\partial_{\alpha}\epsilon^{\beta}A_{\beta}{}^{\text{A}} - \epsilon^{\beta}\partial_{\beta}A_{\alpha}{}^{\text{A}}, \\ \delta g^{\alpha\beta} &= g^{\alpha\gamma}\partial_{\gamma}\epsilon^{\beta} + g^{\beta\gamma}\partial_{\gamma}\epsilon^{\alpha} - \epsilon^{\gamma}\partial_{\gamma}g^{\alpha\beta}. \end{aligned}$$

Отсюда, согласно второй теореме Нетер (20.63), следует зависимость уравнений движения

$$\partial_{\beta}S,{}^{\beta\text{A}}A_{\alpha\text{A}} + S,{}^{\beta\text{A}}F_{\beta\alpha\text{A}} - \partial_{\beta}S,{}^{\beta}{}_{\alpha} - \partial_{\beta}S,{}^{\beta}{}_{\alpha} - S,{}_{\beta\gamma}\partial_{\alpha}g^{\beta\gamma} = 0. \quad (25.261)$$

Или, в ковариантном виде,

$$\tilde{\nabla}_{\beta}S,{}^{\beta\text{A}}A_{\alpha\text{A}} + S,{}^{\beta\text{A}}F_{\beta\alpha\text{A}} - 2\tilde{\nabla}^{\beta}S,{}_{\beta\alpha} = 0. \quad (25.262)$$

Если выполнены уравнения движения для поля Янга–Миллса (25.256), то с учетом свернутых тождеств Бианки соотношение (25.262) принимает вид

$$\tilde{\nabla}_\beta T_{\gamma\mu\alpha}{}^\beta = 0. \quad (25.263)$$

Полученное равенство, выполненное для всех решений уравнений Эйлера–Лагранжа, можно интерпретировать, как ковариантное обобщение закона сохранения тензора энергии-импульса (25.228).

Действие для поля Янга–Миллса инвариантно также относительно калибровочных преобразований (25.203). Этой инвариантности, согласно второй теореме Нетер, соответствует зависимость уравнений движения:

$$\tilde{\nabla}_\alpha S, {}^\alpha_A = 0. \quad (25.264)$$

С учетом уравнения (25.259) это равенство приводит к закону сохранения тока: $\tilde{\nabla}_\alpha J^\alpha_A = 0$. Если ток возникает при варьировании некоторого калибровочно инвариантного действия для полей материи, то сохранение тока будет выполняться автоматически в силу второй теоремы Нетер. Если же ток вводится в уравнения для поля Янга–Миллса “руками”, то условие $\tilde{\nabla}_\alpha J^\alpha_A = 0$ необходимо для самосогласованности уравнений.

25.5.4 Поле Янга–Миллса в общей теории относительности

Если ограничиться рассмотрением четырехмерного пространства-времени, то в теории поля Янга–Миллса возникают определенные специфические свойства. Как и в электродинамике поле Янга–Миллса определяет локальные формы связности

$$A^A := dx^\alpha A_\alpha{}^A, \quad A = 1, \dots, N.$$

Теперь локальная форма кривизны (25.206) не является внешней производной от локальной формы связности. Однако ее можно записать с помощью внешней ковариантной производной, введенной в разделе 13.2,

$$R^A = DA^A := dA^A - \frac{1}{2}A^B \wedge A^C f_{BC}{}^A. \quad (25.265)$$

Тождества Бианки (25.211) также можно записать в компактном виде с помощью оператора внешнего ковариантного дифференцирования (13.45):

$$DR^A = 0.$$

На этом, собственно, преимущество использования обозначений дифференциальных форм в теории полей Янга–Миллса заканчивается.

В общей теории относительности для каждого значения изотопического индекса A можно ввести 2-формы, дуальные к форме локальной кривизны:

$$*F^A := \frac{1}{2}dx^\alpha \wedge dx^\beta (*F^A)_{\alpha\beta}, \quad (25.266)$$

где компоненты дуальной формы кривизны заданы равенством

$$*F_{\alpha\beta}{}^A := \frac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}F^{\gamma\delta A}. \quad (25.267)$$

В приведенной формуле $\varepsilon_{\alpha\beta\gamma\delta}(x)$ – полностью антисимметричный тензор четвертого ранга.

Легко проверить, что оператор Ходжа, дважды примененный к 2-форме кривизны меняет ее знак: $**F^A = -F^A$. По этой причине при лоренцевой сигнатуре метрики кривизну нельзя разложить на самодуальную и антисамодуальную части.

Используя дуальную форму кривизны, тензор энергии-импульса поля Янга–Миллса (25.258) можно переписать в виде

$$T_{\gamma\mu\alpha\beta} = -\frac{1}{2} [F_{\alpha\gamma A} F_{\beta}^{\gamma A} + *F_{\alpha\gamma A} * F_{\beta}^{\gamma A}]. \quad (25.268)$$

Он, очевидно, инвариантен относительно дуальных вращений (25.162) для каждого значения индекса A . Если базис алгебры Ли выбран таким образом, что форма Киллинга–Картана совпадает с символом Кронекера, то эту инвариантность можно расширить до группы $\mathbb{SO}(2N)$.

Так же как и в электродинамике, доказывается

Предложение 25.5.3. *Если метрика на многообразии $\mathbb{M}$ имеет лоренцеву сигнатуру и координата x^0 является временем, то временная компонента тензора энергии-импульса поля Янга–Миллса*

$$T_{\text{EM}00} = -\frac{1}{2} [F_{0\gamma A} F_0^{\gamma A} + *F_{0\gamma A} * F_0^{\gamma A}] \quad (25.269)$$

положительно определена и, следовательно, удовлетворяет слабому энергетическому условию. Она также удовлетворяет сильному энергетическому условию.

Ситуация с полем Янга–Миллса во многих отношениях аналогична ситуации в электродинамике.

Предложение 25.5.4. *Действие для полей Янга–Миллса в аффинной геометрии (25.215) инвариантно относительно преобразований Вейля, не затрагивающих само поле Янга–Миллса:*

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = e^{2\phi} g_{\alpha\beta}, \quad A_{\alpha}^A \mapsto \bar{A}_{\alpha}^A = A_{\alpha}^A, \quad (25.270)$$

где $\phi(x) \in C^2(\mathbb{M})$ – произвольная вещественнозначная функция.

Доказательство. Утверждение следует из равенства

$$\sqrt{|g|} g^{\alpha\beta} g^{\gamma\delta} = \sqrt{|\bar{g}|} \bar{g}^{\alpha\beta} \bar{g}^{\gamma\delta}. \quad \square$$

Как следствие общего утверждения 25.3.7, получаем, что след тензора энергии-импульса поля Янга–Миллса (25.258) для четырехмерного пространства-времени равен нулю, $T_{\gamma\mu\alpha}^{\alpha} = 0$. Это, конечно, легко проверить.

Глава 26

Алгебры Клиффорда

Алгебры Клиффорда представляют собой довольно общую алгебраическую конструкцию, которая в простейших случаях включает в себя поле комплексных чисел, матрицы Паули и Дирака и многие другие хорошо известные понятия. Они играют важную роль в математической физике, например, в описании спинорных представлений групп вращений в евклидовом и псевдоевклидовом пространствах. В настоящей главе мы дадим определение алгебр Клиффорда, опишем основные свойства и покажем, что их можно использовать для построения групп Ли, которые являются универсальными накрывающими для групп вращений. Подробное изложение затронутых ниже вопросов можно найти в книгах [193, 194, 195, 196].

Перед описанием алгебр Клиффорда напомним несколько общих определений.

Определение. Алгеброй $\mathfrak{a}$ над полем $\mathbb{F}$ называется векторное пространство над $\mathbb{F}$, на котором задано умножение

$$\mathfrak{a} \ni f, g \mapsto fg \in \mathfrak{a},$$

которое согласовано с линейной структурой:

$$\begin{aligned} f(g+h) &= fg + fh, & (g+h)f &= gf + hf, & \forall f, g, h \in \mathfrak{a}, \\ a(fg) &= (af)g = f(ag), & \forall f, g \in \mathfrak{a}, \quad \forall a \in \mathbb{F}. \end{aligned} \quad (26.1)$$

Другими словами, алгеброй называется кольцо $\mathfrak{a}$ (сложение элементов и их умножение без условия ассоциативности), которое является векторным пространством над полем $\mathbb{F}$, и удовлетворяет условию (26.1). Алгебра называется *ассоциативной*, если выполнено условие

$$(fg)h = f(gh) = fgh, \quad \forall f, g, h \in \mathfrak{a}.$$

Алгебра называется *коммутативной*, если $fg = gf$. Алгебра называется *алгеброй с единицей* $e \in \mathfrak{a}$, если выполнено условие

$$ef = fe = f, \quad \forall f \in \mathfrak{a}.$$

Ассоциативная алгебра с единицей называется *унитальной*. Элемент унитарной алгебры f называется *обратимым*, если существует такой элемент $f^{-1} \in \mathfrak{a}$, что выполнены равенства

$$ff^{-1} = f^{-1}f = e.$$

Размерностью алгебры называется ее размерность как векторного пространства.

Центром алгебры называется множество тех элементов из $\mathfrak{a}$, которые коммутируют со всеми элементами алгебры $\mathfrak{a}$. □

Легко проверить, что центр алгебры образует подалгебру в $\mathfrak{a}$.

Таким образом, в алгебре определено три операции: сложение, умножение и умножение на числа (скаляры). В дальнейшем мы будем рассматривать алгебры над полями вещественных $\mathbb{R}$ или комплексных $\mathbb{C}$ чисел.

Алгебры, как правило, будут обозначаться готическими буквами.

Пример 26.0.2. Если в полях вещественных $\mathbb{R}$ и комплексных $\mathbb{C}$ чисел ввести базис $e := 1$, то их можно рассматривать как одномерные ассоциативные коммутативные алгебры с единицей над полями вещественных и комплексных чисел, соответственно. $\square$

Пример 26.0.3. Множество вещественных или комплексных квадратных $n \times n$ матриц с обычными операциями сложения, умножения и умножения на числа образует алгебру, которая обозначается $\mathbf{mat}(n, \mathbb{F})$, где $\mathbb{F} = \mathbb{R}$ или $\mathbb{C}$. Эта алгебра ассоциативна и обладает единицей $e = \mathbb{1}$, где $\mathbb{1}$ – единичная $n \times n$ матрица, т.е. унитарна. Ее вещественная и комплексная размерности равны

$$\dim \mathbf{mat}(n, \mathbb{R}) = n^2, \quad \dim_{\mathbb{C}} \mathbf{mat}(n, \mathbb{C}) = n^2.$$

При $n \geq 2$ алгебра матриц некоммутативна. Множество обратимых матриц образует общую линейную группу Ли $\mathbf{GL}(n, \mathbb{F})$.

Центром алгебры $\mathbf{mat}(n, \mathbb{F})$ является множество матриц, кратных единичной, т.е. вида $a\mathbb{1}$, где $a \in \mathbb{F}$.

В дальнейшем для алгебры матриц мы будем использовать также сокращенное обозначение $\mathbf{mat}(n, \mathbb{F}) := F(n)$. $\square$

Каждому элементу алгебры $f \in \mathfrak{a}$ можно поставить в соответствие его действие на векторном пространстве $\mathfrak{a}$ справа (или слева):

$$f : \mathfrak{a} \ni f, h \mapsto R_f h := hf \in \mathfrak{a}.$$

Поскольку умножение в алгебре согласовано с линейной структурой, то каждому элементу алгебры f ставится в соответствие линейный оператор R_f , который в фиксированном базисе задается некоторой матрицей.

Множество всех обратимых элементов $\mathbb{G}(\mathfrak{a})$ в унитарной алгебре $\mathfrak{a}$ образует группу по отношению к умножению. Ясно, что элемент f алгебры $\mathfrak{a}$ обратим тогда и только тогда, когда обратим линейный оператор R_f , т.е. определитель соответствующей матрицы отличен от нуля. Конечномерная алгебра $\mathfrak{a}$ над полем вещественных чисел как векторное пространство представляет собой многообразие, а условие $\det R_f \neq 0$ выделяет в нем открытое подмногообразие той же размерности. Следовательно, множество всех обратимых элементов алгебры представляет собой группу Ли той же размерности, что и сама алгебра, $\dim \mathbb{G} = \dim \mathfrak{a}$.

26.1 Определение алгебр Клиффорда

Мы начнем с общего, а затем дадим конструктивное определение алгебры Клиффорда путем явного построения векторного пространства и введения на нем клиффордова умножения.

Пусть задано векторное пространство $\mathbb{V}$, $\dim \mathbb{V} = n$, над полем вещественных $\mathbb{R}$ или комплексных $\mathbb{C}$ чисел. Зафиксируем базис векторного пространства e^a , $a =$

$1, \dots, n$. Тогда каждый вектор $A \in \mathbb{V}$ представим в виде разложения по данному базису: $A = e^a A_a$, где A_a – вещественные или комплексные числа, а само векторное пространство изоморфно либо $\mathbb{R}^n$, либо $\mathbb{C}^n$, которые рассматриваются как линейные пространства.

Комплексное векторное пространство $\mathbb{V}^{\mathbb{C}}$ является комплексификацией вещественного векторного пространства $\mathbb{V}$ той же размерности, $\dim_{\mathbb{C}}(\mathbb{V}^{\mathbb{C}}) = \dim \mathbb{V} = n$. При фиксированном базисе оно получается простой заменой вещественных компонент вектора на комплексные. Для краткости, индекс $\mathbb{C}$ мы часто писать не будем.

Замечание. В дальнейшем мы отождествим векторное пространство $\mathbb{V}$ с пространством 1-форм. Поэтому для обозначения векторов мы используем заглавные буквы их начала латинского алфавита, как и для форм в главе 3. $\square$

Если рассматривается вещественное векторное пространство, то предположим, что в нем задана симметричная билинейная форма

$$Q : \mathbb{V} \times \mathbb{V} \ni A, B \mapsto Q(A, B) \in \mathbb{R},$$

где симметричность означает справедливость равенства $Q(A, B) = Q(B, A)$ для всех A и B . В компонентах квадратичная форма задается симметричной вещественной матрицей η^{ab} :

$$Q(A, B) = \eta^{ab} A_a B_b.$$

В общем случае мы не требуем положительной определенности матрицы η . Более того, матрица η может быть вырожденной. Если матрица η положительно определена и, следовательно, невырождена, то она определяет в векторном пространстве $\mathbb{V}$ обычное евклидово скалярное произведение.

Если мы рассматриваем векторное пространство над полем комплексных чисел $\mathbb{C}$, то от квадратичной формы

$$Q : \mathbb{V} \times \mathbb{V} \ni A, B \mapsto Q(A, B) \in \mathbb{C},$$

обычно требуют линейность по первому аргументу,

$$Q(\alpha A_1 + \beta A_2, B) = \alpha Q(A_1, B) + \beta Q(A_2, B), \quad \forall \alpha, \beta \in \mathbb{C},$$

а вместо условия симметричности – выполнение равенства

$$Q(A, B) = Q(B, A)^{\dagger},$$

где знак $\dagger$ обозначает комплексное сопряжение. Это значит, что форма Q будет антилинейна по второму аргументу:

$$Q(A, \alpha B_1 + \beta B_2) = \alpha^{\dagger} Q(A, B_1) + \beta^{\dagger} Q(A, B_2), \quad \forall \alpha, \beta \in \mathbb{C}.$$

В компонентах форма Q будет задаваться некоторой эрмитовой матрицей $\eta = \eta^{\dagger}$:

$$Q(A, B) = \eta^{ab} A_a B_b^{\dagger}.$$

Если матрица η положительно определена, то векторное пространство $\mathbb{V}$ становится гильбертовым. Квадратичная форма Q в комплексном случае называется *эрмитовой* и *полуторалинейной*.

Базис и вещественного, и комплексного векторного пространства e^a определен с точностью до невырожденных линейных преобразований. Поэтому его всегда можно выбрать таким образом, что матрица η будет иметь канонический вид.

В приложениях, например, в квантовой механике и квантовой теории поля, рассматривают, как правило, эрмитовы квадратичные формы.

Предложение 26.1.1. Пусть задано вещественное или комплексное векторное пространство $\mathbb{V}$ с симметричной или эрмитовой квадратичной формой η , соответственно. Тогда в $\mathbb{V}$ существует такой базис, в котором квадратичная форма имеет канонический вид

$$\eta = \text{diag} \left(\underbrace{+\dots+}_p \underbrace{-\dots-}_q \underbrace{0\dots0}_{n-p-q} \right), \quad (26.2)$$

где p , q и $n - p - q$ — числа, соответственно, положительных, отрицательных и нулевых собственных значений матрицы η . Путем растяжки координат всегда можно добиться того, чтобы на диагонали стояли 1, -1 или 0.

Доказательство. См., например, [17], глава X, §9. $\square$

Таких базисов, в которых билинейная форма имеет канонический вид (26.2) существует бесконечно много.

В дальнейшем мы всегда предполагаем, что базис выбран таким образом, что квадратичная форма имеет канонический вид. При $n = p + q$ этот базис будем называть *ортонормированным*. В этом случае квадратичная форма является невырожденной.

При построении алгебр Клиффорда над полем комплексных чисел $\mathbb{C}$ мы рассматриваем не эрмитовы квадратичные формы, а просто симметричные билинейные квадратичные формы, которые *линейны* как по первому, так и по второму аргументу. Для таких форм условие симметрии выглядит точно так же как и в вещественном случае $Q(A, B) = Q(B, A)$. Если базис фиксирован, то такие формы задаются некоторой комплексной симметричной матрицей $\eta = \eta^T$. Чтобы не вводить новых обозначений и рассматривать параллельно оба случая $\mathbb{F} = \mathbb{R}$ и $\mathbb{F} = \mathbb{C}$, мы сохраним для симметричных билинейных форм в комплексном случае прежнюю запись $Q(A, B)$. Также для упрощения обозначений квадрат длины вектора (квадратичную форму) $A \in \mathbb{V}$ будем обозначать

$$Q(A) := Q(A, A),$$

опустив один аргумент.

Определение. Квадратичной формой Q на векторном пространстве $\mathbb{V}$ называется отображение $\mathbb{V} \rightarrow \mathbb{F}$ такое, что выполнены следующие условия:

- 1) $Q(\alpha A) = \alpha^2 Q(A)$, $\forall \alpha \in \mathbb{F}, A \in \mathbb{V}$;
- 2) симметричная форма

$$Q(A, B) := \frac{1}{2} [Q(A + B) - Q(A) - Q(B)] \quad (26.3)$$

линейна и по A , и по B .

Симметричная билинейная форма $Q(A, B)$ называется *поляризацией* квадратичной формы $Q(A)$. $\square$

Если рассматриваются билинейные квадратичные формы над полем комплексных чисел, то их канонический вид существенно проще.

Предложение 26.1.2. Пусть в комплексном векторном пространстве $\mathbb{V}$ задана симметричная билинейная квадратичная форма Q . Тогда в $\mathbb{V}$ существует такой базис, в котором форма имеет вид

$$\eta = \text{diag} \left(\underbrace{+\dots+}_m \underbrace{0\dots0}_{n-m} \right),$$

где m – ранг квадратичной формы.

Доказательство. См., например, [12], часть 2, §3, раздел 6. $\square$

Таким образом, количество различных квадратичных форм в комплексном случае существенно меньше, чем в вещественном. Это связано с возможностью умножать векторы базиса на мнимую единицу в комплексном случае. При этом значение квадратичной формы меняет знак: $Q(ie^a, ie^a) = -Q(e^a, e^a)$.

Теперь определим алгебру Клиффорда.

Определение. Пусть задано векторное пространство $\mathbb{V}$ над полем $\mathbb{F}$ с симметричной квадратичной формой Q . Тогда определена ассоциативная тензорная алгебра с единицей

$$\otimes \mathbb{V} := \bigoplus_{r=0}^{\infty} \mathbb{V}^r = \bigoplus_{r=0}^{\infty} \underbrace{\mathbb{V} \otimes \dots \otimes \mathbb{V}}_r,$$

где $\mathbb{V}^0 := \mathbb{F}$ и $\mathbb{V}^1 := \mathbb{V}$. В этой алгебре существует двусторонний идеал $\mathfrak{i}$, свободно порожденный элементами вида $A \otimes A - Q(A)e$, где $A \in \mathbb{V}$ и e – единица поля $\mathbb{F}$. Это значит, что идеал состоит из всех конечных линейных комбинаций:

$$\mathfrak{i} := \left\{ \sum_i U_i \otimes [A_i \otimes A_i - Q(A_i)e] \otimes V_i : \quad \forall U_i, V_i \in \otimes \mathbb{V}, \quad \forall A_i \in \mathbb{V} \right\}. \quad (26.4)$$

Тогда фактор алгебра

$$\text{cl}(\mathbb{V}, Q) := \frac{\otimes \mathbb{V}}{\mathfrak{i}}, \quad (26.5)$$

называется *алгеброй Клиффорда* над векторным пространством $\mathbb{V}$ с квадратичной формой $Q(A)$. $\square$

Поскольку $A \otimes A$ – тензор второго ранга, а $Q(A)e$ – тензор нулевого ранга, то их нельзя складывать, и в определении алгебры Клиффорда следовало бы писать знак прямой суммы:

$$A \otimes A \oplus (-Q(A))e.$$

Для упрощения обозначений в дальнейшем мы будем использовать общепринятую более короткую запись, которая использована в определении (26.4).

Из определения следует, что алгебра Клиффорда существует, ассоциативна и обладает единицей, т.е. унитарна.

Расшифруем данное определение. В определении идеала (26.4) элементы U_i и V_i являются конечными прямыми суммами тензоров различных рангов r . Поэтому элементы идеала не имеют определенного ранга, но всегда содержат слагаемые второго или более высокого ранга благодаря элементу $A \otimes A$. В идеале, что важно, отсутствуют тензоры нулевого и первого ранга.

По построению, элементы вида (26.4) образуют линейное подпространство в тензорном произведении $\otimes \mathbb{V}$. Если любой элемент этого подпространства умножить слева или справа на произвольный элемент тензорного произведения $U \in \otimes \mathbb{V}$, то в результате мы получим элемент из этого же подпространства. Следовательно $\mathfrak{i} \subset \otimes \mathbb{V}$ – двусторонний идеал.

Из формулы (26.3) вытекает равенство

$$\begin{aligned} A \otimes B + B \otimes A - 2Q(A, B)e &= \\ &= [(A + B) \otimes (A + B) - Q(A + B)e] - [A \otimes A - Q(A)e] - [B \otimes B - Q(B)e]. \end{aligned}$$

Поэтому можно считать, что идеал $\mathfrak{i}$ свободно порожден всеми элементами вида

$$A \otimes B + B \otimes A - 2Q(A, B)e, \quad (26.6)$$

где $A, B \in \mathbb{V}$.

Множество элементов алгебры Клиффорда (26.5) состоит из классов эквивалентности элементов тензорного произведения. А именно, два элемента $U, V \in \otimes \mathbb{V}$ называются эквивалентными, если их разность лежит в идеале

$$U \sim V \quad \Leftrightarrow \quad U - V \in \mathfrak{i}.$$

Таким образом, для конструктивного описания алгебры Клиффорда необходимо выбрать из каждого класса эквивалентности по одному представителю.

Прежде всего заметим, что два элемента поля $\alpha, \beta \in \mathbb{F}$ не могут лежать в одном классе эквивалентности, т.к. любой элемент идеала должен содержать тензор не ниже второго ранга. Следовательно, каждый элемент поля $\mathbb{F}$ является представителем некоторого класса эквивалентности. Это значит, что существует каноническое вложение

$$\varrho : \quad \mathbb{F} \hookrightarrow \mathfrak{cl}(\mathbb{V}, Q).$$

При этом вложении единица e поля $\mathbb{F}$ отображается в единицу алгебры Клиффорда. Поэтому мы сохраним обозначение e для единицы алгебры Клиффорда (26.5): $\varrho(e) = e$.

Аналогично, разность двух векторов $A, B \in \mathbb{V}$ является вектором: $A - B \in \mathbb{V}$. Поэтому каждый вектор также является представителем некоторого класса эквивалентности, и существует каноническое вложение

$$\varrho : \quad \mathbb{V} \hookrightarrow \mathfrak{cl}(\mathbb{V}, Q).$$

Тензоры второго ранга могут принадлежать одному классу эквивалентности. Например, поскольку идеал порожден элементами вида (26.6), то элементы

$$A \otimes B \in \mathbb{V}^2 \quad \text{и} \quad -B \otimes A + 2Q(A, B)e \in \mathbb{V}^2 \oplus \mathbb{V}^0$$

принадлежат одному классу эквивалентности.

Кроме того, элементы тензорного произведения $A \otimes A \in \mathbb{V}^2$ и $Q(A)e \in \mathbb{V}^0$ принадлежат одному классу эквивалентности, т.к. их разность принадлежит идеалу:

$$A \otimes A - Q(A)e \in \mathfrak{i}.$$

Эти классы эквивалентности параметризуются элементами поля $Q(A)e \in \mathbb{F}$.

Фактически, равенство (26.5) означает, что для конструктивного описания элементов алгебры Клиффорда необходимо произвести отождествление элементов тензорной алгебры, которые имеют вид

$$A \otimes B + B \otimes A - 2Q(A, B)e \quad \Leftrightarrow \quad A \otimes A - Q(A)e.$$

Каноническое отображение второго равенства в алгебру Клиффорда приводит к равенству $\varrho(A)^2 = Q(A)e$.

Предложение 26.1.3 (Универсальность). Пусть $\mathfrak{a}$ – ассоциативная алгебра с единицей $e_{\mathfrak{a}}$. Если отображение $f : \mathbb{V} \rightarrow \mathfrak{a}$ линейно и выполнено условие

$$f(A)^2 = Q(A)e_{\mathfrak{a}}, \quad \forall A \in \mathbb{V}, \quad (26.7)$$

то отображение f однозначно продолжается до $\mathbb{F}$ -гомоморфизма алгебр $F : \mathfrak{cl}(\mathbb{V}, Q) \rightarrow \mathfrak{a}$ со свойствами:

$$\begin{aligned} F(\alpha) &= \alpha e_{\mathfrak{a}}, & \forall \alpha \in \mathbb{F}, \\ F(A) &= f(A), & \forall A \in \mathbb{V}, \\ F(UV) &= F(U)F(V), \\ F(U + V) &= F(U) + F(V), & \forall U, V \in \mathfrak{cl}(\mathbb{V}, Q). \end{aligned}$$

Более того, алгебра Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$ – это единственная с точностью до изоморфизма ассоциативная алгебра над полем $\mathbb{F}$ с таким свойством.

Доказательство. Каждое линейное отображение $f : \mathbb{V} \rightarrow \mathfrak{a}$ продолжается до единственного гомоморфизма $\hat{f} : \otimes \mathbb{V} \rightarrow \mathfrak{a}$, который определен правилом $\hat{f}(A \otimes B) := f(A)f(B)$ и дальше по линейности. Из свойства (26.7) следует, что $\hat{f}(U) = 0$ для всех $U \in \mathfrak{i}$ (26.4). Поэтому отображение $\hat{f}$ сужается до хорошо определенного отображения F на алгебре Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$, которое имеет необходимые свойства.

Предположим теперь, что $\mathfrak{b}$ является ассоциативной алгеброй с единицей над полем $\mathbb{F}$ и отображение $i : \mathbb{V} \rightarrow \mathfrak{b}$ – вложение с тем свойством, что произвольное линейное отображение $f : \mathbb{V} \rightarrow \mathfrak{a}$ со свойством (26.7) продолжается единственным образом до гомоморфизма алгебр $F : \mathfrak{b} \rightarrow \mathfrak{a}$. Тогда изоморфизм из $\mathbb{V} \subset \mathfrak{cl}(\mathbb{V}, Q)$ в $i(\mathbb{V}) \subset \mathfrak{b}$ индуцирует изоморфизм алгебр $\mathfrak{cl}(\mathbb{V}, Q) \rightarrow \mathfrak{b}$. $\square$

Дадим еще одно эквивалентное определение алгебры Клиффорда.

Определение. Пусть задано векторное пространство $\mathbb{V}$ над полем $\mathbb{F}$ с симметричной квадратичной формой Q . Алгеброй Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$ называется ассоциативная алгебра с единицей $e \in \mathfrak{cl}(\mathbb{V}, Q)$ со следующими свойствами:

- 1) Существует такое $\mathbb{F}$ -линейное вложение $\varrho : \mathbb{V} \hookrightarrow \mathfrak{cl}(\mathbb{V}, Q)$, что для каждого элемента $A \in \mathbb{V}$ выполнено соотношение

$$\varrho(A)^2 = Q(A)e. \quad (26.8)$$

- 2) Элементы $\varrho(\mathbb{V})$ являются мультипликативными образующими, т.е. любой элемент из $\mathfrak{cl}(\mathbb{V}, Q)$ представим в виде конечной линейной комбинации одночленов от элементов $\varrho(A)$.
- 3) В алгебре Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$ не существует других тождеств, кроме тех, которые порождены равенством (26.8).

Если существует другая алгебра $\mathfrak{cl}'(\mathbb{V}, Q)$ и $\mathbb{F}$ линейное вложение $\varrho' : \mathbb{V} \rightarrow \mathfrak{cl}'(\mathbb{V}, Q)$ со свойствами 1)–3), то алгебры $\mathfrak{cl}(\mathbb{V}, Q)$ и $\mathfrak{cl}'(\mathbb{V}, Q)$ изоморфны. $\square$

Условие 3) в определении алгебры Клиффорда можно сформулировать другим, эквивалентным, образом: алгебра Клиффорда не порождается никаким собственным подпространством из $\mathbb{V}$.

Свойство (26.8) является основным и означает, что скалярный квадрат любого вектора $A \in \mathbb{V}$ реализуется в алгебре Клиффорда $\varrho(A) \in \mathfrak{cl}(\mathbb{V}, Q)$ в виде клиффордова умножения. Можно сказать, что в алгебре Клиффорда можно извлечь квадратный корень из произвольной билинейной формы. В разделе 27.1.5 мы увидим, что оператор Дирака для спинорного поля является “квадратным корнем” из уравнения Клейна–Гордона–Фока.

Теорема 26.1.1. Для всякого конечномерного векторного пространства $\mathbb{V}$ над полем $\mathbb{F}$ с симметричной квадратичной формой Q алгебра Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$ существует и имеет размерность 2^n над полем $\mathbb{F}$, где $n = \dim \mathbb{V}$.

Доказательство. Выберем в $\mathbb{V}$ ортонормированный базис e^a , $a = 1, \dots, n$. По определению,

$$[\varrho(e^a + e^b)]^2 = Q(e^a + e^b)e = [Q(e^a) + 2Q(e^a, e^b) + Q(e^b)]e,$$

где мы воспользовались формулой (26.3). С другой стороны, в силу линейности выполнены следующие соотношения

$$[\varrho(e^a + e^b)]^2 = [\varrho(e^a) + \varrho(e^b)]^2 = \varrho(e^a)^2 + \varrho(e^a)\varrho(e^b) + \varrho(e^b)\varrho(e^a) + \varrho(e^b)^2.$$

Сравнение двух последних выражений приводит к определяющему соотношению

$$\varrho(e^a)\varrho(e^b) + \varrho(e^b)\varrho(e^a) = 2Q(e^a, e^b)e = 2\eta^{ab}e. \quad (26.9)$$

Далее, произвольный элемент алгебры Клиффорда представим в виде линейной комбинации произведений $\varrho(A_1) \cdots \varrho(A_k) \in \mathfrak{cl}$, где все $A_i \in \mathbb{V}$, $i = 1, \dots, k$. Разложим векторы по базису $A_i = e^a A_{ia}$. Тогда в силу $\mathbb{F}$ -линейности и ассоциативности любой элемент алгебры Клиффорда представим в виде конечной линейной комбинации произведений базисных векторов $\varrho(e^1) \cdots \varrho(e^k) \in \mathfrak{cl}$. Из равенства (26.9) следует, что если два соседних базисных вектора совпадают, $e^i = e^{i+1}$, то произведение их образов дает элемент из поля $\mathbb{F}$, и его можно отнести к коэффициенту в линейной комбинации. Если же $e^i \neq e^{i+1}$, то образы можно упорядочить по возрастанию $i < i+1$, воспользовавшись их антикоммутативностью (26.9) при разных значениях индексов. Следовательно, любой элемент алгебры Клиффорда представим в виде линейной комбинации элементов

$$e, \quad \varrho(e^{a_1}), \quad \varrho(e^{a_1})\varrho(e^{a_2}), \quad \dots, \quad \varrho(e^{a_1}) \cdots \varrho(e^{a_k}), \quad \dots, \quad \varrho(e^{a_1}) \cdots \varrho(e^{a_n}), \quad (26.10)$$

где все наборы индексов упорядочены по возрастанию $1 \leq a_1 < \dots < a_k \leq n$.

Если в алгебре, порожденной векторами (26.10) существуют какие-либо дополнительные тождества, например, $\varrho(e^1)\varrho(e^2) = \varrho(e^3)$, то множество векторов (26.10) будет линейно зависимым. Поэтому свойство 3) гарантирует линейную независимость векторов (26.10). Следовательно, эти векторы образуют базис в алгебре Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$. Поскольку базисных векторов ровно 2^n штук, то $\dim \mathfrak{cl}(\mathbb{V}, Q) = 2^n$. Существование алгебры Клиффорда следует из построения.

Докажем единственность алгебры Клиффорда с точностью до изоморфизма. Пусть $\varrho' : \mathbb{V} \rightarrow \mathfrak{cl}'(\mathbb{V}, Q)$. Определим биективное отображение алгебр Клиффорда с помощью отображения базисных векторов:

$$\sigma : \mathfrak{cl}(\mathbb{V}, Q) \ni \varrho(e^{a_1, \dots, a_k}) \mapsto \sigma(\varrho(e^{a_1, \dots, a_k})) := \varrho'(e^{a_1, \dots, a_k}) \in \mathfrak{cl}(\mathbb{V}, Q),$$

т.е. $\varrho' = \sigma \circ \varrho$. Нетрудно проверить, что это отображение является гомоморфизмом алгебр. $\square$

Отображение ϱ в определении алгебры Клиффорда является каноническим и в дальнейшем рассмотрении мы его опустим, проведя отождествление

$$\varrho(A) = A, \quad \forall A \in \mathbb{V}.$$

Тогда основное равенство (26.8) примет вид

$$A^2 = Q(A)e. \quad (26.11)$$

Условие 3) в данном выше определении алгебры Клиффорда существенно в размерностях $p - q = 1, \bmod 4$, где $\eta^{ab} = \underbrace{+\dots+}_p \underbrace{-\dots-}_q$ – сигнатура метрики. В частности, его можно опустить в четных размерностях $p + q = n$ (см., например, [195], раздел 14.3). Ниже в примере 26.2.8 будет продемонстрирована его необходимость.

При проведении вычислений в алгебрах Клиффорда удобно пользоваться базисом, который построен в доказательстве теоремы 26.1.1. Поэтому приведем эквивалентное определение алгебры Клиффорда, привязанное к этому базису.

Определение. Пусть задано векторное пространство $\mathbb{V}$, $\dim \mathbb{V} = n$, над полем $\mathbb{F} = \mathbb{R}, \mathbb{C}$, с симметричной квадратичной формой $Q(A, B) = \eta^{ab} A_a B_b$. Будем считать, что в $\mathbb{V}$ задан ортонормированный базис e^a , $a = 1, \dots, n$, в котором квадратичная форма имеет канонический вид. Пусть также задано 2^n -мерное векторное пространство $\mathfrak{cl}(\mathbb{V}, Q)$ над тем же полем $\mathbb{F}$. Обозначим базис векторного пространства $\mathfrak{cl}(\mathbb{V}, Q)$ следующим образом

$$e, \quad e^a, \quad e^{a_1 a_2}, \quad e^{a_1 a_2 a_3}, \dots, \quad e^{a_1 a_2 \dots a_k}, \dots, \quad e^{12 \dots n}, \quad (26.12)$$

где $a, a_1, a_2, \dots = 1, \dots, n$ и $a_1 a_2 \dots a_k$ – упорядоченный набор индексов, $1 \leq a_1 < a_2 < \dots < a_k \leq n$. Произвольный элемент линейного пространства $\mathfrak{cl}(\mathbb{V}, Q)$ взаимно однозначно представим в виде линейной комбинации базисных векторов:

$$U = e u_0 + e^a u_a + e^{a_1 a_2} u_{a_1 a_2} + \dots + e^{a_1 \dots a_n} u_{a_1 \dots a_n} \in \mathfrak{cl}(\mathbb{V}, Q), \quad (26.13)$$

где компоненты $u_0, u_a, u_{a_1 a_2}, \dots$ – вещественные или комплексные числа. Векторное подпространство, натянутое на векторы e^a , отождествим с векторным пространством $\mathbb{V}$. Введем в векторном пространстве $\mathfrak{cl}(\mathbb{V}, Q)$ клиффордово умножение

$$\mathfrak{cl}(\mathbb{V}, Q) \times \mathfrak{cl}(\mathbb{V}, Q) \ni \quad U, V \mapsto UV \quad \in \mathfrak{cl}(\mathbb{V}, Q)$$

с помощью следующих правил:

- 1) $U(V + W) = UV + UW$
 $(U + V)W = UW + VW$ – дистрибутивность;
 - 2) $\alpha(UV) = (\alpha U)V = U(\alpha V)$ – умножение на скаляры;
 - 3) $(UV)W = U(VW) = UVW$ – ассоциативность;
 - 4) $Ue = eU = U$ – существование единицы;
 - 5) $e^a e^b + e^b e^a := 2\eta^{ab} e$;
 - 6) $e^{a_1} e^{a_2} \dots e^{a_k} := e^{a_1 a_2 \dots a_k}$
- (26.14)

для всех $U, V, W \in \mathfrak{cl}(\mathbb{V}, Q)$ и $\alpha, \beta \in \mathbb{F}$. Мы предполагаем, что свойство 5) выполнено для всех значений индексов $a, b = 1, \dots, n$, а свойство 6) – для всех упорядоченных наборов индексов $1 \leq a_1 < a_2 < \dots < a_k \leq n$.

Векторное пространство $\mathfrak{cl}(\mathbb{V}, Q)$ с введенной операцией умножения называется *алгеброй Клиффорда* над векторным пространством $\mathbb{V}$ с квадратичной формой Q . Базисные векторы e^a линейного пространства $\mathbb{V}$ называются *генераторами* или *образующими алгебры Клиффорда*. Эти образующие будем называть *ортонормированными*. □

Требования 1) и 2) говорят о том, что умножение в алгебре Клиффорда должно быть согласовано со структурой векторного пространства. Они совпадают с общим определением алгебры. Требование ассоциативности необходимо для корректного определения умножения в правиле 5). Таким образом, первые четыре требования – это общее определение ассоциативной алгебры с единицей (унитальной алгебры). Специфика алгебры Клиффорда содержится только в правилах 5) и 6).

Поскольку алгебра Клиффорда является векторным пространством, то мы предполагаем, что на ней задана обычная евклидова топология.

Первый базисный вектор e , по определению, является единицей алгебры Клиффорда, т.е. мы считаем, что поле вещественных или комплексных чисел является подпространством рассматриваемого векторного пространства $\mathfrak{cl}(\mathbb{V}, Q)$. В дальнейшем единицу e при записи элементов алгебры Клиффорда мы часто будем опускать.

Свойства 1)–6) в определении (26.14) позволяют перемножать произвольные элементы алгебры Клиффорда. Действительно, определяющее свойство

$$e^a e^b + e^b e^a = 2\eta^{ab} e, \quad (26.15)$$

позволяет упорядочить по возрастанию все возможные наборы индексов, поскольку образующие с различными индексами антикоммутируют.

Пример 26.1.1.

$$e^2 e^1 = -e^1 e^2 = -e^{12}. \quad \square$$

В случае одинаковых индексов справедливо равенство

$$(e^a)^2 = e^a e^a = \eta^{aa} e \quad (\text{суммирования нет}), \quad (26.16)$$

которое позволяет избавиться от индексов, встречающихся дважды.

Пример 26.1.2.

$$e^{12} e^{23} = e^1 e^2 e^2 e^3 = \eta^{22} e^{13}. \quad \square$$

Пример 26.1.3. Клиффордово произведение двух 1-форм $A = e^a A_a$ и $B = e^a B_a$ равно

$$AB = \sum_a e A_a B_a \eta^{aa} + \sum_{a < b} e^{ab} (A_a B_b - A_b B_a). \quad (26.17)$$

При этом скалярное произведение этих форм можно выразить через клиффордово умножение по формуле

$$(A, B) := A_a B_b \eta^{ab} = \frac{1}{2} (AB + BA),$$

где мы опустили единицу алгебры Клиффорда. Формулу для клиффордова произведения 1-форм (26.17) можно условно записать в виде

$$AB = (A, B) + A \wedge B,$$

если отождествить базисный вектор e^{ab} с внешним произведением $e^a \wedge e^b$ при $a < b$. $\square$

Из определения следует, что при $n = 1$ алгебра Клиффорда коммутативна, а при $n \geq 2$ – некоммутативна.

В дальнейшем мы увидим, что каждая алгебра Клиффорда изоморфна некоторой подалгебре в алгебре $m \times m$ матриц $\mathbf{mat}(m, \mathbb{F})$ с обычным умножением. Поэтому элемент алгебры Клиффорда можно представлять себе в виде матрицы, а клиффордово умножение – как обычное умножение матриц.

Для клиффордова умножения мы не будем вводить специального обозначения, потому что оно рассматривается только в настоящей главе. Кроме того, в приложениях встречаются, как правило, матричные представления алгебр Клиффорда, а для умножения матриц мы не используем специальных символов.

Пусть в исходном векторном пространстве задано подпространство $\mathbb{W} \subset \mathbb{V}$, натянутое на векторы e^i , $i = 1, \dots, m$, где $m < n$, тогда они удовлетворяют антикомму- тационным соотношениям

$$e^i e^j + e^j e^i = 2\eta^{ij}, \quad \eta^{ij} = \text{diag}(+\dots + -\dots - 0\dots 0).$$

где на диагонали матрицы η^{ij} стоит некоторое количество положительных и отрицательных единиц, а также нулей. Эта матрица задает квадратичную форму Q' на W , которая является сужением формы Q на подпространство $\mathbb{W}$. Отсюда следует, что векторы e^i сами по себе генерируют некоторую алгебру Клиффорда $\mathbf{cl}(\mathbb{W}, Q')$. Ясно, что эта алгебра является подалгеброй в исходной алгебре Клиффорда, $\mathbf{cl}(\mathbb{W}, Q') \subset \mathbf{cl}(\mathbb{V}, Q)$.

В одном и том же векторном пространстве $\mathbb{V}$ можно задать много симметричных билинейных форм. Каждой такой форме соответствует своя алгебра Клиффорда. Все такие алгебры Клиффорда изоморфны между собой как векторные пространства, но отличаются правилом умножения вектором. В дальнейшем мы увидим, что некоторые из этих алгебр Клиффорда изоморфны, несмотря на то, что сами билинейные формы различны.

В определении алгебр Клиффорда квадратичная форма Q может быть вырождена. В частном случае, если уравнение (26.15) заменить на перестановочное соотношение

$$e^a e^b + e^b e^a = 0, \quad (26.18)$$

соответствующее нулевой “метрике”, то получим *алгебру Грассмана* [51] или внешнюю алгебру, которая была подробно рассмотрена в разделе 3.1. То есть $\mathbf{cl}(\mathbb{V}, 0) = \Lambda(\mathbb{V})$.

По определению, каждое слагаемое в представлении (26.13) элементов алгебры Клиффорда является элементом векторного подпространства. То есть мы имеем разложение алгебры Клиффорда в прямую сумму

$$\mathbf{cl}(\mathbb{V}, Q) = \bigoplus_{r=0}^n \mathbf{cl}_r(\mathbb{V}, Q), \quad (26.19)$$

где элементы векторного пространства $\mathbf{cl}_r(\mathbb{V}, Q)$ имеют вид

$$U_r := e^{a_1 \dots a_r} u_{a_1 \dots a_r} \in \mathbf{cl}_r(\mathbb{V}, Q), \quad \dim \mathbf{cl}_r(\mathbb{V}, Q) = C_n^r.$$

То есть элементы алгебры Клиффорда $U_r \in \mathbf{cl}(\mathbb{V}, Q)$, степени r , разлагаются по базисным векторам алгебры Клиффорда с r индексами. Очевидно, что $\mathbf{cl}_0(\mathbb{V}, Q) = \mathbb{R}, \mathbb{C}$ и $\mathbf{cl}_1(\mathbb{V}, Q) = \mathbb{V}$.

Из определения следует, что все индексы базисного вектора $e^{a_1 \dots a_r}$ различны, и компоненты $u_{a_1 \dots a_r}$ антисимметричны относительно любой перестановки индексов.

Это позволяет использовать запись элементов алгебры Клиффорда, аналогичную записи элементов внешней алгебры. А именно, введем векторы

$$\gamma^{a_1 \dots a_r} := e^{[a_1} e^{a_2} \dots e^{a_r]}, \quad (26.20)$$

где квадратные скобки обозначают антисимметризацию по всем индексам. Эти векторы отличны от нуля только в том случае, если все индексы различны. Тогда элемент алгебры Клиффорда (26.13) запишется в виде суммы

$$U = \sum_{r=0}^n \frac{1}{r!} \gamma^{a_1 \dots a_r} u_{a_1 \dots a_r}, \quad (26.21)$$

где суммирование ведется по всем значениям индексов, а не только по упорядоченным.

Отсюда сразу следует, что алгебра Клиффорда $\mathbf{cl}(\mathbb{V}, Q)$, как векторное пространство, изоморфна внешней алгебре $\Lambda(\mathbb{V})$. Однако, они не изоморфны как алгебры при нетривиальной квадратичной форме Q .

Назвать элементы $\gamma^{a_1 \dots a_r}$ базисом векторного подпространства $\mathbf{cl}_r(\mathbb{V}, Q)$ нельзя, т.к. их существенно больше, чем нужно. Поэтому в формуле (26.21) в знаменателе возник факториал. Тем не менее запись (26.21) удобна и иногда упрощает вычисления.

Мы будем называть элементы подпространств $\mathbf{cl}_0(\mathbb{V}, Q)$ скалярами, $\mathbf{cl}_1(\mathbb{V}, Q)$ – векторами, $\mathbf{cl}_2(\mathbb{V}, Q)$ – бивекторами и т.д. Последний базисный вектор $e^{a_1 \dots a_n}$ назовем объемом.

Если квадратичная форма Q отлична от нуля, то умножение в алгебре Клиффорда существенно отличается от умножения во внешней алгебре. Для умножения элементов алгебры степеней $r \leq s$ справедливы включения:

$$\begin{aligned} \Lambda_r \times \Lambda_s \ni \quad A_r, A_s &\mapsto A_r \wedge A_s && \in \Lambda_{r+s}, \\ \mathbf{cl}_r \times \mathbf{cl}_s \ni \quad U_r, U_s &\mapsto U_r U_s && \in \mathbf{cl}_{s+r} \oplus \mathbf{cl}_{s+r-2} \dots \oplus \mathbf{cl}_{s-r}. \end{aligned} \quad (26.22)$$

Если все индексы элементов U_r и U_s различны, то $U_r U_s \in \mathbf{cl}_{r+s}$. Если совпадает один индекс, то $U_r U_s \in \mathbf{cl}_{r+s-2}$ и т.д. В предельном случае, когда все индексы элемента U_r содержатся в U_s , справедливо включение $U_r U_s \in \mathbf{cl}_{s-r}$.

Свойство (26.22) означает, что все алгебры Клиффорда являются фильтрованными. С каждой фильтрацией ассоциируется градуированная алгебра (см. раздел 1.3.3). Можно проверить, что ассоциированной градуированной алгеброй для алгебры Клиффорда $\mathbf{cl}(\mathbb{V}, Q)$ будет внешняя алгебра $\Lambda(\mathbb{V})$.

Из определения следует, что *четные элементы*

$$\mathbf{cl}_+ := \mathbf{cl}_0 \oplus \mathbf{cl}_2 \oplus \dots, \quad \dim \mathbf{cl}_+ = 2^{n-1}, \quad (26.23)$$

образуют подалгебру алгебры Клиффорда

$$\mathbf{cl}_+ \subset \mathbf{cl}.$$

Нечетные элементы алгебры Клиффорда

$$\mathbf{cl}_- := \mathbf{cl}_1 \oplus \mathbf{cl}_3 \oplus \dots, \quad \dim \mathbf{cl}_- = 2^{n-1}, \quad (26.24)$$

образуют только линейное подпространство, но не подалгебру в алгебре Клиффорда $\mathbf{cl}(\mathbb{V}, Q)$, т.к., например, не содержат единичного элемента.

Таким образом, существует разложение алгебры Клиффорда в прямую сумму

$$\mathbf{cl} = \mathbf{cl}_+ \oplus \mathbf{cl}_-$$

со свойством

$$\mathbf{cl}_i \mathbf{cl}_j \subseteq \mathbf{cl}_{i+j}, \quad \text{mod } 2,$$

где $i, j = (0, 1) = (+, -)$. Такие алгебры называются $\mathbb{Z}_2$ -градуированными.

Допустим, что квадратичная форма Q в векторном пространстве $\mathbb{V}$ не вырождена. Тогда векторное пространство $\mathbb{V}$ можно отождествить с (псевдо)евклидовым пространством $\mathbb{R}^{p,q}$ или $\mathbb{C}^{p,q}$, и рассматривать квадратичную форму в качестве метрики. Эта метрика в каноническом базисе представляет собой диагональную матрицу, у которой на диагонали стоят положительные и отрицательные единицы:

$$\eta^{ab} = \text{diag}(\underbrace{+\dots+}_p, \underbrace{-\dots-}_q), \quad p+q=n. \quad (26.25)$$

В этом случае для алгебры Клиффорда $\mathbf{cl}(\mathbb{V}, Q)$ мы будем использовать более содержательное обозначение $\mathbf{cl}(p, q, \mathbb{R})$ или $\mathbf{cl}(p, q, \mathbb{C})$.

Очевидно, что вещественные алгебры Клиффорда образуют подалгебру в комплексной алгебре Клиффорда для квадратичной формы той же сигнатуры,

$$\mathbf{cl}(p, q, \mathbb{R}) \subset \mathbf{cl}(p, q, \mathbb{C}).$$

Конечно, в комплексном случае замена $e^a \mapsto ie^a$ при $a = p+1, \dots, p+q$ отображает алгебру Клиффорда $\mathbf{cl}(p, q, \mathbb{C})$ в алгебру $\mathbf{cl}(p+q, 0, \mathbb{C})$. То есть все комплексные алгебры Клиффорда одинаковой размерности изоморфны.

Ортонормальный базис e^a подпространства $\mathbf{cl}_1 \subset \mathbf{cl}$, как и в векторном пространстве $\mathbb{V}$ над полем $\mathbb{F}$ с невырожденной квадратичной формой сигнатуры (p, q) определен с точностью до $\mathbb{O}(p, q, \mathbb{F})$ -вращений, где $\mathbb{F} = \mathbb{R}, \mathbb{C}$. Поэтому можно совершить (псевдо)вращение

$$e^a \mapsto e'^a = e^b S_b^a, \quad S \in \mathbb{O}(p, q, \mathbb{F}), \quad (26.26)$$

которое сохранит квадратичную форму. Элемент каждого подпространства $\mathbf{cl}_r$ можно записать в виде суммы

$$U_r = \frac{1}{r!} \gamma^{a_1 \dots a_r} u_{a_1 \dots a_r} \in \mathbf{cl}_r(p, q, \mathbb{F}). \quad (26.27)$$

Поскольку векторы $\gamma^{a_1 \dots a_r}$, по определению, антисимметричны относительно перестановки всех индексов (26.20), то все подпространства $\mathbf{cl}_r$ инвариантны относительно вращений (26.26):

$$\gamma^{a_1 \dots a_r} \mapsto \gamma^{b_1 \dots b_r} S_{b_1}^{a_1} \dots S_{b_r}^{a_r} \in \mathbf{cl}_r.$$

Перед рассмотрением примеров сделаем еще одно замечание, которое будет использовано в дальнейшем. Умножение в алгебре Клиффорда было определено с помощью квадратичной формы в выделенном подпространстве $\mathbb{V}$. Если умножение определено, то в алгебре Клиффорда, как в векторном пространстве, можно выбрать произвольный базис. То есть от базиса (26.12) перейти к произвольному базису:

$$\{e, e^a, e^{a_1 a_2}, \dots, e^{a_1 a_2 \dots a_r}, \dots, e^{12 \dots n}\} \mapsto \{e^A\}, \quad A = 1, \dots, 2^n.$$

Конечно, при этом алгебра Клиффорда останется алгеброй Клиффорда, которая будет изоморфна исходной. Однако образующие и подпространство $\mathbb{V}$ не будут явно выделены. В дальнейшем мы часто будем этим пользоваться при доказательстве изоморфизмов некоторых алгебр Клиффорда.

26.2 Алгебры Клиффорда низших размерностей

Рассмотрим примеры алгебр Клиффорда низших размерностей. Эти примеры важны для последующей классификации алгебр Клиффорда на основе рекуррентных соотношений.

Пример 26.2.1. По определению, положим $\mathfrak{cl}(0, 0, \mathbb{F}) \simeq \mathbb{F}$. $\square$

Пример 26.2.2. Алгебра Клиффорда $\mathfrak{cl}(1, 0, \mathbb{R})$ двумерна. Она имеет одну образующую e^1 и единичный элемент e . При этом выполнено равенство $(e^1)^2 = e$. Произвольный элемент имеет вид

$$U = eu_0 + e^1u_1 \in \mathfrak{cl}(1, 0, \mathbb{R}). \quad (26.28)$$

Клиффордово произведение двух элементов равно

$$UV = (eu_0 + e^1u_1)(ev_0 + e^1v_1) = e(u_0v_0 + u_1v_1) + e^1(u_0v_1 + u_1v_0).$$

Ясно, что оно коммутативно.

Отождествим алгебру Клиффорда $\mathfrak{cl}(1, 0, \mathbb{R})$ с плоскостью: $(u_0, u_1) \in \mathbb{R}^2$. Тогда клиффордово умножение двух точек на плоскости дает новую точку с координатами

$$(u_0, u_1)(v_0, v_1) = (u_0v_0 + u_1v_1, u_0v_1 + u_1v_0).$$

При этом единичному элементу и единственной образующей соответствуют точки $e = (1, 0)$ и $e^1 = (0, 1)$.

На алгебру Клиффорда $\mathfrak{cl}(1, 0, \mathbb{R})$ можно взглянуть с другой точки зрения. Пусть задано двумерное векторное пространство, т.е. плоскость, которую можно представить в виде прямой суммы $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$. Введем на плоскости коммутативное и ассоциативное умножение

$$\mathbb{R}^2 \times \mathbb{R}^2 \ni (a_1 \oplus b_1), (a_2 \oplus b_2) \mapsto (a_1 \oplus b_1)(a_2 \oplus b_2) := (a_1a_2 \oplus b_1b_2) \in \mathbb{R}^2. \quad (26.29)$$

Отождествим единицу e и единственный образующей элемент алгебры Клиффорда e^1 с точками плоскости:

$$\begin{aligned} e &= (1, 1) \in \mathbb{R}^2 && \text{— скаляр,} \\ e^1 &= (1, -1) \in \mathbb{R}^2 && \text{— вектор (он же объем).} \end{aligned}$$

Тогда нетрудно проверить равенства $ee = e$, $ee^1 = e^1e = e^1$, $(e^1)^2 = e$. Это значит, что алгебра Клиффорда $\mathfrak{cl}(1, 0, \mathbb{R})$ изоморфна прямой сумме двух вещественных прямых $\mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2$, рассматриваемых как алгебры, с умножением (26.29). Мы пишем

$$\mathfrak{cl}(1, 0, \mathbb{R}) \simeq \mathbb{R} \oplus \mathbb{R}. \quad (26.30)$$

Во втором определении алгебры Клиффорда на с. 1030 есть третье условие об отсутствии дополнительных тождеств. Если наложить дополнительное соотношение $e^1 = -e$, то получим одномерную алгебру, изоморфную $\mathbb{R}$. Это говорит о необходимости условия 3) в определении.

Алгебра Клиффорда $\mathfrak{cl}(1, 0, \mathbb{C})$ над полем комплексных чисел имеет комплексную размерность два, $\dim_{\mathbb{C}} [\mathfrak{cl}(1, 0, \mathbb{C})] = 2$, и те же образующие, что и в вещественном случае. Отличие сводится к тому, что компоненты u_0 и u_1 в разложении (26.28) теперь являются комплексными числами, и их нельзя рассматривать в качестве декартовых координат на евклидовой плоскости. Аналогично вещественному случаю существует изоморфизм

$$\mathfrak{cl}(1, 0, \mathbb{C}) \simeq \mathbb{C} \oplus \mathbb{C}. \quad \square \quad (26.31)$$

Пример 26.2.3. Алгебра Клиффорда $\mathfrak{cl}(0, 1, \mathbb{R})$ с одной образующей e^1 , удовлетворяющей соотношению $(e^1)^2 = -1$, изоморфна полю комплексных чисел $\mathbb{C}$. Действительно, клиффордово произведение двух элементов имеет вид

$$UV = (eu_0 + e^1u_1)(ev_0 + e^1v_1) = e(u_0v_0 - u_1v_1) + e^1(u_0v_1 + u_1v_0).$$

Если отождествить $i := e^1$ и ввести комплексные числа $z := u_0 + iu_1$, то клиффордово умножение совпадает с обычным умножением комплексных чисел. Таким образом, доказан следующий изоморфизм

$$\mathfrak{cl}(0, 1, \mathbb{R}) \simeq \mathbb{C}. \quad (26.32)$$

Теперь рассмотрим комплексную одномерную алгебру Клиффорда $\mathfrak{cl}(0, 1, \mathbb{C})$. Введем в $\mathbb{C} \oplus \mathbb{C} = \mathbb{C}^2$ коммутативное и ассоциативное умножение:

$$\mathbb{C}^2 \times \mathbb{C}^2 \ni (a_1 \oplus b_1), (a_2 \oplus b_2) \mapsto (a_1 \oplus b_1)(a_2 \oplus b_2) := (a_1a_2 \oplus b_1b_2) \in \mathbb{C}^2. \quad (26.33)$$

Положим $e := (1, 1)$ и $e^1 := (i, -i)$. Тогда справедлива таблица умножения

$$ee = e, \quad ee^1 = e^1e = e, \quad (e^1)^2 = -e.$$

Следовательно, алгебра Клиффорда над полем комплексных чисел изоморфна прямой сумме

$$\mathfrak{cl}(0, 1, \mathbb{C}) \simeq \mathbb{C} \oplus \mathbb{C}. \quad (26.34)$$

То есть $\mathfrak{cl}(0, 1, \mathbb{C}) \simeq \mathfrak{cl}(1, 0, \mathbb{C})$ в соответствии с изоморфизмом (26.31), как и должно быть согласно предложению 26.1.2. $\square$

В следующих примерах одновременно рассматриваются и вещественные, и комплексные алгебры Клиффорда. Для упрощения обозначений будем использовать одну букву $\mathbb{F} = (\mathbb{R} \text{ или } \mathbb{C})$.

Пример 26.2.4. Алгебра Клиффорда $\mathfrak{cl}(2, 0, \mathbb{F})$ четырехмерна. Ее элементы имеют вид

$$U = eu_0 + e^1u_1 + e^2u_2 + e^{12}u_{12}.$$

В случае $\mathbb{F} = \mathbb{R}$ компоненты u_0, u_1, u_2 и u_{12} являются вещественными числами, а в случае $\mathbb{F} = \mathbb{C}$ – комплексными. Таблица умножения базисных векторов выглядит следующим образом:

	e	e^1	e^2	e^{12}
e	e	e^1	e^2	e^{12}
e^1	e^1	e	e^{12}	e^2
e^2	e^2	$-e^{12}$	e	$-e^1$
e^{12}	e^{12}	$-e^2$	e^1	$-e$

Квадрат элемента объема $v := e^1e^2$ равен $v^2 = -1$. Поэтому подалгебра в $\mathfrak{cl}(2, 0, \mathbb{R})$, натянутая на векторы e, v , изоморфна алгебре комплексных чисел $\mathbb{C}$.

Произвольный вектор можно записать в полярной форме

$$A = e^1A_1 + e^2A_2 = e^1(A_1 + vA_2) = |A|e^1e^{iv\varphi},$$

где $|A| := \sqrt{A_1^2 + A_2^2}$ и $\varphi := \operatorname{arctg} A_2/A_1$.

Нетрудно проверить, что базисные векторы алгебры Клиффорда $\mathfrak{cl}(2, 0, \mathbb{F})$ имеют точное представление в алгебре вещественных 2×2 матриц $\mathbf{mat}(2, \mathbb{R})$:

$$\begin{aligned} e &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} && \text{— скаляр,} \\ e^1 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e^2 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} && \text{— вектор,} \\ e^{12} &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && \text{— объем.} \end{aligned} \quad (26.35)$$

Совокупность матриц e, e^1, e^2, e^{12} образует базис в пространстве 2×2 матриц. Поэтому между элементами алгебры Клиффорда и вещественными (комплексными) 2×2 матрицами существует взаимно однозначное соответствие. При этом клиффово умножение соответствует обычному умножению матриц. Следовательно, существуют изоморфизмы:

$$\mathfrak{cl}(2, 0, \mathbb{R}) \simeq \mathbf{mat}(2, \mathbb{R}), \quad \mathfrak{cl}(2, 0, \mathbb{C}) \simeq \mathbf{mat}(2, \mathbb{C}). \quad (26.36)$$

Отметим, что вещественная алгебра Клиффорда $\mathfrak{cl}(2, 0, \mathbb{R})$ имеет более богатую структуру, чем алгебра $\mathbf{mat}(2, \mathbb{R})$. Это связано с тем, что в алгебре Клиффорда выделено двумерное подпространство, натянутое на векторы e^1 и e^2 , которое вместе с евклидовой квадратичной формой изоморфно евклидовой плоскости $\mathbb{R}^2$.

Алгебра Клиффорда $\mathfrak{cl}(2, 0, \mathbb{R})$ позволяет “извлечь квадратный корень” из двумерного оператора Лапласа. А именно, справедливо представление

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \left(e^1 \frac{\partial}{\partial x} + e^2 \frac{\partial}{\partial y} \right)^2.$$

Если образующие алгебры Клиффорда e^1 и e^2 заданы матрицами (26.35), то дифференциальный оператор первого порядка, действуя на функцию-столбец

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\partial}{\partial x} + \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \frac{\partial}{\partial y} \right] \begin{pmatrix} u \\ v \end{pmatrix} = 0,$$

приводит к уравнениям Коши-Римана:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

□

Пример 26.2.5. Алгебра Клиффорда $\mathfrak{cl}(1, 1, \mathbb{F})$ четырехмерна и, так же, как и для алгебры $\mathfrak{cl}(2, 0, \mathbb{F})$, ее элементы имеют точное представление в виде 2×2 матриц:

$$\begin{aligned} e &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} && \text{— скаляр,} \\ e^1 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && \text{— вектор,} \\ e^{12} &= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} && \text{— объем.} \end{aligned}$$

В вещественном случае это представление алгебры Клиффорда отличается от соответствующего представления алгебры $\mathbf{cl}(2, 0, \mathbb{R})$ только тем, что выделенное двумерное подпространство, натянутое на векторы e^1 и e^2 , изоморфно пространству Минковского $\mathbb{R}^{1,1}$.

Сравнение представлений алгебр Клиффорда $\mathbf{cl}(2, 0, \mathbb{F})$ и $\mathbf{cl}(1, 1, \mathbb{F})$ говорит о том, что эти алгебры изоморфны. Действительно, изоморфизм задается отображением

$$e \mapsto e, \quad e^1 \mapsto e^1, \quad e^2 \mapsto e^{12}, \quad e^{12} \mapsto e^1.$$

В этом изоморфизме нетрудно убедиться и с помощью таблицы умножения, не используя явное представление.

Поскольку алгебры Клиффорда $\mathbf{cl}(2, 0, \mathbb{F})$ изоморфны матричным алгебрам (26.36), то имеют место также следующие изоморфизмы:

$$\mathbf{cl}(1, 1, \mathbb{R}) \simeq \mathbf{mat}(2, \mathbb{R}), \quad \mathbf{cl}(1, 1, \mathbb{C}) \simeq \mathbf{mat}(2, \mathbb{C}). \quad \square \quad (26.37)$$

Пример 26.2.6. Вещественная алгебра Клиффорда $\mathbf{cl}(0, 2, \mathbb{R})$ имеет две образующие e^1, e^2 , удовлетворяющими соотношениям: $(e^1)^2 = (e^2)^2 = -1$. Если отождествить $e^1 = i$, $e^2 = j$ и $e^{12} = k$, то таблица умножения этой алгебры совпадает с таблицей умножения кватернионов (см. приложение 37.3). Таким образом, вещественная алгебра Клиффорда $\mathbf{cl}(0, 2, \mathbb{R})$ изоморфна алгебре кватернионов,

$$\mathbf{cl}(0, 2, \mathbb{R}) \simeq \mathbb{H}. \quad (26.38)$$

Алгебра Клиффорда $\mathbf{cl}(0, 2, \mathbb{R})$ не имеет представления в алгебре вещественных 2×2 матриц, однако она имеет точное представление в алгебре комплексных 2×2 матриц $\mathbf{mat}(2, \mathbb{C})$:

$$e = \mathbb{1}, \quad e^1 = i\sigma^3, \quad e^2 = i\sigma^2, \quad e^{12} = i\sigma^1, \quad (26.39)$$

где $\sigma^{1,2,3}$ – матрицы Паули.

В разложении элементов алгебры Клиффорда $\mathbf{cl}(0, 2, \mathbb{R})$ по базису (26.39), $U = eu_0 + e^1u_1 + e^2u_2 + e^{12}u_{12}$, все коэффициенты разложения u_0, u_1, u_2, u_{12} – вещественные числа. Поэтому алгебра Клиффорда $\mathbf{cl}(0, 2, \mathbb{R})$ не изоморфна алгебре комплексных матриц $\mathbf{mat}(2, \mathbb{C})$.

Для алгебры Клиффорда $\mathbf{cl}(0, 2, \mathbb{C})$ над полем комплексных чисел коэффициенты u_0, u_1, u_2, u_{12} являются комплексными числами. Поэтому в соответствии с общим утверждением о том, что комплексные алгебры Клиффорда при разной сигнатуре метрики изоморфны, справедливы изоморфизмы

$$\mathbf{cl}(2, 0, \mathbb{C}) \simeq \mathbf{cl}(1, 1, \mathbb{C}) \simeq \mathbf{cl}(0, 2, \mathbb{C}) \simeq \mathbf{mat}(2, \mathbb{C}). \quad \square \quad (26.40)$$

Рассмотренные выше три примера показывают, что существует всего три вещественные алгебры Клиффорда с двумя образующими, из которых две являются изоморфными $\mathbf{cl}(2, 0, \mathbb{R}) \simeq \mathbf{cl}(1, 1, \mathbb{R})$. В то же время все три комплексные алгебры Клиффорда с двумя образующими изоморфны между собой (26.40).

Пример 26.2.7. Алгебра Клиффорда $\mathbf{cl}(3, 0, \mathbb{F})$ восьмимерна. Произвольный элемент алгебры имеет вид

$$U = eu_0 + e^1u_1 + e^2u_2 + e^3u_3 + e^{12}u_{12} + e^{23}u_{23} + e^{13}u_{13} + e^{123}u_{123}. \quad (26.41)$$

Квадрат элемента объема $v := e^{123}$ равен $v^2 = -e$. Поэтому подалгебра, натянутая на векторы e, v изоморфна алгебре комплексных чисел. Более того, в алгебре Клиффорда $\mathfrak{cl}(3, 0, \mathbb{F})$ существует бесконечно много подалгебр, изоморфных $\mathbb{C}$, т.к. для любого вектора $A = e^1 A_1 + e^2 A_2 + e^3 A_3$ на единичной сфере $A_1^2 + A_2^2 + A_3^2 = 1$ квадрат элемента

$$Av = e^{23} A_1 - e^{13} A_2 + e^{12} A_3$$

также равен $(Av)^2 = AvAv = A^2 v^2 = -e$.

Произвольный элемент алгебры Клиффорда (26.41) можно параметризовать с помощью двух скаляров $\alpha, \beta \in \mathbb{F}$ и двух векторов $A, B \in \mathbb{V}$:

$$U = \alpha + A + Bv + \beta v. \quad (26.42)$$

Точнее, β и B являются, соответственно, псевдоскаляром и псевдовектором, т.к. меняют знак при изменении ориентации (например $e^1 \mapsto -e^1$).

Базисные векторы алгебры Клиффорда $\mathfrak{cl}(3, 0, \mathbb{F})$ имеют точное представление в алгебре комплексных 2×2 -матриц $\mathbf{mat}(2, \mathbb{C})$:

$$\begin{aligned} e &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} && \text{— скаляр,} \\ e^1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad e^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad e^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} && \text{— вектор,} \\ e^{12} &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad e^{13} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad e^{23} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} && \text{— бивектор,} \\ e^{123} &= \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} && \text{— объем,} \end{aligned}$$

у которой образующие e^1, e^2, e^3 совпадают с матрицами Паули σ_1, σ_2 и σ_3 . Квадрат элемента объема равен $(e^{123})^2 = -1$. Элемент объема коммутирует со всеми матрицами и, следовательно, принадлежит центру алгебры.

Матрицы e, e^a, e^{ab}, e^{abc} образуют базис 2×2 комплексных матриц: любую комплексную 2×2 матрицу можно представить в виде суммы $U = eu_0 + e^a u_a + e^{ab} u_{ab} + e^{abc} u_{abc}$ с вещественными коэффициентами u_0, u_a, u_{ab} и u_{abc} . Поэтому имеет место изоморфизм

$$\mathfrak{cl}(3, 0, \mathbb{R}) \simeq \mathbf{mat}(2, \mathbb{C}). \quad (26.43)$$

В разделе 26.3 мы докажем изоморфизм $\mathfrak{cl}(3, 0, \mathbb{F}) \simeq \mathfrak{cl}(1, 2, \mathbb{F})$.

В комплексном случае коэффициенты разложения u_0, u_a, u_{ab} и u_{abc} являются комплексными числами. Поэтому имеет место изоморфизм

$$\mathfrak{cl}(3, 0, \mathbb{C}) \simeq \mathbf{mat}(2, \mathbb{C}) \oplus \mathbf{mat}(2, \mathbb{C}). \quad \square \quad (26.44)$$

Пример 26.2.8. Рассмотрим алгебру Клиффорда $\mathfrak{cl}(0, 3, \mathbb{R})$. Эта алгебра имеет три образующих и восьмимерна. Любой вектор из алгебры Клиффорда $\mathfrak{cl}(0, 3, \mathbb{R})$ имеет вид

$$eu_0 + e^1 u_1 + e^2 u_2 + e^3 u_3 + e^{12} u_{12} + e^{13} u_{13} + e^{23} u_{23} + e^{123} u_{123}. \quad (26.45)$$

Рассмотрим теперь прямую сумму двух алгебр кватернионов $\mathbb{H} \oplus \mathbb{H}$. Это – восьмимерное линейное пространство над полем вещественных чисел. Пусть $(q_1 \oplus q_2) \in \mathbb{H} \oplus \mathbb{H}$. Введем в прямой сумме $\mathbb{H} \oplus \mathbb{H}$ ассоциативное покомпонентное умножение:

$$(\mathbb{H} \oplus \mathbb{H}) \times (\mathbb{H} \oplus \mathbb{H}) \ni (q_1 \oplus q_2), (q'_1 \oplus q'_2) \mapsto (q_1 q'_1 \oplus q_2 q'_2) \in \mathbb{H} \oplus \mathbb{H}, \quad (26.46)$$

и отождествим

$$e = 1 \oplus 1, \quad e^1 = i \ominus i, \quad e^2 = j \ominus j, \quad e^3 = k \ominus k,$$

где i, j, k – образующие алгебры кватернионов (см. приложение 37.3) и $i \ominus i := i \oplus (-i)$. Тогда справедливы равенства

$$e^{12} = k \oplus k, \quad e^{13} = -j \ominus j, \quad e^{23} = i \oplus i, \quad e^{123} = -1 \oplus 1.$$

Каждый элемент алгебры Клиффорда (26.45) взаимно однозначно определяет два кватерниона:

$$\begin{aligned} q_1 &= (u_0 - u_{123}) + i(u_1 + u_{23}) + j(u_2 - u_{13}) + k(u_3 + u_{12}), \\ q_2 &= (u_0 + u_{123}) + i(-u_1 + u_{23}) + j(-u_2 - u_{13}) + k(-u_3 + u_{12}). \end{aligned}$$

при этом умножение в алгебре Клиффорда переходит в умножение (26.46). Следовательно существует изоморфизм

$$\mathfrak{cl}(0, 3, \mathbb{R}) \simeq \mathbb{H} \oplus \mathbb{H}. \quad (26.47)$$

В алгебре Клиффорда $\mathfrak{cl}(0, 3, \mathbb{R})$ можно ввести дополнительные тождества. А именно, элементы вида $(0 \oplus q) \in \mathbb{H} \oplus \mathbb{H}$ образуют в алгебре Клиффорда двусторонний идеал. Поэтому определена фактор алгебра

$$\frac{\mathbb{H} \oplus \mathbb{H}}{\mathbb{H}} \simeq \mathbb{H}.$$

В этой фактор алгебре мы отождествляем те элементы алгебры Клиффорда, которые отличаются только на элементы вида $(0 \oplus q)$. Другими словами, элементы алгебры Клиффорда $\mathfrak{cl}(0, 3, \mathbb{R})$ проектируются на первое слагаемое в прямой сумме (26.47). Это означает, что базисные векторы из $\mathbb{V}$ отображаются в векторы

$$e = 1, \quad e^1 = i, \quad e^2 = j, \quad e^3 = k.$$

Легко проверить, что определяющие соотношения (26.15) для алгебры Клиффорда $\mathfrak{cl}(0, 3, \mathbb{R})$ выполнены. Теперь получим выражения для остальных базисных векторов:

$$e^{12} = k, \quad e^{13} = -j, \quad e^{23} = i, \quad e^{123} = -1.$$

Отсюда вытекает, что в данном случае векторы $e, e^a, e^{a_1 a_2}, e^{123}$ не образуют базиса в алгебре Клиффорда, т.к. линейно зависимы. Это доказывает необходимость условия 3) во втором определении алгебры Клиффорда на с. 1030.

В комплексном случае алгебра кватернионов $\mathbb{H}$ в разложении (26.47) заменяется на алгебру комплексных матриц $\mathbf{mat}(2, \mathbb{C})$, и возникает изоморфизм

$$\mathfrak{cl}(0, 3, \mathbb{C}) \simeq \mathbf{mat}(2, \mathbb{C}) \oplus \mathbf{mat}(2, \mathbb{C})$$

в соответствии с общим правилом. □

Пример 26.2.9. Построим матричное представление для шестнадцатимерной алгебры Клиффорда $\mathfrak{cl}(4, 0, \mathbb{R})$. Пусть задана алгебра кватернионов $\mathbb{H}$ с образующими i, j, k (см. приложение 37.3). Отождествим образующие e^1, e^2, e^3 и e^4 с 2×2 матрицами с элементами из алгебры кватернионов:

$$e^1 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad e^2 := \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}, \quad e^3 := \begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix}, \quad e^4 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Нетрудно проверить, что в данном представлении определяющее соотношение (26.15) для образующих алгебры Клиффорда $\mathfrak{cl}(4, 0, \mathbb{R})$ выполнено. Для остальных базисных векторов получаем следующие матрицы представления

$$\begin{aligned} e^{12} &= \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}, & e^{13} &= \begin{pmatrix} j & 0 \\ 0 & j \end{pmatrix}, & e^{14} &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \\ e^{23} &= \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, & e^{24} &= \begin{pmatrix} 0 & j \\ j & 0 \end{pmatrix}, & e^{34} &= \begin{pmatrix} 0 & k \\ k & 0 \end{pmatrix}, \\ e^{123} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, & e^{124} &= \begin{pmatrix} k & 0 \\ 0 & -k \end{pmatrix}, & e^{134} &= \begin{pmatrix} j & 0 \\ 0 & -j \end{pmatrix}, & e^{234} &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \\ e^{1234} &= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \end{aligned}$$

Отсюда следует, что любую 2×2 матрицу с элементами из алгебры кватернионов можно представить в виде суммы

$$\mathbb{1}u_0 + e^a u_a + e^{ab} u_{ab} + e^{abc} u_{abc} + e^{abcd} u_{abcd}$$

с вещественными коэффициентами, где суммирование производится только по упорядоченным наборам индексов. При этом клиффордово умножение совпадает с матричным. Таким образом, существует изоморфизм

$$\mathfrak{cl}(4, 0, \mathbb{R}) \simeq \mathbf{mat}(2, \mathbb{H}). \quad \square \quad (26.48)$$

Пример 26.2.10. Рассмотрим алгебру Клиффорда $\mathfrak{cl}(3, 1, \mathbb{R})$. Она шестнадцатимерна и имеет четыре образующих. Для образующих этой алгебры существует представление в алгебре вещественных 4×4 матриц:

$$e^1 = \begin{pmatrix} \sigma^3 & 0 \\ 0 & -\sigma^3 \end{pmatrix}, \quad e^2 = \begin{pmatrix} \sigma^1 & 0 \\ 0 & \sigma^1 \end{pmatrix}, \quad e^3 = \begin{pmatrix} 0 & \sigma^3 \\ \sigma^3 & 0 \end{pmatrix}, \quad e^4 = \begin{pmatrix} i\sigma^2 & 0 \\ 0 & i\sigma^2 \end{pmatrix}.$$

где $\sigma^i = \sigma_i$, $i = 1, 2, 3$, – матрицы Паули. Как и в предыдущем примере можно вычислить остальные базисные векторы e^{ab} , e^{abc} и e^{abcd} и убедиться, что вместе с единицей они образуют базис вещественных 4×4 матриц. Поэтому существует изоморфизм

$$\mathfrak{cl}(3, 1, \mathbb{R}) \simeq \mathbf{mat}(4, \mathbb{R}). \quad (26.49)$$

В разделе 26.3 мы докажем изоморфизм алгебр Клиффорда $\mathfrak{cl}(3, 1, \mathbb{F}) \simeq \mathfrak{cl}(2, 2, \mathbb{F})$.

Широко используется также следующее представление образующих алгебры Клиффорда $\mathfrak{cl}(3, 1, \mathbb{F})$ в алгебре комплексных 4×4 матриц:

$$e^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}, \quad e^4 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \quad \square$$

Пример 26.2.11. Алгебра Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$ шестнадцатимерна. Ее образующие имеют точное представление в алгебре комплексных 4×4 матриц $\mathbf{mat}(4, \mathbb{C})$. Ее генераторами являются γ -матрицы Дирака с определяющим равенством:

$$\gamma^a \gamma^b + \gamma^b \gamma^a = 2\eta^{ab}, \quad \eta^{ab} = \text{diag}(+ - - -). \quad (26.50)$$

Здесь мы использовали общепринятое обозначение образующих алгебры Клиффорда: $e^a \mapsto \gamma^a$. Представление образующих можно выбрать в алгебре комплексных 4×4 матриц $\mathbf{mat}(4, \mathbb{C})$ в следующем блочном виде (представление Дирака):

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad (26.51)$$

где $\{\gamma^a\} = \{\gamma^0, \gamma^i\}$, и $\sigma^i := \sigma_i$, $i = 1, 2, 3$, – матрицы Паули.

Матрицы Дирака можно также представить в виде 2×2 матриц над алгеброй кватернионов $\mathbf{mat}(2, \mathbb{H})$:

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & j \\ j & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & k \\ k & 0 \end{pmatrix},$$

где i, j, k – базисные векторы кватернионов (см. приложение 37.3). Если вычислить остальные базисные векторы γ^{ab} , γ^{abc} и γ^{abcd} , то можно убедиться, что вместе с единицей они образуют базис в алгебре 2×2 матриц над алгеброй кватернионов. Следовательно, имеет место изоморфизм

$$\mathbf{cl}(1, 3, \mathbb{R}) \simeq \mathbf{mat}(2, \mathbb{H}). \quad (26.52)$$

Сравнение полученного результата с изоморфизмом (26.48) доказывает изоморфизм алгебр Клиффорда $\mathbf{cl}(1, 3, \mathbb{R}) \simeq \mathbf{cl}(4, 0, \mathbb{R})$.

Заметим, что алгебры Клиффорда $\mathbf{cl}(3, 1, \mathbb{R})$ и $\mathbf{cl}(1, 3, \mathbb{R})$ не изоморфны, хотя и отличаются только общим знаком метрики. Это значит, что спиноры различают многообразие с метриками сигнатуры $(+ - - -)$ и $(- + + +)$. $\square$

Конструкция алгебр Клиффорда относительно сложна, но содержательна. Рассмотренные примеры показывают, что многие интересные с точки зрения приложений алгебры вкладываются в эту конструкцию.

26.3 Классификация алгебр Клиффорда

Как и ранее, $\mathbb{F}$ обозначает либо поле вещественных $\mathbb{R}$, либо комплексных $\mathbb{C}$ чисел.

В настоящем разделе мы докажем путем явного построения, что каждая алгебра Клиффорда изоморфна некоторой матричной алгебре. При этом мы часто будем использовать понятие прямой суммы и тензорного произведения алгебр. Напомним

Определение. Рассмотрим две алгебры $\mathfrak{a}$ и $\mathfrak{b}$ над одним полем $\mathbb{F}$. Поскольку они являются векторными пространствами, то для них определено понятие прямой суммы $\mathfrak{a} \oplus \mathfrak{b}$ и тензорного произведения $\mathfrak{a} \otimes \mathfrak{b} = \mathfrak{a} \otimes_{\mathbb{F}} \mathfrak{b}$ (см. раздел 1.3.3). Определим умножение разложимых элементов в прямой сумме и тензорном произведении алгебр покомпонентно:

$$(a_1 \oplus b_1)(a_2 \oplus b_2) := (a_1 a_2) \oplus (b_1 b_2), \quad (26.53)$$

$$(a_1 \otimes b_1)(a_2 \otimes b_2) := (a_1 a_2) \otimes (b_1 b_2), \quad (26.54)$$

где $a_1, a_2 \in \mathfrak{a}$ и $b_1, b_2 \in \mathfrak{b}$ – произвольные элементы первой и второй алгебры, и продолжим его по линейности. $\square$

Легко проверить, что векторные пространства с введенным выше произведением являются алгебрами. Они имеют следующие размерности:

$$\dim_{\mathbb{F}}(\mathfrak{a} \oplus \mathfrak{b}) = \dim_{\mathbb{F}} \mathfrak{a} + \dim_{\mathbb{F}} \mathfrak{b}, \quad \dim_{\mathbb{F}}(\mathfrak{a} \otimes \mathfrak{b}) = \dim_{\mathbb{F}} \mathfrak{a} \times \dim_{\mathbb{F}} \mathfrak{b}.$$

Если в алгебрах существуют единицы, то в прямой сумме и тензорном произведении алгебр также существуют единичные элементы, которые имеют вид $e \oplus e$ и $e \otimes e$. Кроме того,

$$\begin{aligned} (a \oplus e) \cap (e \oplus b) &= \mathbb{F}, \\ (a \otimes e) \cap (e \otimes b) &= \mathbb{F}. \end{aligned}$$

Очевидно, что элементы алгебр $\mathfrak{a}$ и $\mathfrak{b}$ коммутируют:

$$\begin{aligned} (a \oplus e)(e \oplus b) &= (e \oplus b)(a \oplus e), \\ (a \otimes e)(e \otimes b) &= (e \otimes b)(a \otimes e). \end{aligned}$$

Предложение 26.3.1. *Прямая сумма и тензорное произведение алгебр ассоциативны:*

$$\begin{aligned} (\mathfrak{a} \oplus \mathfrak{b}) \oplus \mathfrak{c} &= \mathfrak{a} \oplus (\mathfrak{b} \oplus \mathfrak{c}) := \mathfrak{a} \oplus \mathfrak{b} \oplus \mathfrak{c}, \\ (\mathfrak{a} \otimes \mathfrak{b}) \otimes \mathfrak{c} &= \mathfrak{a} \otimes (\mathfrak{b} \otimes \mathfrak{c}) := \mathfrak{a} \otimes \mathfrak{b} \otimes \mathfrak{c}. \end{aligned}$$

Для трех алгебр $\mathfrak{a}$, $\mathfrak{b}$ и $\mathfrak{c}$ выполнены условия дистрибутивности:

$$\begin{aligned} (\mathfrak{a} \oplus \mathfrak{b}) \otimes \mathfrak{c} &= (\mathfrak{a} \otimes \mathfrak{c}) \oplus (\mathfrak{b} \otimes \mathfrak{c}), \\ \mathfrak{c} \otimes (\mathfrak{a} \oplus \mathfrak{b}) &= (\mathfrak{c} \otimes \mathfrak{a}) \oplus (\mathfrak{c} \otimes \mathfrak{b}). \end{aligned} \tag{26.55}$$

Доказательство. Прямая проверка. □

Пример 26.3.1. Рассмотрим поле вещественных чисел $\mathbb{R}$ как алгебру. Любое вещественное число представимо в виде $a = a \cdot 1 \in \mathbb{R}$, где 1 – базисный вектор. Тензорное произведение двух прямых одномерно, т.к. при тензорном произведении размерности перемножаются. В этом произведении определено умножение

$$(\mathbb{R} \otimes \mathbb{R}) \times (\mathbb{R} \otimes \mathbb{R}) \ni (a_1 \otimes b_1), (a_2 \otimes b_2) \mapsto (a_1 a_2 \otimes b_1 b_2) = a_1 a_2 b_1 b_2 \cdot \mathbb{1} \in \mathbb{R},$$

где $\mathbb{1} := 1 \otimes 1$. Таким образом, мы доказали изоморфизм $\mathbb{R} \otimes \mathbb{R} \simeq \mathbb{R}$.

Как векторное пространство $\mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2$. На плоскости $\mathbb{R}^2$ мы имеем покомпонентное умножение декартовых координат:

$$\mathbb{R}^2 \times \mathbb{R}^2 \ni (a_1 \oplus b_1), (a_2 \oplus b_2) \mapsto (a_1 \oplus b_1)(a_2 \oplus b_2) = (a_1 a_2 \oplus b_1 b_2) \in \mathbb{R}^2.$$

Таким образом, прямая сумма двух вещественных прямых, рассматриваемых как алгебры, изоморфна евклидовой плоскости с покомпонентным умножением точек. □

Пример 26.3.2. Рассмотрим тензорное произведение комплексной плоскости на вещественную прямую $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}$, рассматриваемых как вещественные алгебры. Пусть $z = z \cdot 1 \in \mathbb{C}$ и $a = a \cdot 1 \in \mathbb{R}$. Тогда произведение элементов из $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}$ примет вид

$$(\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}) \times (\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}) \ni (z_1 \otimes_{\mathbb{R}} a_1), (z_2 \otimes_{\mathbb{R}} a_2) \mapsto (z_1 z_2 \otimes_{\mathbb{R}} a_1 a_2) = z_1 z_2 a_1 a_2 \cdot \mathbb{1} \in \mathbb{C}.$$

где $\mathbb{1} := 1 \otimes_{\mathbb{R}} 1$. Тем самым мы доказали изоморфизм $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R} \simeq \mathbb{C}$.

Как вещественные векторные пространства $\mathbb{C} \oplus \mathbb{R} = \mathbb{R}^3$. Умножение в этой прямой сумме задается следующим образом

$$(\mathbb{C} \oplus \mathbb{R}) \times (\mathbb{C} \oplus \mathbb{R}) \ni (z_1 \oplus a_1), (z_2 \oplus a_2) \mapsto (z_1 z_2 \oplus a_1 a_2) \in (\mathbb{C} \oplus \mathbb{R}). \quad \square$$

Пример 26.3.3. Рассмотрим тензорное произведение двух комплексных плоскостей $\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C}$, которые рассматриваются как алгебры над полем комплексных чисел. Пусть $z = z \cdot 1 \in \mathbb{C}$ и $w = w \cdot 1 \in \mathbb{C}$. Тогда умножение в тензорном произведении задается следующим образом

$$(\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C}) \times (\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C}) \ni (z_1 \otimes_{\mathbb{C}} w_1), (z_2 \otimes_{\mathbb{C}} w_2) \mapsto (z_1 z_2 \otimes_{\mathbb{C}} w_1 w_2) = z_1 z_2 w_1 w_2 \cdot \mathbb{1} \in \mathbb{C}. \quad (26.56)$$

где $\mathbb{1} := 1 \otimes_{\mathbb{C}} 1$. То есть существует изоморфизм $\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C} \simeq \mathbb{C}$.

Как комплексное векторное пространство $\mathbb{C} \oplus \mathbb{C} = \mathbb{C}^2$. В этой прямой сумме определено покомпонентное умножение

$$\mathbb{C}^2 \times \mathbb{C}^2 \ni (z_1 \oplus w_1), (z_2 \oplus w_2) \mapsto (z_1 z_2 \oplus w_1 w_2) \in \mathbb{C}^2.$$

Мы видим полную аналогию с тензорным произведением и прямой суммой вещественных прямых.

Если комплексную плоскость рассматривать не как одномерное комплексное пространство, а как двумерное вещественное пространство, $\mathbb{C} \simeq \mathbb{R}^2$, то тензорное произведение алгебр изменится. Обозначим $z = x + iy \in \mathbb{R}^2$, где 1 и i рассматриваются как базисные векторы в $\mathbb{R}^2$. Аналогично, $w = u + iv \in \mathbb{R}^2$. Тогда умножение в $\mathbb{R}^2 \simeq \mathbb{C}$ задается формулами:

$$\begin{aligned} z_1 z_2 &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1), \\ w_1 w_2 &= (u_1 u_2 - v_1 v_2) + i(u_1 v_2 + u_2 v_1). \end{aligned}$$

Теперь для умножения элементов тензорного произведения $\mathbb{R}^2 \otimes_{\mathbb{R}} \mathbb{R}^2 \simeq \mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}$ справедлива формула

$$\begin{aligned} (z_1 \otimes_{\mathbb{R}} w_1)(z_2 \otimes_{\mathbb{R}} w_2) &= (z_1 z_2 \otimes_{\mathbb{R}} w_1 w_2) = \\ &= [(x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)] \otimes_{\mathbb{R}} [(u_1 u_2 - v_1 v_2) + i(u_1 v_2 + u_2 v_1)] = \\ &= (x_1 x_2 - y_1 y_2)(u_1 u_2 - v_1 v_2) \mathbb{1} \otimes_{\mathbb{R}} \mathbb{1} + (x_1 y_2 + x_2 y_1)(u_1 u_2 - v_1 v_2) i \otimes_{\mathbb{R}} \mathbb{1} + \\ &\quad + (x_1 x_2 - y_1 y_2)(u_1 v_2 + u_2 v_1) \mathbb{1} \otimes_{\mathbb{R}} i + (x_1 y_2 + x_2 y_1)(u_1 v_2 + u_2 v_1) i \otimes_{\mathbb{R}} i. \end{aligned}$$

Как вещественное векторное пространство тензорное произведение $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}$ четырехмерно с базисными векторами $\mathbb{1} \otimes_{\mathbb{R}} \mathbb{1}$, $i \otimes_{\mathbb{R}} \mathbb{1}$, $\mathbb{1} \otimes_{\mathbb{R}} i$ и $i \otimes_{\mathbb{R}} i$. Нетрудно видеть, что $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} = \mathbb{C}^2$. Чтобы вернуться к комплексному тензорному произведению (26.56) необходимо положить

$$\mathbb{1} \otimes_{\mathbb{R}} \mathbb{1} := \mathbb{1}, \quad i \otimes_{\mathbb{R}} \mathbb{1} := i\mathbb{1}, \quad \mathbb{1} \otimes_{\mathbb{R}} i := i\mathbb{1}, \quad i \otimes_{\mathbb{R}} i := -\mathbb{1}. \quad \square$$

Мы видим, что если одно и то же множество элементов можно рассматривать, как алгебры над различными полями, то тензорное произведение таких алгебр зависит от того, над каким полем они рассматриваются. Поэтому там, где это существенно, мы будем писать соответствующий индекс у тензорного произведения $\mathfrak{a} \otimes_{\mathbb{F}} \mathfrak{a}$. В противном случае, когда выбор поля ясен из контекста, индекс, для краткости, будет опущен.

Предложение 26.3.2. *Существуют изоморфизмы:*

$$\begin{aligned} \mathbb{H} \otimes_{\mathbb{R}} \mathbb{R} &= \mathbb{H}, \\ \mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} &= \mathbf{mat}(2, \mathbb{C}), \\ \mathbb{H} \otimes_{\mathbb{R}} \mathbb{H} &= \mathbf{mat}(4, \mathbb{R}). \end{aligned}$$

Доказательство. См., например, [193], глава I, предложение 4.2. $\square$

Поскольку алгебры Клиффорда являются алгебрами, то имеет смысл рассматривать алгебры матриц, элементы которых принадлежат алгебрам Клиффорда.

Рассмотрим алгебру Клиффорда $\mathbf{cl}(p, q, \mathbb{F})$ с образующими e^a , $a = 1, \dots, p + q$. Определим $p + q + 2$ матрицы $\check{e}^A$, $A = 1, \dots, p + q + 2$, следующим образом:

$$\check{e}^a := \begin{pmatrix} e^a & 0 \\ 0 & e^a \end{pmatrix}, \quad \check{e}^{p+q+1} := \check{e}^+ := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \check{e}^{p+q+2} := \check{e}^- := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (26.57)$$

Тогда эти матрицы будут удовлетворять следующим антикоммутиационным соотношениям

$$\check{e}^A \check{e}^B + \check{e}^B \check{e}^A = 2\eta^{AB}, \quad \eta^{AB} := \text{diag}(\underbrace{+ \dots +}_p \underbrace{- \dots -}_q + -).$$

Следовательно, матрицы $\check{e}^A$ в свою очередь можно рассматривать, как образующие новой алгебры Клиффорда $\mathbf{cl}(p + 1, q + 1, \mathbb{F})$ (с точностью до перестановки образующих). Таким образом, мы построили диагональное вложение $\mathbf{cl}(p, q, \mathbb{F}) \hookrightarrow \mathbf{cl}(p + 1, q + 1, \mathbb{F})$:

$$\mathbf{cl}(p, q, \mathbb{F}) \ni U \mapsto \check{U} = \begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix} \in \mathbf{cl}(p + 1, q + 1, \mathbb{F}).$$

Поскольку матрица $\check{U}$ пропорциональна единичной 2×2 матрице, то она коммутирует со всеми другими матрицами. В частности, с четырьмя матрицами

$$\begin{aligned} \frac{1}{2}(1 + \check{e}^+ \check{e}^-) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, & \frac{1}{2}(\check{e}^+ - \check{e}^-) &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ \frac{1}{2}(\check{e}^+ + \check{e}^-) &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, & \frac{1}{2}(1 - \check{e}^+ \check{e}^-) &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

которые образуют базис в $\mathbf{mat}(2, \mathbf{cl}(p, q, \mathbb{F}))$. Поэтому произвольный элемент алгебры $\mathbf{mat}(2, \mathbf{cl}(p, q, \mathbb{F}))$ представим в виде

$$\begin{aligned} \begin{pmatrix} U & V \\ W & T \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} U + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} V + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} W + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} T = \\ &= \frac{1}{2}(1 + \check{e}^+ \check{e}^-)U + \frac{1}{2}(\check{e}^+ - \check{e}^-)V + \frac{1}{2}(\check{e}^+ + \check{e}^-)W + \frac{1}{2}(1 - \check{e}^+ \check{e}^-)T, \end{aligned} \quad (26.58)$$

где U , V , W и T – элементы алгебры Клиффорда $\mathbf{cl}(p, q, \mathbb{F})$. С другой стороны, произвольный элемент алгебры Клиффорда $\mathbf{cl}(p + 1, q + 1, \mathbb{F})$ можно записать в виде

$$S + e^+ S_+ + e^- S_- + e^+ e^- S_{+-},$$

где S , S_+ , S_- , S_{+-} – некоторые элементы подалгебры $\mathbf{cl}(p, q, \mathbb{F})$. Сравнивая полученное выражение с представлением (26.58), можно установить взаимно однозначное соответствие между элементами алгебр $\mathbf{cl}(p + 1, q + 1, \mathbb{F})$ и $\mathbf{mat}(2, \mathbf{cl}(p, q, \mathbb{F}))$, т.е. между элементами S , S_+ , S_- , S_{+-} и U , V , W , T , которое мы не будем выписывать. Поскольку все соотношения линейны, то мы доказали изоморфизм алгебр Клиффорда

$$\mathbf{cl}(p + 1, q + 1, \mathbb{F}) \simeq \mathbf{mat}(2, \mathbf{cl}(p, q, \mathbb{F})). \quad (26.59)$$

Пример 26.3.4. В примере 26.2.6 мы доказали изоморфизм алгебр $\mathfrak{cl}(0, 2, \mathbb{R}) \simeq \mathbb{H}$. Согласно приведенному выше построению отсюда вытекает изоморфизм $\mathfrak{cl}(1, 3, \mathbb{R}) \simeq \mathfrak{mat}(2, \mathbb{H})$. Этот изоморфизм уже был явно построен в примере 26.2.11. $\square$

Теперь вернемся к началу нашего построения (26.57). Поскольку матрицы $\{e^+, e^-\} = \{e^i\}$, $i = 1, 2$, удовлетворяют антикоммутационным соотношениям

$$e^i e^j + e^j e^i = 2\eta^{ij}, \quad \eta^{ij} := \text{diag}(+-),$$

то они сами по себе генерируют алгебру Клиффорда $\mathfrak{cl}(1, 1, \mathbb{F})$. Поэтому из представления (26.57) следует изоморфизм

$$\mathfrak{mat}(2, \mathfrak{cl}(p, q, \mathbb{F})) \simeq \mathfrak{cl}(p, q, \mathbb{F}) \otimes \mathfrak{cl}(1, 1, \mathbb{F}).$$

Полученный изоморфизм вместе с изоморфизмом (26.59) доказывает третью часть следующего утверждения.

Теорема 26.3.1. *Существуют изоморфизмы:*

$$\mathfrak{cl}(p+2, 0, \mathbb{F}) \simeq \mathfrak{cl}(0, p, \mathbb{F}) \otimes \mathfrak{cl}(2, 0, \mathbb{F}), \quad (26.60)$$

$$\mathfrak{cl}(0, q+2, \mathbb{F}) \simeq \mathfrak{cl}(q, 0, \mathbb{F}) \otimes \mathfrak{cl}(0, 2, \mathbb{F}), \quad (26.61)$$

$$\mathfrak{cl}(p+1, q+1, \mathbb{F}) \simeq \mathfrak{cl}(p, q, \mathbb{F}) \otimes \mathfrak{cl}(1, 1, \mathbb{F}). \quad (26.62)$$

Доказательство. Пусть e^A , $A = 1, \dots, p+2$, – ортонормальный базис в алгебре Клиффорда $\mathfrak{cl}(p+2, 0, \mathbb{F})$. Пусть e^a , $a = 1, \dots, p$, и $\check{e}^1, \check{e}^2$ – ортонормальные базисы, соответственно, в $\mathfrak{cl}(0, p, \mathbb{F})$ и $\mathfrak{cl}(2, 0, \mathbb{F})$. Определим отображение $\mathfrak{cl}(p+2, 0, \mathbb{F}) \rightarrow \mathfrak{cl}(0, p, \mathbb{F}) \otimes \mathfrak{cl}(2, 0, \mathbb{F})$, отождествив базисные векторы:

$$e^A := \begin{cases} e^a \otimes \check{e}^1 \check{e}^2, & 1 \leq A = a \leq p, \\ 1 \otimes \check{e}^{A-p}, & A = p+1, p+2 \end{cases} \quad (26.63)$$

Теперь нетрудно проверить, что выполнено равенство

$$\{e^A, e^B\} = 2\delta^{AB}\mathbb{1},$$

где мы отождествили единицы в алгебрах Клиффорда $\mathbb{1} := 1 \otimes 1$. Поскольку отображение (26.63), продолженное по линейности, сюръективно, то оно определяет изоморфизм (26.60).

Аналогично доказывается изоморфизм (26.61).

Чтобы доказать существование изоморфизма (26.62) проведем следующее построение. Пусть e^A , $A = 1, \dots, p+q+2$, – ортонормальный базис в алгебре Клиффорда $\mathfrak{cl}(p+1, q+1, \mathbb{F})$ с квадратичной формой $Q(A) := \eta^{AB}A_A A_B$. Пусть e^a , $a = 1, \dots, p+q$, и $\check{e}^1, \check{e}^2$ – ортонормальные базисы в $\mathfrak{cl}(p, q, \mathbb{F})$ и $\mathfrak{cl}(1, 1, \mathbb{F})$, соответственно. Проведем отождествление:

$$e^A := \begin{cases} e^a \otimes \check{e}^1 \check{e}^2, & 1 \leq A = a \leq p, \\ 1 \otimes \check{e}^1, & A = p+1, \\ e^a \otimes \check{e}^1 \check{e}^2, & p+2 \leq A = a+1 \leq p+q+1, \\ 1 \otimes \check{e}^2, & A = p+q+2. \end{cases} \quad (26.64)$$

Тогда нетрудно проверить равенство

$$\{e^A, e^B\} = 2\eta^{AB}\mathbb{1}, \quad \eta^{AB} = \text{diag}(\underbrace{+\dots+}_{p+1} \underbrace{-\dots-}_{q+1}).$$

где мы отождествили единицы. Поскольку отождествление базисных векторов взаимно однозначно, то мы доказали изоморфизм (26.62). $\square$

Аналогичные построения позволяют доказать наличие других изоморфизмов.

Теорема 26.3.2. *При $p \geq 1$ существует изоморфизм*

$$\mathfrak{cl}(p, q, \mathbb{F}) \simeq \mathfrak{cl}(q + 1, p - 1, \mathbb{F}). \quad (26.65)$$

Доказательство. Пусть задана алгебра Клиффорда $\mathfrak{cl}(p, q, \mathbb{F})$ с ортонормированными образующими e^a , $a = 1, \dots, p + q$. Перейдем к новым образующим

$$e'^1 := e^1, \quad e'^a := e^a e^1, \quad a = 2, \dots, p + q.$$

Тогда новые образующие с различными индексами будут антикоммутировать между собой, и, кроме того, $(e'^1)^2 = 1$ и $(e'^a)^2 = -(e^a)^2$ при $a \geq 2$. То есть справедливы следующие антикоммутационные соотношения

$$e'^a e'^b + e'^b e'^a = 2\eta'^{ab}, \quad \eta'^{ab} = \text{diag} \left(+ \underbrace{- \dots -}_{p-1} + \underbrace{\dots +}_p \right).$$

Следовательно, с точностью до перестановки новые образующие являются генераторами алгебры Клиффорда $\mathfrak{cl}(q + 1, p - 1, \mathbb{F})$. $\square$

Если данную теорему применить дважды, то получим тождество

$$\mathfrak{cl}(p, q, \mathbb{F}) \simeq \mathfrak{cl}(q + 1, p - 1, \mathbb{F}) \simeq \mathfrak{cl}(p, q, \mathbb{F}).$$

Поэтому ее нужно использовать только один раз.

Пример 26.3.5.

$$\begin{aligned} \mathfrak{cl}(2, 0, \mathbb{F}) &\simeq \mathfrak{cl}(1, 1, \mathbb{F}), & \mathfrak{cl}(3, 0, \mathbb{F}) &\simeq \mathfrak{cl}(2, 1, \mathbb{F}), \\ \mathfrak{cl}(4, 0, \mathbb{F}) &\simeq \mathfrak{cl}(1, 3, \mathbb{F}), & \mathfrak{cl}(3, 1, \mathbb{F}) &\simeq \mathfrak{cl}(2, 2, \mathbb{F}). \end{aligned}$$

Первый изоморфизм был установлен в примерах 26.2.4 и 26.2.5. Третий изоморфизм – в примере 26.2.11. $\square$

Теорема 26.3.3. *При $p \geq 4$ существует изоморфизм*

$$\mathfrak{cl}(p, q, \mathbb{F}) \simeq \mathfrak{cl}(p - 4, q + 4, \mathbb{F}). \quad (26.66)$$

Доказательство. Пусть задана алгебра Клиффорда $\mathfrak{cl}(p, q, \mathbb{F})$ с ортонормированными образующими e^a , $a = 1, \dots, p + q$. Перейдем к новым образующим

$$\begin{aligned} e'^a &:= e^a h, & a &= 1, 2, 3, 4, \\ e'^a &:= e^a, & a &= 5, \dots, p + q, \end{aligned}$$

где $h := e^1 e^2 e^3 e^4$. Можно проверить, что новые образующие удовлетворяют антикоммутационным соотношениям

$$e'^a e'^b + e'^b e'^a = 2\eta'^{ab}, \quad \eta'^{ab} = \text{diag} \left(\underbrace{+ \dots +}_{p-4} \underbrace{- \dots -}_{q+4} \right).$$

Следовательно, новые образующие являются генераторами алгебры Клиффорда $\mathfrak{cl}(p - 4, q + 4, \mathbb{F})$. $\square$

В отличие от предыдущей теоремы полученное утверждение можно применять к алгебрам Клиффорда несколько раз пока $p - 4 \geq 0$.

Пример 26.3.6.

$$\mathfrak{cl}(6, 0, \mathbb{F}) \simeq \mathfrak{cl}(2, 4, \mathbb{F}), \quad \mathfrak{cl}(5, 2, \mathbb{F}) \simeq \mathfrak{cl}(1, 6, \mathbb{F}). \quad \square$$

Теорема 26.3.4. *Существуют изоморфизмы*

$$\mathfrak{cl}(p, q + 8, \mathbb{F}) \simeq \mathbf{mat}(16, \mathfrak{cl}(p, q, \mathbb{F})). \quad (26.67)$$

$$\mathfrak{cl}(p + 8, q, \mathbb{F}) \simeq \mathbf{mat}(16, \mathfrak{cl}(p, q, \mathbb{F})). \quad (26.68)$$

Доказательство. Сначала докажем изоморфизм (26.67). С этой целью выберем ортонормальный базис $\{e^1, \dots, e^n, e^{n+1}, \dots, e^{n+8}\}$ в псевдоевклидовом векторном пространстве $\mathbb{R}^{p,q+8}$, где $n = p + q$. Положим

$$e'^a := e^a e^{n+1} \dots e^{n+8}, \quad a = 1, \dots, n = p + q.$$

Тогда набор векторов $\{e'^1, \dots, e'^n\}$ генерирует подалгебру, изоморфную $\mathfrak{cl}(p, q, \mathbb{F})$. Набор векторов $\{e^{n+1}, \dots, e^{n+8}\}$ генерирует подалгебру, которая изоморфна $\mathbf{mat}(16, \mathbb{R})$. Эти две подалгебры коммутируют между собой. Поэтому существует изоморфизм

$$\mathfrak{cl}(p, q + 8, \mathbb{F}) \simeq \mathfrak{cl}(p, q, \mathbb{F}) \otimes \mathbf{mat}(16, \mathbb{F}) \simeq \mathbf{mat}(16, \mathfrak{cl}(p, q, \mathbb{F})).$$

Аналогично доказывается второй изоморфизм (26.67). $\square$

Две последние теоремы были доказаны Э. Картаном [197].

Теорема 26.3.4 позволяет классифицировать алгебры Клиффорда. Действительно, если известны все алгебры Клиффорда $\mathfrak{cl}(p, q, \mathbb{F})$ при $p = 0, 1, \dots, 7$ и $q = 0, 1, \dots, 7$, то изоморфизм (26.67) можно использовать в качестве рекуррентного соотношения для построения всех остальных алгебр. Это свойство известно как *периодичность 8* или *периодичность Картана–Ботта*.

Приведем примеры использования последней теоремы.

Пример 26.3.7. В примере 26.2.3 доказан изоморфизм $\mathfrak{cl}(0, 1, \mathbb{R}) \simeq \mathbb{C}$. поэтому из теоремы 26.3.4 следуют изоморфизмы $\mathfrak{cl}(8, 1, \mathbb{R}) \simeq \mathfrak{cl}(0, 9, \mathbb{R}) \simeq \mathbf{mat}(16, \mathbb{C})$. $\square$

Пример 26.3.8. В примере 26.2.5 доказан изоморфизм $\mathfrak{cl}(1, 1, \mathbb{R}) \simeq \mathbf{mat}(2, \mathbb{R})$. Следовательно, $\mathfrak{cl}(9, 1, \mathbb{R}) \simeq \mathfrak{cl}(1, 9, \mathbb{R}) \simeq \mathbf{mat}(16, \mathbf{mat}(2, \mathbb{R})) \simeq \mathbf{mat}(32, \mathbb{R})$. $\square$

Сначала рассмотрим классификацию алгебр Клиффорда в вещественном случае. Теоремы 26.3.1–26.3.4 позволяют для невырожденных квадратичных форм Q из знания алгебр Клиффорда низших размерностей определить все остальные алгебры Клиффорда с помощью рекуррентных соотношений. А именно, четыре неизоморфные между собой алгебры

$$\begin{aligned} \mathfrak{cl}(1, 0, \mathbb{R}) &\simeq \mathbb{R} \oplus \mathbb{R}, & \mathfrak{cl}(1, 1, \mathbb{R}) &\simeq \mathbf{mat}(2, \mathbb{R}), \\ \mathfrak{cl}(0, 1, \mathbb{R}) &\simeq \mathbb{C}, & \mathfrak{cl}(0, 2, \mathbb{R}) &\simeq \mathbb{H}, \end{aligned} \quad (26.69)$$

определяют все остальные алгебры Клиффорда путем сдвига аргументов p и q .

$\begin{smallmatrix} p-q \\ p+q \end{smallmatrix}$	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
0								$\mathbb{R}$							
1							$\mathbb{C}$		${}^2\mathbb{R}$						
2						$\mathbb{H}$		$\mathbb{R}(2)$		$\mathbb{R}(2)$					
3					${}^2\mathbb{H}$		$\mathbb{C}(2)$		${}^2\mathbb{R}(2)$		$\mathbb{C}(2)$				
4				$\mathbb{H}(2)$		$\mathbb{H}(2)$		$\mathbb{R}(4)$		$\mathbb{R}(4)$		$\mathbb{H}(2)$			
5			$\mathbb{C}(4)$		${}^2\mathbb{H}(2)$		$\mathbb{C}(4)$		${}^2\mathbb{R}(4)$		$\mathbb{C}(4)$		${}^2\mathbb{H}(2)$		
6		$\mathbb{R}(8)$		$\mathbb{H}(4)$		$\mathbb{H}(4)$		$\mathbb{R}(8)$		$\mathbb{R}(8)$		$\mathbb{H}(4)$		$\mathbb{H}(4)$	
7	${}^2\mathbb{R}(8)$		$\mathbb{C}(8)$		${}^2\mathbb{H}(4)$		$\mathbb{C}(8)$		${}^2\mathbb{R}(8)$		$\mathbb{C}(8)$		${}^2\mathbb{H}(4)$		$\mathbb{C}(8)$

Таблица 26.1: Вещественные алгебры Клиффорда низших размерностей, которые определяют все остальные алгебры Клиффорда, благодаря периодичности 8 (периодичность Картана–Ботта).

Поместим алгебры Клиффорда низших размерностей, $p \leq 7$ и $q \leq 7$, в таблицу 26.1, где мы ввели следующие сокращения:

$$\begin{aligned}
 \mathbb{R}(k) &:= \mathbf{mat}(k, \mathbb{R}), & {}^2\mathbb{R} &:= \mathbb{R} \oplus \mathbb{R}, \\
 \mathbb{C}(k) &:= \mathbf{mat}(k, \mathbb{C}), & {}^2\mathbb{H} &:= \mathbb{H} \oplus \mathbb{H}, \\
 {}^2\mathbb{R}(k) &:= \mathbb{R}(k) \oplus \mathbb{R}(k), & {}^2\mathbb{H}(k) &:= \mathbb{H}(k) \oplus \mathbb{H}(k).
 \end{aligned}$$

В приведенной таблице четыре алгебры (26.69) обведены рамкой. Формулы 26.65 и 26.66 устанавливают изоморфизм некоторых алгебр Клиффорда, которые находятся в одной строке. Переход от строк 1–4 к строкам 5–7 осуществляется с помощью изоморфизма 26.62. При этом надо воспользоваться следующими изоморфизмами:

$$\begin{aligned}
 \mathbb{C} \otimes \mathbb{R}(2) &\simeq \mathbb{C}(2), & \mathbb{H} \otimes \mathbb{R}(2) &\simeq \mathbb{H}(2), \\
 \mathbb{R}(2) \otimes \mathbb{R}(2) &\simeq \mathbb{R}(4), & \mathbb{R}(4) \otimes \mathbb{R}(2) &\simeq \mathbb{R}(8), \\
 \mathbb{C}(2) \otimes \mathbb{R}(2) &\simeq \mathbb{C}(4), & \mathbb{H}(2) \otimes \mathbb{R}(2) &\simeq \mathbb{H}(4), \\
 \mathbb{C}(4) \otimes \mathbb{R}(2) &\simeq \mathbb{C}(8), & {}^2\mathbb{R} \otimes \mathbb{R}(2) &\simeq {}^2\mathbb{R}(2), \\
 {}^2\mathbb{R}(2) \otimes \mathbb{R}(2) &\simeq {}^2\mathbb{R}(4), & {}^2\mathbb{H}(2) \times \mathbb{R}(2) &\simeq {}^2\mathbb{H}(4),
 \end{aligned}$$

которые легко проверяются.

Приведенная классификация вещественных алгебр Клиффорда имеет несколько далеко идущих следствий. Одним из примеров является периодичность Ботта, которая устанавливает периодичность групп гомотопий для унитарных, ортогональных и симплектических групп. Например,

Теорема 26.3.5. *Для $n \geq k + 2$ и всех $k \geq 1$ имеют место следующие изоморфизмы гомотопических групп:*

$$\pi_k(\mathbb{O}(n)) \simeq \pi_k(\mathbb{SO}(n)) \simeq \begin{cases} \{e\}, & k = 2, 4, 5, 6 \pmod{8}, \\ \mathbb{Z}_2, & k = 0, 1 \pmod{8}, \\ \mathbb{Z}, & k = 3, 7 \pmod{8}. \end{cases} \quad (26.70)$$

Доказательство. См., например, [193]. $\square$

Классификация комплексных алгебр Клиффорда существенно проще вещественных. Рассмотрим алгебру Клиффорда $\text{cl}(p, q, \mathbb{C})$ над полем комплексных чисел. В соответствующем векторном пространстве $\mathbb{V}$ произведем замену базиса

$$e^a \mapsto e^a, \quad a = 1, \dots, p; \quad e^a \mapsto ie^a, \quad a = p+1, \dots, p+q.$$

Очевидно, что в новом базисе квадратичная форма станет единичной матрицей. Поэтому существует изоморфизм

$$\text{cl}(p, q, \mathbb{C}) \simeq \text{cl}(p+q, 0, \mathbb{C}).$$

Отсюда следует, что в комплексном случае достаточно классифицировать только алгебры Клиффорда $\text{cl}(n, 0, \mathbb{C})$, и эта классификация вытекает сразу из классификации вещественных алгебр Клиффорда.

Построение ведется следующим образом. В примерах 26.2.2 и 26.2.4 были доказаны изоморфизмы

$$\text{cl}(1, 0, \mathbb{C}) \simeq \mathbb{C} \oplus \mathbb{C}, \quad \text{cl}(2, 0, \mathbb{C}) \simeq \text{mat}(2, \mathbb{C}).$$

Учитывая изоморфизм комплексных алгебр Клиффорда одинаковой размерности, изоморфизм (26.60) можно переписать в виде

$$\text{cl}(p+2, 0, \mathbb{C}) \simeq \text{cl}(p, 0, \mathbb{C}) \otimes \text{cl}(2, 0, \mathbb{C}) \simeq \text{cl}(p, 0, \mathbb{C}) \otimes \text{mat}(2, \mathbb{C}).$$

Поэтому существует изоморфизм

$$\text{cl}(3, 0, \mathbb{C}) \simeq \text{cl}(1, 0, \mathbb{C}) \otimes \text{cl}(2, 0, \mathbb{C}) \simeq (\mathbb{C} \oplus \mathbb{C}) \otimes \text{mat}(2, \mathbb{C}) \simeq \text{mat}(2, \mathbb{C}) \oplus \text{mat}(2, \mathbb{C}).$$

Аналогично доказывается изоморфизм

$$\text{cl}(4, 0, \mathbb{C}) \simeq \text{cl}(2, 0, \mathbb{C}) \otimes \text{cl}(2, 0, \mathbb{C}) \simeq \text{mat}(2, \mathbb{C}) \otimes \text{mat}(2, \mathbb{C}) \simeq \text{mat}(2^2, \mathbb{C}).$$

Нетрудно проверить, что для произвольных размерностей справедливы изоморфизмы

$$\begin{aligned} (\mathbb{C} \oplus \mathbb{C}) \otimes \text{mat}(n, \mathbb{C}) &\simeq \text{mat}(n, \mathbb{C}) \oplus \text{mat}(n, \mathbb{C}), \\ \text{mat}(n, \mathbb{C}) \otimes \text{mat}(2, \mathbb{C}) &\simeq \text{mat}(2n, \mathbb{C}). \end{aligned} \tag{26.71}$$

Поэтому продолжение данного построения по n доказывает следующее утверждение.

Теорема 26.3.6. *Имеют место следующие изоморфизмы:*

$$\text{cl}(n, 0, \mathbb{C}) \simeq \begin{cases} \text{mat}(2^{\frac{n}{2}}, \mathbb{C}), & n \text{ четное,} \\ 2\text{mat}(2^{\frac{n-1}{2}}, \mathbb{C}) := \text{mat}(2^{\frac{n-1}{2}}, \mathbb{C}) \oplus \text{mat}(2^{\frac{n-1}{2}}, \mathbb{C}), & n \text{ нечетное.} \end{cases} \tag{26.72}$$

Таким образом, в комплексном случае имеет место периодичность 2.

Следствие. Комплексные алгебры Клиффорда $\text{cl}(p, q, \mathbb{C})$ имеют точное представление в алгебре комплексных матриц $\text{mat}(m, \mathbb{C})$. Минимальная размерность представления равна $m = 2^{\lceil (n+1)/2 \rceil}$, где $n := p+q$ и квадратные скобки обозначают целую часть.

26.4 Представления алгебр Клиффорда

В приложениях встречаются, как правило, представления алгебр Клиффорда. Начнем с определений.

Определение. Пусть задано конечномерное векторное пространство $\mathbb{W}$ над полем $\mathbb{k}$, содержащим поле $\mathbb{F} \subseteq \mathbb{k}$. Тогда $\mathbb{k}$ -представлением алгебры Клиффорда $\mathbf{cl}(\mathbb{V}, Q)$ над полем $\mathbb{F}$ называется гомоморфизм

$$\rho : \mathbf{cl}(\mathbb{V}, Q) \ni U \mapsto \rho(U) \in \mathbf{hom}_{\mathbb{k}}(\mathbb{W}, \mathbb{W}) \quad (26.73)$$

алгебры Клиффорда в алгебру $\mathbb{k}$ -линейных преобразований векторного пространства $\mathbb{W}$. Векторное пространство $\mathbb{W}$ называется $\mathbf{cl}(\mathbb{V}, Q)$ -модулем над полем $\mathbb{k}$. $\square$

При фиксированном базисе в $\mathbb{W}$ каждому элементу алгебры Клиффорда ставится в соответствие некоторая матрица с элементами из поля $\mathbb{k}$. При этом выполнено равенство

$$\rho(UV) = \rho(U)\rho(V),$$

где в правой части стоит обычное умножение матриц.

В дальнейшем мы упростим обозначения:

$$\rho(U)(X) := UX, \quad \forall U \in \mathbf{cl}(\mathbb{V}, Q), \quad \forall X \in \mathbb{W}.$$

Нас будут интересовать случаи $\mathbb{k} = \mathbb{R}, \mathbb{C}$ или $\mathbb{H}$. При $\mathbb{k} = \mathbb{H}$ векторное пространство $\mathbb{W}$ рассматривается не над полем, а над алгеброй, т.е. является $\mathbb{H}$ -модулем.

Заметим, что комплексное векторное пространство изоморфно вещественному векторному пространству $\mathbb{W}$ вдвое большей размерности с линейным отображением (комплексной структурой)

$$J : \mathbb{W} \rightarrow \mathbb{W},$$

удовлетворяющим условию $J^2 = -\text{id}$. Поэтому комплексное представление алгебры Клиффорда есть просто вещественное представление

$$\rho : \mathbf{cl}(\mathbb{V}, Q) \rightarrow \mathbf{hom}_{\mathbb{R}}(\mathbb{W}, \mathbb{W}),$$

которое удовлетворяет равенству

$$\rho(U)J = J\rho(U), \quad \forall U \in \mathbf{cl}(\mathbb{V}, Q). \quad (26.74)$$

Следовательно, образ отображения ρ коммутирует с подалгеброй в $\mathbf{hom}_{\mathbb{R}}(\mathbb{W}, \mathbb{W})$, которая натянута на матрицы id , J и изоморфна полю комплексных чисел $\mathbb{C}$.

Пример 26.4.1. Рассмотрим двумерную вещественную алгебру Клиффорда $\mathbf{cl}(0, 1, \mathbb{R}) \simeq \mathbb{C}$. Эта алгебра имеет точное неприводимое одномерное комплексное представление

$$\mathbf{cl}(0, 1, \mathbb{R}) \ni z := u_0 + iu_1 \mapsto \rho(z) = z \in \mathbf{hom}_{\mathbb{C}}(\mathbb{W}, \mathbb{W}),$$

где $\mathbb{W} \simeq \mathbb{C}$ – одномерное комплексное пространство. Элементы алгебры Клиффорда $z \in \mathbf{cl}(0, 1, \mathbb{R})$ действуют в $\mathbb{W}$ простым умножением:

$$\mathbb{W} \ni w \mapsto zw \in \mathbb{W}.$$

Рассмотрим теперь $\mathbb{W}$ как двумерное вещественное векторное пространство. Пусть $w = w_0 + iw_1 \in \mathbb{R}^2$. Тогда действие алгебры Клиффорда в $\mathbb{W}$ можно записать в вещественной форме:

$$\mathbb{W} \ni \begin{pmatrix} w_0 \\ w_1 \end{pmatrix} \mapsto \begin{pmatrix} u_0 & -u_1 \\ u_1 & u_0 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \end{pmatrix} = \begin{pmatrix} u_0 w_0 - u_1 w_1 \\ u_0 w_1 + u_1 w_0 \end{pmatrix} \in \mathbb{W}.$$

То есть каждому элементу алгебры Клиффорда z ставится во взаимно однозначное соответствие вещественная 2×2 матрица. При этом базис алгебры Клиффорда отображается в матрицы:

$$e \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad e^1 = i \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Отсюда следует выражение для комплексной структуры

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Теперь нетрудно проверить, что $J^2 = -\text{id}$ и равенства (26.74) выполнены. В рассматриваемом случае подалгебра в $\text{hom}_{\mathbb{R}}(\mathbb{W}, \mathbb{W})$, натянутая на векторы id и J совпадает со всем образом представления ρ .

Построенное вещественное представление алгебры Клиффорда $\text{cl}(0, 1, \mathbb{R})$ является неприводимым. Действительно, если бы оно было приводимым, то комплексная структура должна была бы представляться диагональной матрицей, что противоречит условию $J^2 = -\text{id}$. $\square$

Кватернионные представления алгебр Клиффорда также эквивалентны вещественным представлениям, но вчетверо большей размерности. В этом случае в векторном пространстве $\mathbb{W}$ задано три линейных оператора $I, J, K : \mathbb{W} \rightarrow \mathbb{W}$, которые удовлетворяют равенствам:

$$I^2 = J^2 = K^2 = -\text{id}, \\ IJ = -JI = K, \quad JK = -KJ = I, \quad KI = -IK = J.$$

При этом векторное пространство $\mathbb{W}$ становится $\mathbb{H}$ -модулем. Представление $\rho : \mathbb{W} \rightarrow \mathbb{W}$ является кватернионным, если выполнены следующие коммутационные соотношения:

$$\rho(U)I = I\rho(U), \quad \rho(U)J = J\rho(U), \quad \rho(U)K = K\rho(U)$$

для всех $U \in \text{cl}(\mathbb{V}, Q)$. То есть алгебра $\text{hom}_{\mathbb{R}}(\mathbb{W}, \mathbb{W})$ имеет коммутативную подалгебру, натянутую на матрицы id, I, J, K , которая изоморфна алгебре кватернионов $\mathbb{H}$.

Каждое комплексное представление вещественной алгебры Клиффорда $\text{cl}(p, q, \mathbb{R})$ автоматически порождает комплексное представление вещественной алгебры $\text{cl}(p, q, \mathbb{R}) \otimes_{\mathbb{R}} \mathbb{C} \simeq \text{cl}(p + q, 0, \mathbb{R})$.

Каждое кватернионное представление вещественной алгебры $\text{cl}(p, q, \mathbb{R})$ является автоматически комплексным путем сужения алгебры кватернионов до поля комплексных чисел $\mathbb{C} \subset \mathbb{H}$. Конечно, комплексная размерность каждого $\mathbb{H}$ -модуля является четной.

Определение. $\mathbb{k}$ -представление (26.73) называется *приводимым*, если векторное пространство $\mathbb{W}$ можно представить в виде нетривиальной прямой суммы

$$\mathbb{W} = \mathbb{W}_1 \oplus \mathbb{W}_2,$$

такой, что выполнено равенство $\rho(U)(\mathbb{W}_i) \subseteq \mathbb{W}_i$ для $i = 1, 2$ и для всех $U \in \mathbf{cl}(\mathbb{V}, Q)$. В этом случае можно писать

$$\rho = \rho_1 \oplus \rho_2,$$

где $\rho_i(U) := \rho(U)|_{\mathbb{W}_i}$, $i = 1, 2$, – сужение представления ρ на $\mathbb{W}_i \subset \mathbb{W}$.

В противном случае представление называется *неприводимым*. $\square$

Поскольку алгебры конечномерны, то представление будет неприводимым тогда и только тогда, когда векторное пространство $\mathbb{W}$ не будет содержать собственных инвариантных подпространств.

Предложение 26.4.1. Каждое $\mathbb{k}$ -представление алгебры Клиффорда $\mathbf{cl}(\mathbb{V}, Q)$ можно разложить в прямую сумму $\rho = \rho_1 \oplus \dots \oplus \rho_m$ неприводимых представлений.

Доказательство. Если представление неприводимо, то оно уже имеет требуемый вид. Если представление приводимо, то $\rho = \rho_1 \oplus \rho_2$. Если какое-либо из представлений ρ_1 или ρ_2 приводимо, то его также можно разложить, и так далее. Этот процесс прекратится, поскольку $\mathbf{cl}(\mathbb{V}, Q)$ -модуль $\mathbb{W}$ имеет конечную размерность. $\square$

Определение. Два представления $\rho_i : \mathbf{cl}(\mathbb{V}, Q) \rightarrow \mathbf{hom}_{\mathbb{k}}(\mathbb{W}_i, \mathbb{W}_i)$, $i = 1, 2$ эквивалентны, если существует $\mathbb{k}$ -линейный изоморфизм $f : \mathbb{W}_1 \rightarrow \mathbb{W}_2$ такой, что $\rho_2(U) = f \circ \rho_1(U) \circ f^{-1}$ для всех $U \in \mathbf{cl}(\mathbb{V}, Q)$. $\square$

Из классификации алгебр Клиффорда, которая была дана в предыдущем разделе, мы знаем, что каждая алгебра Клиффорда $\mathbf{cl}(p, q, \mathbb{F})$ изоморфна одной из двух матричных алгебр $\mathbf{mat}(2^m, \mathbb{k})$ или $\mathbf{mat}(2^m, \mathbb{k}) \oplus \mathbf{mat}(2^m, \mathbb{k})$, где $\mathbb{k} = \mathbb{R}, \mathbb{C}$ или $\mathbb{H}$, для некоторого натурального m . Теория представлений для таких алгебр особенно проста.

Теорема 26.4.1. Пусть $\mathbb{k} = \mathbb{R}, \mathbb{C}$ или $\mathbb{H}$. Рассмотрим кольцо $n \times n$ матриц $\mathbf{mat}(n, \mathbb{k})$ как алгебру над $\mathbb{R}$. Тогда фундаментальное представление $\rho : \mathbb{k}^n \rightarrow \mathbb{k}^n$ алгебры $\mathbf{mat}(n, \mathbb{k})$ с точностью до эквивалентности является единственным неприводимым представлением алгебры $\mathbf{mat}(n, \mathbb{k})$.

Алгебра $\mathbf{mat}(n, \mathbb{k}) \oplus \mathbf{mat}(n, \mathbb{k})$ с точностью до эквивалентности имеет два неприводимых представления, которые действуют в $\mathbb{k}^n$:

$$\rho_1(U_1 \oplus U_2) := \rho(U_1) \quad \text{или} \quad \rho_2(U_1 \oplus U_2) := \rho(U_2),$$

где $U_1 \oplus U_2 \in \mathbf{mat}(n, \mathbb{k}) \oplus \mathbf{mat}(n, \mathbb{k})$.

Доказательство. Доказательство вытекает из классического результата: алгебра матриц $\mathbf{mat}(n, \mathbb{k})$ проста (т.е., не имеет двусторонних идеалов, отличных от нуля и всей алгебры) и имеет только одно точное неприводимое представление с точностью до эквивалентности (см., например, [198]). $\square$

Из классификации вещественных (таблица 26.1) и комплексных (теорема 26.72) алгебр Клиффорда вытекает

Теорема 26.4.2. Обозначим через $\nu(p, q, \mathbb{R})$ число неэквивалентных неприводимых вещественных представлений алгебры Клиффорда $\text{cl}(p, q, \mathbb{R})$ и через $\nu(n, 0, \mathbb{C})$ – число неэквивалентных неприводимых комплексных представлений алгебры Клиффорда $\text{cl}(n, 0, \mathbb{C})$. Тогда

$$\begin{aligned} \nu(p, q, \mathbb{R}) &= \begin{cases} 2, & \text{если } p - q = 1 \pmod{4}, \\ 1, & \text{в противном случае,} \end{cases} \\ \nu(n, 0, \mathbb{C}) &= \begin{cases} 2, & n - \text{нечетное,} \\ 1, & n - \text{четное.} \end{cases} \end{aligned} \quad (26.75)$$

Доказательство. В первом случае достаточно обратить внимание на вертикальные колонки в таблице 26.1 при $p - q = -7, -3, 1, 5$. Вторым случаем прямо следует из теоремы 26.72. $\square$

26.5 Свойства алгебр Клиффорда

Рассмотрим кратко основные свойства алгебр Клиффорда.

Алгебра Клиффорда содержит подалгебру cl_+ , натянутую на базисные векторы с четным числом индексов. Ее всегда можно представить в виде прямой суммы

$$\text{cl} = \text{cl}_+ \oplus \text{cl}_-. \quad (26.76)$$

Элементы алгебры Клиффорда, натянутые на базисные векторы с четным и нечетным числом индексов, будем, соответственно, называть *четными* и *нечетными*. Нетрудно проверить, что произведение четных или нечетных элементов дает четный элемент, а произведение четного и нечетного элемента – нечетный:

$$\begin{aligned} \text{cl}_+ \text{cl}_+ &\subset \text{cl}_+, & \text{cl}_- \text{cl}_- &\subset \text{cl}_+, \\ \text{cl}_+ \text{cl}_- &\subset \text{cl}_-, & \text{cl}_- \text{cl}_+ &\subset \text{cl}_-. \end{aligned}$$

Это значит, что алгебра Клиффорда является $\mathbb{Z}_2$ градуированной алгеброй. В соответствии с разложением (26.76) произвольный элемент алгебры Клиффорда можно представить единственным образом в виде суммы четного и нечетного элемента:

$$U = U_+ + U_-, \quad U_{\pm} \in \text{cl}_{\pm}.$$

Отметим, что, в отличие от внешней алгебры, в алгебре Клиффорда нельзя ввести $\mathbb{Z}$ градуировку, т.к. в общем случае произведение двух элементов степеней r и s содержит не только элементы степени $r + s$ (26.22).

Введем несколько отображений в алгебрах Клиффорда. Сначала напомним общее

Определение. *Инволюцией*, которую мы обозначим звездочкой $*$, в произвольной алгебре $\mathfrak{a}$ над полем вещественных или комплексных чисел называется отображение

$$\mathfrak{a} \ni U \mapsto U^* \in \mathfrak{a}, \quad (26.77)$$

удовлетворяющее следующим условиям:

$$U^{**} = U, \quad (\alpha U + \beta V)^* = \alpha^\dagger U^* + \beta^\dagger V^*, \quad (UV)^* = V^* U^* \quad (26.78)$$

для всех $U, V \in \mathfrak{a}$, где символ $\dagger$ обозначает комплексное сопряжение в случае $\alpha, \beta \in \mathbb{C}$. Мы также предполагаем, что отображение (26.77) является непрерывным в естественной топологии векторного пространства. $\square$

Отметим изменение порядка элементов при сопряжении произведения (26.78).

В теории алгебр Клиффорда вводится несколько полезных отображений, которые называются инволюциями, но не всегда являются таковыми с точки зрения общего определения, приведенного выше.

Определение. В соответствии с разложением в прямую сумму (26.19) произвольный элемент алгебры Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$ в ортонормальном базисе имеет вид

$$U = U_0 + U_1 + U_2 + \dots + U_n = U_+ + U_-.$$

Определим *инволюцию градуировки*, которую обозначим шляпкой, следующим образом

$$\widehat{U} := U_+ - U_-. \quad (26.79)$$

Введем операцию *реверса*, которую обозначим тильдой,

$$\widetilde{U} := U_0 + U_1 - U_2 - U_3 + U_4 + U_5 - U_6 - U_7 + \dots \quad (26.80)$$

Введем также *клиффордово сопряжение*, которое обозначим чертой, как композицию инволюции градуировки и реверса:

$$\overline{U} := \widetilde{\widehat{U}}. \quad \square \quad (26.81)$$

Инволюция градуировки (26.79) означает изменение знака образующих $e^a \mapsto -e^a$. Эта операция наследуется из векторного пространства $\mathbb{V}$, в котором определена инволюция (отражение)

$$\mathbb{V} \ni \quad A \mapsto -A \quad \in \mathbb{V}.$$

Эта инволюция естественным образом продолжается в тензорную алгебру $\otimes \mathbb{V}$, и оттуда спускается в алгебру Клиффорда, т.к. идеал (26.4), входящий в определение алгебры Клиффорда, инвариантен относительно изменения знака всех векторов.

Операция (26.80) названа реверсом, т.к. соответствует изменению порядка следования базисных векторов:

$$\mathfrak{cl}_r \ni \quad e^1 e^2 \dots e^{a_r} \mapsto e^{a_r} e^{a_{r-1}} \dots e^1 \quad \in \mathfrak{cl}_r.$$

Действительно, для того чтобы вернуть порядок индексов в произведении $e^{a_r} e^{a_{r-1}} \dots e^1$ к первоначальному, необходимо произвести

$$(r-1) + (r-2) + \dots + 1 = \frac{1}{2}r(r-1)$$

перестановок. Это приводит к изменению знака слагаемых в сумме (26.80) при $r = 2, 3, 6, 7, \dots$

Операция реверса также наследуется из тензорной алгебры. Действительно, в тензорной алгебре $\otimes \mathbb{V}$ определена инволюция (реверс), которая меняет порядок сомножителей во всех тензорных произведениях:

$$\otimes \mathbb{V} \ni \quad A_1 \otimes A_2 \otimes \dots \otimes A_r \mapsto A_r \otimes A_{r-1} \otimes \dots \otimes A_1 \quad \in \otimes \mathbb{V}.$$

Поскольку идеал (26.4), входящий в определение алгебры Клиффорда, инвариантен относительно этой инволюции, то она спускается на алгебру Клиффорда.

Пример 26.5.1. В алгебре Клиффорда $\mathfrak{cl}(0, 1, \mathbb{R}) \simeq \mathbb{C}$ операции инволюции градуировки и клиффордову сопряжению соответствует комплексное сопряжение. $\square$

Пример 26.5.2. В алгебре Клиффорда $\mathfrak{cl}(0, 2, \mathbb{R}) \simeq \mathbb{H}$ клиффордово сопряжение соответствует сопряжению кватернионов, т.к.

$$\overline{eu_0 + e^1u_1 + e^2u_2 + e^{12}u_{12}} = eu_0 - e^1u_1 - e^2u_2 - e^{12}u_{12}. \quad \square$$

Инволюции градуировки, реверса и клиффордова сопряжения можно записать в следующем виде

$$\widehat{U} = \sum_{k=0}^n (-1)^k U_k, \quad (26.82)$$

$$\widetilde{U} = \sum_{k=0}^n (-1)^{\frac{k(k-1)}{2}} U_k, \quad (26.83)$$

$$\overline{U} = \sum_{k=0}^n (-1)^{\frac{k(k+1)}{2}} U_k. \quad (26.84)$$

который удобен в вычислениях.

Пример 26.5.3. Для каждого вектора $A \in \mathbb{V} = \mathfrak{cl}_1$ справедливы равенства:

$$A^2 = -\hat{A}A = \tilde{A}A = -\bar{A}A = Q(A)e. \quad \square$$

Предложение 26.5.1. Реверс, инволюция градуировки и клиффордово сопряжение обладают следующими свойствами:

$$\begin{aligned} \widehat{\widehat{U}} &= U, & \widehat{\widehat{UV}} &= \widehat{U}\widehat{V}, \\ \widetilde{\widetilde{U}} &= U, & \widetilde{\widetilde{UV}} &= \widetilde{V}\widetilde{U}, \\ \overline{\overline{U}} &= U, & \overline{\overline{UV}} &= \overline{V}\overline{U}, \end{aligned}$$

для всех $U, V \in \mathfrak{cl}(\mathbb{V}, Q)$. Введенные операции (анти-)инволюции коммутируют между собой:

$$\widehat{\widetilde{U}} = \widetilde{\widehat{U}}, \quad \widetilde{\overline{U}} = \overline{\widetilde{U}}, \quad \widehat{\overline{U}} = \overline{\widehat{U}}. \quad (26.85)$$

Кроме того

$$\widehat{U^{-1}} = \widehat{U}^{-1}, \quad \widetilde{U^{-1}} = \widetilde{U}^{-1}, \quad \overline{U^{-1}} = \overline{U}^{-1}. \quad (26.86)$$

Доказательство. Прямая проверка. $\square$

Если алгебра Клиффорда рассматривается над полем вещественных чисел, то клиффордово сопряжение и реверс являются инволюциями, а инволюция градуировки – антиинволюцией, т.к. при действии на произведение элементов их порядок сохраняется. В случае алгебр Клиффорда над полем комплексных чисел ни одна из введенных операций не является инволюцией в смысле общего определения. Действительно, из общего определения инволюции для клиффордова сопряжения следует равенство $\widehat{iU} = -i\widehat{U}$. С другой стороны для $U = e$ справедливо соотношение $\widehat{ie} = ie$, что приводит к противоречию. Аналогичные противоречия возникают для инволюции градуировки и реверса.

Определение. Скалярная часть элемента алгебры Клиффорда называется *следом*

$$\mathrm{tr} U := \mathrm{tr}(eu_0) := u_0. \quad (26.87)$$

Квадратичная форма Q на векторном пространстве $\mathbb{V}$ индуцирует квадратичную форму в алгебре Клиффорда, для которой мы сохраним прежнее обозначение

$$Q(U) := \mathrm{tr}(\tilde{U}U) \in \mathbb{F}. \quad \square \quad (26.88)$$

Предложение 26.5.2. *Квадратичная форма в алгебре Клиффорда обладает следующим свойством*

$$Q(UV \dots W) = Q(U)Q(V) \dots Q(W).$$

Доказательство. Прямая проверка. $\square$

Отсюда вытекает равенство

$$Q(UV) = Q(VU) \Leftrightarrow \mathrm{tr}(\tilde{U}V) = \mathrm{tr}(U\tilde{V}).$$

Квадратичной форме в алгебре Клиффорда соответствует симметричная билинейная форма (поляризация), для которой мы также сохраним прежнее обозначение

$$Q(U, V) := \mathrm{tr}(\tilde{U}V) \in \mathbb{F}. \quad (26.89)$$

Пример 26.5.4. Если сузить алгебру Клиффорда на векторное подпространство, то получим исходную билинейную форму

$$Q(A, B) = \mathrm{tr}(\tilde{A}B) = \mathrm{tr}(e^a e^b) A_a B_b = \eta^{ab} A_a B_b$$

для всех $A = e^a A_a$ и $B = e^a B_a$ из $\mathbb{V}$. $\square$

Предложение 26.5.3. *Билинейная форма Q в алгебре Клиффорда невырождена тогда и только тогда, когда невырождена исходная билинейная форма Q в векторном пространстве $\mathbb{V}$.*

Доказательство. Прямая проверка. $\square$

Предложение 26.5.4. *Операция реверса является левой или правой операцией сопряжения для билинейной формы Q :*

$$Q(WU, V) = Q(U, \tilde{W}V), \quad Q(UW, V) = Q(U, V\tilde{W}).$$

Доказательство. Прямая проверка. $\square$

26.6 Группы (псевдо)вращений

В алгебрах Клиффорда некоторые подмножества элементов представляют собой группы, которые играют важную роль в математической физике. А именно, в следующих разделах с помощью алгебр Клиффорда мы построим группы Ли, которые являются накрытиями групп (псевдо)вращений $\mathcal{O}(p, q, \mathbb{F})$. Эти накрытия лежат в основе определения спинорных полей, которые служат для описания частиц с полужелым спином. В настоящем разделе мы напомним некоторые свойства групп вращений.

Рассмотрим (псевдо)евклидово пространство $\mathbb{R}^{p,q}$, $p + q = n \geq 2$, с декартовыми координатами $x^1, \dots, x^n$ и метрикой

$$\eta = \text{diag}(\underbrace{+ \dots +}_p \underbrace{- \dots -}_q). \quad (26.90)$$

Первые p координат будем называть *временными*, а последние q координат – *пространственными*.

В дальнейшем приставку “псевдо”, для краткости, мы часто будем опускать.

Определение. Вещественной группой вращений $\mathbb{O}(p, q) := \mathbb{O}(p, q, \mathbb{R})$ называется множество вещественных матриц S , сохраняющих метрику:

$$S_a^c S_b^d \eta_{cd} = \eta_{ab} \quad \Leftrightarrow \quad S \eta S^T = \eta \quad (26.91)$$

с обычным умножением матриц. $\square$

При $n \geq 3$ это множество матриц образует неабелеву группу Ли размерности $n(n-1)/2$. Группы $\mathbb{SO}(2)$ и $\mathbb{SO}_0(1, 1)$ являются абелевыми (см. разделы 1.6 и 1.7).

Предложение 26.6.1. *Существует изоморфизм $\mathbb{O}(p, q) \simeq \mathbb{O}(q, p)$. В частности, $\mathbb{O}(n, 0) \simeq \mathbb{O}(0, n)$.*

Доказательство. Определяющее соотношение (26.91) инвариантно относительно изменения знака метрики $\eta \mapsto -\eta$. Поэтому изоморфизм достигается простым изменением нумерации координат. $\square$

Обозначим через

$$\det_p S := \det \begin{pmatrix} S_1^1 & \dots & S_1^p \\ \vdots & \ddots & \vdots \\ S_p^1 & \dots & S_p^p \end{pmatrix}$$

– определитель верхнего левого $p \times p$ блока и

$$\det_q S := \det \begin{pmatrix} S_{p+1}^{p+1} & \dots & S_{p+1}^n \\ \vdots & \ddots & \vdots \\ S_n^{p+1} & \dots & S_n^n \end{pmatrix}$$

– определитель нижнего правого $q \times q$ блока матрицы S .

Предложение 26.6.2. *Пусть $S \in \mathbb{O}(p, q)$. Тогда выполнены неравенства*

$$|\det_p S| \geq 1, \quad |\det_q S| \geq 1.$$

Доказательство. См., например, [196], предложение 13.1. $\square$

С частным случаем этого утверждения мы уже встречались в разделе 1.9.1, где было доказано неравенство $|S_0^0| \geq 1$ для группы Лоренца $\mathbb{O}(1, 3)$.

Предложение 26.6.3. *Группы вращений $\mathbb{O}(n) := \mathbb{O}(n, 0, \mathbb{R}) \simeq \mathbb{O}(0, n, \mathbb{R})$, $n \geq 2$, состоят из двух компонент связности: связной компоненты единицы $\mathbb{SO}(n)$ (матрицы с единичным определителем, $\det S = 1$) и смежного класса $\mathbb{S}_-$ (матрицы с отрицательным определителем, $\det S = -1$). Если ввести оператор четности*

$$P : (x^1, \dots, x^{n-1}, x^n) \mapsto (x^1, \dots, x^{n-1}, -x^n), \quad P^2 = 1, \quad \det P = -1,$$

то множество элементов смежного класса представимо в виде

$$\mathbb{S}_- = \text{PSO}(n).$$

Группы вращений $\mathbb{O}(p, q) := \mathbb{O}(p, q, \mathbb{R})$, $p \geq 1$ и $q \geq 1$ состоят из четырех компонент связности: связной компоненты единицы $\mathbb{SO}_0(p, q)$ и трех смежных классов:

$$\begin{aligned}\mathbb{SO}_0(p, q) &:= \{S \in \mathbb{O}(p, q) : \det_p S \geq 1, \det_q S \geq 1\}, \\ \mathbb{S}_{+-} &:= \{S \in \mathbb{O}(p, q) : \det_p S \geq 1, \det_q S \leq -1\}, \\ \mathbb{S}_{-+} &:= \{S \in \mathbb{O}(p, q) : \det_p S \leq -1, \det_q S \geq 1\}, \\ \mathbb{S}_{--} &:= \{S \in \mathbb{O}(p, q) : \det_p S \leq -1, \det_q S \leq -1\},\end{aligned}$$

Если ввести операторы обращения времени и четности:

$$T : (x^1, x^2, \dots, x^n) \mapsto (-x^1, x^2, \dots, x^n), \quad \det_p T = -1, \det_q T = 1, \quad (26.92)$$

$$P : (x^1, \dots, x^{n-1}, x^n) \mapsto (x^1, \dots, x^{n-1}, -x^n), \quad \det_p P = 1, \det_q P = -1, \quad (26.93)$$

то смежные классы можно представить в виде

$$\begin{aligned}\mathbb{S}_{-+} &:= T\mathbb{SO}_0(p, q), \\ \mathbb{S}_{+-} &:= P\mathbb{SO}_0(p, q), \\ \mathbb{S}_{--} &:= PT\mathbb{SO}_0(p, q).\end{aligned}$$

Доказательство. См., например, [196], раздел 13.3. □

Матрицы обращения времени и четности принадлежат смежным классам:

$$T \in \mathbb{S}_{-+}, \quad P \in \mathbb{S}_{+-},$$

и меняют ориентацию, соответственно, временных и пространственных координат. Их совместное действие $PT \in \mathbb{S}_{--}$ меняет одновременно ориентацию и временных, и пространственных координат. Конечно, матрицы обращения времени и четности определены не однозначно. Вместо матриц (26.92) и (26.93) можно выбрать произвольных представителей из соответствующих смежных классов.

Связная компонента единицы $\mathbb{SO}_0(p, q)$ является нормальной подгруппой в группе вращений. При $p \geq 1$ и $q \geq 1$ определена фактор группа

$$\frac{\mathbb{O}(p, q)}{\mathbb{SO}_0(p, q)} = \mathbb{K}_4 \simeq \mathbb{Z}_2 \times \mathbb{Z}_2,$$

где $\mathbb{K}_4$ – группа Клейна, состоящая из четырех элементов (см., раздел 1.7).

Определение. Подгруппа $\mathbb{SO}_0(p, q)$ группы псевдовращений $\mathbb{O}(p, q)$ называется *собственной ортохронной группой* псевдовращений. Подгруппа

$$\begin{aligned}\mathbb{SO}(p, q) &= \{S \in \mathbb{O}(p, q) : \det S = 1\} = \\ &= \mathbb{SO}_0(p, q) \cup \mathbb{S}_{--}\end{aligned} \quad (26.94)$$

– *собственной группой* псевдовращений. Подгруппа

$$\begin{aligned}\mathbb{O}_{+-}(p, q) &= \{S \in \mathbb{O}(p, q) : \det S = 1, \det_p S \geq 1\} = \\ &= \mathbb{SO}_0(p, q) \cup \mathbb{S}_{+-}\end{aligned} \quad (26.95)$$

– *ортохронной группой* псевдовращений. Подгруппа

$$\begin{aligned}\mathbb{O}_{-+}(p, q) &= \{S \in \mathbb{O}(p, q) : \det S = 1, \det_q S \geq 0\} = \\ &= \mathbb{SO}_0(p, q) \cup \mathbb{S}_{-+}\end{aligned} \quad (26.96)$$

– *ортохронной группой* псевдовращений. □

Из определения следует, что собственные вращения сохраняют ориентацию всего псевдоевклидова пространства $\mathbb{R}^{p,q}$. Ортохронная группа вращений состоит из тех преобразований, которые сохраняют ориентацию временных координат. Ортохронные преобразования сохраняют ориентацию пространственных координат. Специальные ортохронные преобразования сохраняют ориентацию и временных, и пространственных координат.

Комплексные группы вращений действуют в комплексном псевдоевклидовом пространстве $\mathbb{C}^{p,q}$. Пусть z^a , $a = 1, \dots, n = p+q$ – декартовы координаты в $\mathbb{C}^{p,q}$, в которых задана псевдоевклидова метрика (26.90). Тогда комплексная группа вращений определяется тем же уравнением (26.91), что и в вещественном случае, только теперь матрицы S комплексные. Группа комплексных вращений $\mathbb{O}(p, q, \mathbb{C})$ является группой Ли в два раза большей вещественной размерности $\dim \mathbb{O}(p, q, \mathbb{C}) = n(n-1)$. Из определения (26.91) следует равенство

$$|\det S| = 1.$$

То есть, как и в вещественном случае, возможны два случая $\det S = \pm 1$. Матрицы с единичным определителем образуют собственную подгруппу в группе вращений $\mathbb{SO}(p, q, \mathbb{C}) \subset \mathbb{O}(p, q, \mathbb{C})$.

В случае поля вещественных чисел группа вращений была определена в разделе 8.14.2. Напомним, что унитарная группа, которая была определена в том же разделе, определяется соотношением

$$S\eta S^\dagger = \eta. \quad (26.97)$$

Отсюда следует, что, если матрица $S \in \mathbf{mat}(n, \mathbb{C})$ удовлетворяет сразу двум условиям (26.91) и (26.97), то она вещественна. Следовательно, справедливы равенства

$$\mathbb{O}(p, q, \mathbb{R}) := \mathbb{O}(p, q, \mathbb{C}) \cap \mathbb{GL}(p+q, \mathbb{R}) = \mathbb{O}(p, q, \mathbb{C}) \cap \mathbb{U}(p, q).$$

Напомним размерности рассматриваемых групп Ли:

$$\begin{aligned} \dim \mathbb{O}(p, q, \mathbb{R}) &= \frac{1}{2}(p+q)(p+q-1), \\ \dim \mathbb{O}(p, q, \mathbb{C}) &= (p+q)(p+q-1), \\ \dim \mathbb{U}(p, q) &= (p+q)^2. \end{aligned}$$

26.7 Накрытия групп вращений

В настоящем разделе мы покажем, что в алгебрах Клиффорда можно выделить часть элементов, которые образуют группу по умножению. Эта группа имеет много подгрупп, имеющих важное значение для приложений. В частности, как будет показано ниже, группы $\mathbb{PIN}$ и ее подгруппы накрывают группы вращений в (псевдо)евклидовом пространстве. Сначала приведем

Пример 26.7.1. Рассмотрим восьмимерную вещественную алгебру Клиффорда $\mathbf{cl}(3, 0, \mathbb{R})$ с образующими e^a , $a = 1, 2, 3$. Эта алгебра порождена обычным трехмерным евклидовым пространством $\mathbb{R}^3$, и образующие e^a можно рассматривать в качестве ортонормированных базисных векторов декартовой системы координат.

Рассмотрим вращение в плоскости, натянутой на векторы e^1, e^2 на угол ω . При этом базисные векторы e^1 и e^2 повернутся:

$$e^2 \mapsto e^b S_b^a = (e^1, e^2) \begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix} = (e^1 \cos \omega + e^2 \sin \omega, -e^1 \sin \omega + e^2 \cos \omega).$$

Поскольку в алгебре Клиффорда выполнено равенство

$$(e^1 e^2)^2 := e^1 e^2 e^1 e^2 = -1,$$

то базисный вектор e^{12} играет роль мнимой единицы. Поэтому вращение в плоскости e^1, e^2 можно переписать, используя клиффордово умножение:

$$\begin{aligned} e^1 &\mapsto e^1 \cos \omega + e^1 e^1 e^2 \sin \omega = e^1 (\cos \omega + e^1 e^2 \sin \omega) = e^1 e^{e^1 e^2 \omega}, \\ e^2 &\mapsto -e^2 e^2 e^1 \sin \omega + e^2 \cos \omega = e^2 (e^1 e^2 \sin \omega + \cos \omega) = e^2 e^{e^1 e^2 \omega}. \end{aligned}$$

Вращение базисных векторов можно записать в виде

$$\begin{aligned} e^1 e^{e^1 e^2 \omega} &= e^1 e^{e^1 e^2 \frac{\omega}{2}} e^{e^1 e^2 \frac{\omega}{2}} = e^1 \left(\cos \frac{\omega}{2} + e^1 e^2 \sin \frac{\omega}{2} \right) e^{e^1 e^2 \frac{\omega}{2}} = \\ &= \left(\cos \frac{\omega}{2} - e^1 e^2 \sin \frac{\omega}{2} \right) e^1 e^{e^1 e^2 \frac{\omega}{2}} = e^{-e^1 e^2 \frac{\omega}{2}} e^1 e^{e^1 e^2 \frac{\omega}{2}}. \end{aligned}$$

Точно так же преобразуется второй базисный вектор

$$e^2 e^{e^1 e^2 \omega} = e^{-e^1 e^2 \frac{\omega}{2}} e^2 e^{e^1 e^2 \frac{\omega}{2}}.$$

Кроме того, базисный вектор e^3 , который перпендикулярен оси вращения, не преобразуется:

$$e^3 \mapsto e^3 = e^{-e^1 e^2 \frac{\omega}{2}} e^3 e^{e^1 e^2 \frac{\omega}{2}},$$

т.к. $e^3 e^1 e^2 = e^1 e^2 e^3$. По линейности, любое вращение в плоскости e^1, e^2 произвольного вектора $A = e^a A_a$ можно записать в виде клиффордова умножения

$$A \mapsto e^b S_b^a A_a = U A U^{-1},$$

где $U := e^{-e^1 e^2 \frac{\omega}{2}} \in \mathbf{cl}_2(3, 0, \mathbb{R})$.

Поскольку каждое вращение в $\mathbb{R}^3$ можно представить в виде трех вращений вокруг осей e^1, e^2 и e^3 , то произвольное вращение евклидова пространства можно представить в виде

$$A \mapsto A S = U A U^{-1},$$

где $U \in \mathbf{cl}_2(3, 0, \mathbb{R})$ – некоторый обратимый элемент алгебры Клиффорда второй степени. В дальнейшем мы увидим, что элемент алгебры Клиффорда U принадлежит группе $\mathbf{SPIN}(3)$, которая будет определена безотносительно к группе вращений. Поскольку двум элементам алгебры Клиффорда $\pm U \in \mathbf{SPIN}(3, 0, \mathbb{R})$ соответствует одна матрица вращения $S \in \mathbf{SO}(3)$, то гомоморфизм $\mathbf{SPIN}(3, 0, \mathbb{R}) \rightarrow \mathbf{SO}(3)$, как будет показано ниже, является двулистным накрытием. $\square$

Перейдем к общей конструкции. Сначала будут определены группы $\mathbf{PIN}(p, q, \mathbb{F})$, $\mathbf{SPIN}(p, q, \mathbb{F})$ и $\mathbf{SPIN}_0(p, q, \mathbb{F})$, а затем будет доказано, что они являются двулистными накрытиями групп вращений.

В алгебре Клиффорда общего вида $\mathbf{cl}(\mathbb{V}, Q)$, как и в любой другой ассоциативной алгебре, можно ввести структуру алгебры Ли, определив коммутатор двух произвольных элементов:

$$[U, V] := UV - VU. \quad (26.98)$$

При этом ассоциативность достаточна для выполнения тождеств Якоби, которые содержат клиффордовы произведения трех элементов. Из контекста всегда ясно,

когда алгебра Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$ рассматривается как алгебра Ли. Поэтому для этой алгебры Ли мы не будем вводить нового обозначения.

Поскольку справедливо равенство

$$[e^a, e^b] = 2e^a e^b - 2\eta^{ab}e,$$

то сами образующие алгебры Клиффорда никакой алгебры Ли не генерируют.

Теперь выделим в алгебре Ли множество элементов, которые образуют группу по отношению к клиффордову умножению. Поскольку любая алгебра Клиффорда изоморфна некоторой матричной алгебре, то ясно, что часть элементов алгебры Клиффорда будет обратимой, т.е. для каждого обратимого элемента U в алгебре $\mathfrak{cl}(\mathbb{V}, Q)$ существует такой элемент U^{-1} , что выполнены равенства:

$$UU^{-1} = U^{-1}U = 1.$$

Обратимым элементам и только им соответствуют матрицы с отличным от нуля определителем. Если матрица представления имеет равный нулю определитель, то она соответствует необратимому элементу алгебры Клиффорда.

Пример 26.7.2. В примере 26.2.5 было доказано, что произвольный элемент алгебры Клиффорда $\mathfrak{cl}(1, 1, \mathbb{F})$ имеет точное матричное представление:

$$U = eu_0 + e^1u_1 + e^2u_2 + e^{12}u_{12} = \begin{pmatrix} u_0 + u_1 & u_2 + u_{12} \\ -u_2 + u_{12} & u_0 - u_1 \end{pmatrix}.$$

Условие

$$\det U = u_0^2 - u_1^2 + u_2^2 - u_{12}^2 = 0,$$

выделяет в четырехмерном многообразии $(u_0, u_1, u_2, u_{12}) \in \mathbb{F}^4$ трехмерное многообразие необратимых элементов. То, что останется – это открытое подмногообразие обратимых элементов, которое четырехмерно. $\square$

В дальнейшем исходное векторное пространство $\mathbb{V} = \mathfrak{cl}_1 \subset \mathfrak{cl}$ будет играть важную роль. Чтобы подчеркнуть принадлежность какого-либо элемента алгебры Клиффорда этому подпространству будем использовать для него буквы из начала алфавита $A, B, \dots$, так же, как и для 1-форм.

Пример 26.7.3. Рассмотрим элемент алгебры Клиффорда степени 1, $A = e^a A_a \in \mathfrak{cl}_1 = \mathbb{V}$, такой, что его длина отлична от нуля, $Q(A) := \eta^{ab} A_a A_b \neq 0$. Тогда у него существует обратный элемент, который имеет вид

$$A^{-1} = \frac{A}{Q(A)}. \quad \square$$

Пример 26.7.4. Рассмотрим внешнюю алгебру $\mathfrak{cl}(\mathbb{V}, 0) = \Lambda(\mathbb{V})$. Поскольку длина любого вектора равна нулю, то обратимых элементов первой степени во внешней алгебре не существует. Среди внешних форм не выше второй степени обратимые элементы есть. Например, нетрудно проверить, что для формы $e + e^1 e^2 \in \Lambda(\mathbb{V})$ обратным будет элемент $e - e^1 e^2 \in \Lambda(\mathbb{V})$. $\square$

Предложение 26.7.1. Множество обратимых элементов алгебры Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$ образует группу. Эта группа является группой Ли размерности 2^n , которую мы обозначим $\mathbb{CL}(\mathbb{V}, Q)$. Алгеброй Ли группы Ли $\mathbb{CL}(\mathbb{V}, Q)$ является алгебра Клиффорда $\mathfrak{cl}(\mathbb{V}, Q)$, рассматриваемая как алгебра Ли с коммутатором (26.98).

Доказательство. Прямая проверка. $\square$

Все группы, рассмотренные дальше в этом разделе, являются подгруппами группы $\mathbb{CL}(\mathbb{V}, Q)$.

Группа Ли $\mathbb{CL}(\mathbb{V}, Q)$ не является векторным пространством. Она является открытым подмногообразием в 2^n -мерном векторном пространстве $\mathfrak{cl}(\mathbb{V}, Q)$ с естественной дифференцируемой структурой. Поэтому можно говорить о том, что некоторый элемент группы принадлежит одному из подпространств определенной степени.

Поскольку произведение четных элементов является четным и единица является четным элементом, то в группе $\mathbb{CL}$ есть подгруппа $\mathbb{CL}_+$, состоящая из четных элементов.

Для конечномерных векторных пространств $\mathbb{V}$ над полем $\mathbb{R}$ или $\mathbb{C}$ существует естественное экспоненциальное отображение алгебры Ли $\mathfrak{cl}$ в группу Ли $\mathbb{CL}$:

$$\mathfrak{cl} \ni U \mapsto \exp U := \sum_{k=0}^{\infty} \frac{U^k}{k!} \in \mathbb{CL}.$$

Этот ряд равномерно сходится в любой ограниченной области значений компонент элементов алгебры Клиффорда, так же, как и ряд от матрицы.

Предложение 26.7.2. *Справедливы равенства*

$$e^U e^{-U} = e^{-U} e^U = e \quad (26.99)$$

для всех элементов $U \in \mathfrak{cl}(\mathbb{V}, Q)$.

Доказательство. Следствие равенства $[U, U] = 0$. $\square$

То есть, если некоторому элементу группы Ли $\mathbb{CL}$ соответствует элемент алгебры Ли $U \in \mathfrak{cl}$, то обратному элементу соответствует противоположный элемент (вектор) $-U$. Элементы группы мы будем обозначать латинскими буквами $a, b, \dots \in \mathbb{CL}$.

Определение. Группа Ли $\mathbb{CL}(\mathbb{V}, Q)$ порождает естественное действие в своей алгебре Ли $\mathfrak{cl}(\mathbb{V}, Q)$ с помощью автоморфизмов

$$\text{Ad} : \mathbb{CL} \times \mathfrak{cl} \ni a, U \mapsto \text{Ad}(a)U := aUa^{-1} \in \mathfrak{cl}, \quad (26.100)$$

которое называется *присоединенным представлением*. $\square$

Если элемент группы a умножить на произвольный отличный от нуля элемент поля $t \in \mathbb{F}^\times$, то из определения вытекает равенство

$$\text{Ad}(ta)U = \text{Ad}(a)U.$$

Взятие “производной” от определения (26.100) приводит к гомоморфизму алгебры Клиффорда в алгебру дифференцирований этой алгебры:

$$\mathfrak{cl} \times \mathfrak{cl} \ni V, U \mapsto \text{ad}_V U := [V, U] \in \mathfrak{cl}.$$

Действительно, для элементов группы, близких к единичному, справедливо представление $a = e + tV + o(t)$, где $t \in \mathbb{F}$ и $V \in \mathfrak{cl}$ – элемент алгебры Ли. Поэтому

$$\text{Ad}(e + tV + o(t))U = (e + tV + o(t))U(e - tV + o(t)) = U + t[V, U] + o(t).$$

Отсюда вытекает равенство

$$\frac{d}{dt} \text{Ad}(e + tV)U \Big|_{t=0} = \text{ad}_V U.$$

Предложение 26.7.3. Пусть $A \in \mathbb{V} = \mathfrak{cl}_1$ – обратимый элемент алгебры Клиффорда, рассматриваемой как алгебра Ли. Тогда для произвольного элемента первой степени $B \in \mathbb{V}$ справедливо равенство

$$\text{Ad}(A)B = -B + 2\frac{Q(A, B)}{Q(A)}A. \quad (26.101)$$

Доказательство. Поскольку $A^{-1} = A/Q(A)$, то справедливы следующие равенства:

$$\begin{aligned} Q(A) \text{Ad}(A)B &= Q(A)ABA^{-1} = ABA = \\ &= -AAB + 2Q(A, B)A = -Q(A)B + 2Q(A, B)A. \quad \square \end{aligned}$$

Следствие. Подпространство элементов первой степени $\mathbb{V} = \mathfrak{cl}_1$ инвариантно относительно присоединенного действия своих элементов:

$$\text{Ad}(A)\mathbb{V} = \mathbb{V}$$

для всех обратимых элементов $A \in \mathbb{V}$. $\square$

Доказательство. Включение $\text{Ad}(A)\mathbb{V} \subset \mathbb{V}$ очевидно. Для каждого обратимого элемента A условие $ABA^{-1} = C$ эквивалентно равенству $B = A^{-1}CA$. Отсюда следует сюръективность отображения (26.101). $\square$

Предложение 26.7.4. Действие $\text{Ad}(A)$ обратимого элемента $A \in \mathbb{V}$ на произвольный вектор $B \in \mathbb{V}$ сохраняет его длину:

$$Q(\text{Ad}(A)B) = Q(B). \quad (26.102)$$

Доказательство. Из равенства (26.101) вытекают следующие соотношения:

$$\begin{aligned} Q(\text{Ad}(A)B) &= Q\left(-B + 2\frac{Q(A, B)}{Q(A)}A\right) = \\ &= \eta^{ab} \left(-B_a + 2\frac{Q(A, B)}{Q(A)}A_a\right) \left(-B_b + 2\frac{Q(A, B)}{Q(A)}A_b\right) = \\ &= \eta^{ab} B_a B_b - 4\frac{Q^2(A, B)}{Q(A)} + 4\frac{Q^2(A, B)}{Q(A)} = Q(B). \quad \square \end{aligned}$$

Последнее утверждение указывает на возможность рассмотрения тех обратимых элементов алгебры Клиффорда $a \in \mathbb{CL}$, которые сохраняют векторное подпространство: $\text{Ad}(a)\mathbb{V} = \mathbb{V}$. Ясно, что это множество элементов образует группу. Из следствия предложения 26.7.3 вытекает, что данная группа содержит все обратимые элементы из $\mathbb{V}$.

Определение. Группой $\mathbb{P}(\mathbb{V}, Q)$ называется подгруппа в $\mathbb{CL}(\mathbb{V}, Q)$, которая свободно порождена всеми обратимыми элементами первой степени:

$$\mathbb{P}(\mathbb{V}, Q) := \{A_1 \dots A_r \in \mathbb{CL}(\mathbb{V}, Q) : A_1, \dots, A_r \in \mathbb{V}\}, \quad (26.103)$$

где $A_1 \dots A_r$ – клиффордово произведение произвольного, но конечного числа r обратимых векторов.

Группой $\text{PIN}(\mathbb{V}, Q)$ векторного пространства $\mathbb{V}$ с квадратичной формой Q называется подгруппа в $\mathbb{P}(\mathbb{V}, Q)$, которая свободно порождена векторами $A \in \mathbb{V}$ единичной длины, $Q(A) = \pm 1$:

$$\text{PIN}(\mathbb{V}, Q) := \{A_1 \dots A_r \in \mathbb{CL}(\mathbb{V}, Q) : A_1, \dots, A_r \in \mathbb{V}, Q(A_k) = \pm 1\}. \quad (26.104)$$

Группой $\text{SPIN}(\mathbb{V}, Q)$ называется подгруппа в группе $\text{PIN}(\mathbb{V}, Q)$, состоящая из четных элементов:

$$\text{SPIN}(\mathbb{V}, Q) := \text{PIN}(\mathbb{V}, Q) \cap \mathbb{CL}_+(\mathbb{V}, Q). \quad \square \quad (26.105)$$

Из определения следует, что любой элемент группы $\mathbb{P}(\mathbb{V}, Q)$ представляет собой конечное произведение векторов (но не их конечные линейные комбинации). Поэтому произвольный элемент группы $\mathbb{P}(\mathbb{V}, Q)$ является либо четным, либо нечетным. Кроме того, группа $\mathbb{P}(\mathbb{V}, Q)$ является минимальной группой, содержащей все обратимые векторы.

Можно записать равенство

$$\mathbb{P}(\mathbb{V}, Q) = \mathbb{F}^\times \text{PIN}(\mathbb{V}, Q),$$

т.е. каждый элемент группы $\mathbb{P}$ является произведением некоторого элемента группы PIN на отличное от нуля число.

В следующем разделе мы рассмотрим несколько примеров явного описания групп $\mathbb{P}(\mathbb{V}, Q)$, $\text{PIN}(\mathbb{V}, Q)$ и $\text{SPIN}(\mathbb{V}, Q)$ с невырожденной квадратичной формой в низших размерностях.

Поскольку каждый из сомножителей в $A_1 \dots A_r$ сохраняет длину вектора при присоединенном действии, то каждый элемент группы $\mathbb{P}(\mathbb{V}, Q)$ сохраняет квадратичную форму Q . Зафиксируем элемент группы $a \in \mathbb{P}(\mathbb{V}, Q)$. Тогда для каждого вектора A существует отображение

$$\mathbb{V} \ni A \mapsto \text{Ad}(a)A := aAa^{-1} \in \mathbb{V}.$$

Это значит, что определено отображение

$$\mathbb{P}(\mathbb{V}, Q) \ni a \mapsto S \in \mathbb{O}(\mathbb{V}, Q),$$

группы $\mathbb{P}(\mathbb{V}, Q)$ в группу вращений векторного пространства $\mathbb{V}$, сохраняющую квадратичную форму:

$$\mathbb{O}(\mathbb{V}, Q) := \{S \in \mathbb{GL}(\mathbb{V}) : S^*Q = Q\},$$

где S^* – возврат отображения $S : \mathbb{V} \rightarrow \mathbb{V}$. В компонентах каждому элементу $a \in \mathbb{P}(\mathbb{V}, Q)$ мы ставим в соответствие некоторую матрицу вращения $S = (S_a^b)$:

$$a : A = e^a A_a \mapsto SA = e^a S_a^b(a) A_b.$$

Поскольку для всех векторов выполнены равенства

$$\text{Ad}(a_1 a_2)A = S_{12}A = a_1 a_2 A a_2^{-1} a_1^{-1} = a_1 (S_2 A) a_1^{-1} = S_1 S_2 A,$$

то присоединенное представление осуществляет гомоморфизм групп

$$\mathbb{P}(\mathbb{V}, Q) \xrightarrow{\text{Ad}} \mathbb{O}(\mathbb{V}, Q).$$

Если квадратичная форма Q невырождена, то отображение (26.101) представляет собой отражение вектора B в гиперплоскости, перпендикулярной вектору A (все

компоненты вектора, перпендикулярные A , меняют знак). Если размерность векторного пространства n четна, то это отображение меняет ориентацию, а если нечетна, то сохраняет.

В дальнейшем мы хотим построить накрытия групп вращений $\mathbb{O}(p, q)$, которые не являются связными, и переход между компонентами связности осуществляется, в том числе, с помощью преобразования четности, меняющим ориентацию (см. раздел 1.6). Для того, чтобы достичь желаемого результата, введем новое отображение (автоморфизм), которое меняет ориентацию векторного пространства в любой размерности,

$$\widehat{\text{Ad}} : \mathbb{CL} \times \mathfrak{cl} \ni a, U \mapsto \widehat{\text{Ad}}(a)U := aU\hat{a}^{-1} \in \mathfrak{cl} \quad (26.106)$$

для всех обратимых элементов $a \in \mathbb{CL}(\mathbb{V}, Q)$, где шляпка обозначает инволюцию градуировки (26.79).

Из предложения (26.5.1) следует, что введенный автоморфизм так же, как и присоединенное представление, является представлением группы $\mathbb{CL}$:

$$\widehat{\text{Ad}}(ab) = \widehat{\text{Ad}}(a) \circ \widehat{\text{Ad}}(b).$$

Кроме того, $\widehat{\text{Ad}}(a) = \text{Ad}(a)$ для всех четных элементов $a \in \mathbb{CL}_+$.

Если сузить отображение $\widehat{\text{Ad}}(A)$ на векторное пространство обратимых элементов $A \in \mathbb{V} \subset \mathbb{CL}$, то этот автоморфизм будет отличаться от присоединенного представления $\text{Ad}(A)$ только знаком. Поэтому так же, как и присоединенное представление, автоморфизм $\widehat{\text{Ad}}(A)$ сохраняет квадратичную форму и является сюръективным отображением $\mathbb{V} \rightarrow \mathbb{V}$.

Поскольку

$$Q(\hat{A}) = Q(A), \quad A\hat{A} = -Q(A)e \quad \text{и} \quad Q(\hat{A}, B) = -Q(A, B)$$

для всех $A, B \in \mathbb{V}$, то из предложения 26.7.4 следует

Предложение 26.7.5. Пусть $A \in \mathbb{V}$ – обратимый элемент алгебры Клиффорда. Тогда для произвольного элемента первой степени $B \in \mathbb{V}$ справедливо равенство

$$\widehat{\text{Ad}}(A)B = B - 2\frac{Q(A, B)}{Q(A)}A. \quad (26.107)$$

Отсюда вытекает, что при отображении $\widehat{\text{Ad}}(A)$ отражается только та компонента вектора B , которая параллельна A , и поэтому ориентация векторного пространства $\mathbb{V}$ меняется на противоположную в любом числе измерений. В частности, для любого обратимого вектора

$$\widehat{\text{Ad}}(A)A = -A.$$

Определение. Группой Липшица $\widehat{\mathbb{P}}(\mathbb{V}, Q)$ называется множество обратимых элементов алгебры Клиффорда, которые сохраняют векторное подпространство при автоморфизме $\widehat{\text{Ad}}$:

$$\widehat{\mathbb{P}}(\mathbb{V}, Q) := \{a \in \mathbb{CL}(\mathbb{V}, Q) : \widehat{\text{Ad}}(a)\mathbb{V} := a\mathbb{V}\hat{a}^{-1} = \mathbb{V}\}. \quad (26.108)$$

Группу Липшица называют также *группой Клиффорда*. □

Поскольку инволюция градуировки меняет только знак нечетного произведения векторов, то $\mathbb{P}(\mathbb{V}, Q) \subseteq \widehat{\mathbb{P}}(\mathbb{V}, Q)$. Можно доказать более сильное утверждение.

Предложение 26.7.6. Рассмотрим вещественное векторное пространство $\mathbb{V}$ с невырожденной квадратичной формой Q сигнатуры (p, q) . Тогда

$$\widehat{\mathbb{P}}(p, q, \mathbb{R}) = \mathbb{P}(p, q, \mathbb{R}).$$

Доказательство. См., например, [193], равенство (2.33). $\square$

Это предложение проясняет строение элементов группы Липшица. В частности, любой элемент группы Липшица представляет собой конечное произведение векторов и либо четен, либо нечетен.

Опишем важные свойства группы Липшица.

Предложение 26.7.7. Предположим, что квадратичная форма Q невырождена и имеет сигнатуру (p, q) . Тогда ядро гомоморфизма

$$\widehat{\mathbb{P}}(p, q, \mathbb{F}) \xrightarrow{\widehat{\text{Ad}}} \mathbb{O}(p, q, \mathbb{F})$$

представляет собой группу $\mathbb{F}^\times$ (множество ненулевых элементов поля $\mathbb{F}$, рассматриваемое, как группа по умножению).

Доказательство. Выберем ортонормированный базис e^a в $\mathbb{V}$. Пусть элемент $a \in \mathbb{CL}$ лежит в ядре отображения $\widehat{\text{Ad}}$, т.е.

$$aU\hat{a}^{-1} = U \quad \Leftrightarrow \quad aU = U\hat{a} \quad \forall U \in \mathfrak{cl}.$$

Представим $a = a_+ + a_-$. Тогда для всех $A \in \mathbb{V}$ должны быть выполнены равенства

$$a_+A = Aa_+, \quad a_-A = -Aa_-. \quad (26.109)$$

Элементы a_+ и a_- являются полиномами по базисным векторам e^a . Поэтому, используя определяющее соотношение (26.15), четную часть a_+ можно представить в виде $a_+ = b_+ + e^1b_-$, где элементы b_+ и b_- не зависят от e^1 . Положим в первом равенстве (26.109) $A = e^1$. Тогда выполнены соотношения

$$(b_+ + e^1b_-)e^1 = e^1(b_+ + e^1b_-) \quad \Leftrightarrow \quad -(e^1)^2b_- = (e^1)^2b_-.$$

Откуда следует, что $b_- = 0$. Поэтому элемент a_+ не зависит от e^1 . Повторяя эту процедуру, мы видим, что элемент a_+ не зависит от всех e^a . Следовательно, он имеет вид et , где $t \in \mathbb{F}$.

Применим аналогичное рассуждение к элементу a_- . Представим его в виде $a_- = c_- + e^1c_+$, где элементы c_- и c_+ не зависят от e^1 . Подстановка этого представления в равенство (26.109) при $A = e^1$ приводит к соотношениям:

$$(c_- + e^1c_+)e^1 = -e^1(c_- + e^1c_+) \quad \Leftrightarrow \quad (e^1)^2c_+ = -(e^1)^2c_+.$$

Отсюда следует, что $c_+ = 0$. Поэтому элемент a_- не зависит от e^1 . Затем по индукции получаем, что a_- не зависит также от $e^2, \dots, e^n$. Следовательно, $a_- = 0$.

Таким образом, $a = a_+ + a_- = et$. Поскольку $a \neq 0$, то $a \in \mathbb{F}^\times$. $\square$

Замечание. Инволюция градуировки в определении автоморфизма $\widehat{\text{Ad}}$ (26.106) существенна, т.к. именно она привела к появлению знака минус во втором равенстве (26.109), который является решающим.

Предположение о невырожденности квадратичной формы Q также является существенным. Для этого приведем

Пример 26.7.5. Рассмотрим внешнюю алгебру $\mathfrak{cl}(\mathbb{V}, 0) = \Lambda(\mathbb{V})$. Для двух базисных векторов $e^1, e^2 \in \mathbb{V}$ элемент внешней алгебры $e + e^1 e^2 \in \mathbb{CL}(\mathbb{V}, 0)$ обратим. Его обратный имеет вид $(e + e^1 e^2)^{-1} = e - e^1 e^2$. С другой стороны, для любого вектора $A \in \mathbb{V}$ выполнено равенство

$$(e + e^1 e^2)A(\widehat{e + e^1 e^2})^{-1} = (e + e^1 e^2)A(e - e^1 e^2) = A.$$

Поэтому ядро гомоморфизма $\widehat{\text{Ad}}$ включает много элементов, отличных от скаляров. $\square$

Каждый элемент группы обратимых элементов $a \in \mathbb{CL}$ определяет отображение (26.106). Определим *норму отображения* следующим равенством

$$N(\widehat{\text{Ad}}(a)) := N(a) := a\hat{a}. \quad (26.110)$$

Легко проверить, что $\hat{a} = \tilde{a}$ и

$$N(A) = -Q(A), \quad \forall A \in \mathbb{V}. \quad (26.111)$$

Название “норма” в данном случае не соответствует общему определению нормы, данному в разделе 1.3.3. Действительно, определение даже не гарантирует того, что в общем случае норма принимает значения в поле вещественных чисел $\mathbb{R}$.

Важность понятия нормы отображения дает следующее

Предложение 26.7.8. Пусть квадратичная форма Q невырождена. Тогда сужение нормы (26.110) на элементы группы Липшица $\widehat{\mathbb{P}}(p, q, \mathbb{F})$ является гомоморфизмом

$$N : \widehat{\mathbb{P}}(p, q, \mathbb{F}) \ni a \mapsto N(a) := a\hat{a} \in \mathbb{F}^\times \quad (26.112)$$

в мультипликативную группу основного поля $\mathbb{F}^\times$.

Доказательство. Сначала докажем включение $N(\widehat{\mathbb{P}}(p, q, \mathbb{F})) \subset \mathbb{F}^\times$, т.е. что норма любого элемента из группы $\widehat{\mathbb{P}}$ является отличным от нуля числом. По определению, $aA\hat{a}^{-1} \in \mathbb{V}$. Поскольку операция реверса не меняет векторов, то взятие реверса от выражения $aA\hat{a}^{-1}$ приводит к соотношениям

$$\widetilde{aA\hat{a}^{-1}} = \tilde{a}^{-1}A\tilde{a} = aA\hat{a}^{-1}.$$

Умножим это равенство слева на $\tilde{a}$ и справа – на $\tilde{a}^{-1}$:

$$\tilde{a}aA\hat{a}^{-1}\tilde{a}^{-1} = \tilde{a}aA(\widehat{\tilde{a}a})^{-1} = \widehat{\text{Ad}}(\tilde{a}a)A = A.$$

Тогда из предложения 26.7.7 следует, что элемент $\tilde{a}a$ лежит в ядре отображения $\widehat{\text{Ad}}$, т.е. в $\mathbb{F}^\times$. Поэтому $\tilde{a}a = a\hat{a} \in \mathbb{F}^\times$. Инволюция градуировки приводит к включению $\tilde{a}\hat{a} = N(\tilde{a}) \in \mathbb{F}^\times$. Поскольку реверс сохраняет группу $\widehat{\mathbb{P}}(p, q, \mathbb{F})$, то отсюда вытекает, что $N(a) \in \mathbb{F}^\times$ для всех $a \in \widehat{\mathbb{P}}(p, q, \mathbb{F})$.

Для произвольных элементов $a, b \in \widehat{\mathbb{P}}(\mathbb{V}, Q)$ справедливо равенство $N(ab) = ab\hat{a}\hat{b} = ab\hat{b}\hat{a} = N(a)N(b)$. Поэтому отображение N является гомоморфизмом. $\square$

Следствие. Отображение $\widehat{\text{Ad}}(a) : \mathbb{V} \rightarrow \mathbb{V}$ для всех $a \in \widehat{\mathbb{P}}(p, q, \mathbb{F})$ сохраняет квадратичную форму Q . Поэтому существует гомоморфизм

$$\widehat{\text{Ad}} : \widehat{\mathbb{P}}(p, q, \mathbb{F}) \rightarrow \mathcal{O}(p, q, \mathbb{F}). \quad \square \quad (26.113)$$

Доказательство. Справедливы равенства:

$$N(\hat{a}) = \hat{a}\tilde{a} = \widehat{N(a)} = N(a).$$

Множество обратимых векторов мы обозначили через $\mathbb{CL}_1$. Очевидно, что $\mathbb{CL}_1 \subset \widehat{\mathbb{P}}(p, q, \mathbb{F})$. Тогда для каждого $A \in \mathbb{CL}_1$ справедливы равенства:

$$N(\widehat{\text{Ad}}(a)A) = N(aA\hat{a}^{-1}) = N(a)N(A)N^{-1}(\hat{a}) = N(A).$$

Поскольку $N(A) = -AA = -Q(A)$, то отображение $\widehat{\text{Ad}}(a)$ сохраняет длину всех векторов из $\mathbb{CL}_1$. Применение обратного отображения $\widehat{\text{Ad}}(a^{-1})$ показывает, что $\widehat{\text{Ad}}(a)\mathbb{CL}_1 = \mathbb{CL}_1$, и отображение $\widehat{\text{Ad}}(a)$ оставляет инвариантным подпространство векторов нулевой длины. Следовательно, существует гомоморфизм (26.113). $\square$

Вернемся к группе $\mathbb{P}(\mathbb{V}, Q)$. Поскольку каждый элемент этой группы является конечным произведением элементов $A_1, \dots, A_k$ из $\mathbb{CL}_1$, то действие $\widehat{\text{Ad}} : \mathbb{P}(p, q, \mathbb{F}) \rightarrow \mathbb{O}(p, q, \mathbb{F})$ на векторы из $\mathbb{V}$ сводится к последовательности преобразований

$$\widehat{\text{Ad}}(A_1 \dots A_k) = \rho_{A_1} \dots \rho_{A_k},$$

где каждое преобразование

$$\rho_A := \widehat{\text{Ad}}(A)$$

имеет вид (26.107) и представляет собой отражение вдоль вектора A . Следовательно, образ группы $\mathbb{P}(p, q, \mathbb{F})$ при гомоморфизме (26.113) совпадает со всей группой вращений согласно следующему классическому утверждению.

Теорема 26.7.1 (Картан–Дьюдонé). Пусть Q – невырожденная симметричная квадратичная форма сигнатуры (p, q) на конечномерном векторном пространстве $\mathbb{V}$. Тогда каждое (псевдо)вращение из группы $S \in \mathbb{O}(p, q, \mathbb{F})$ может быть записано в виде конечного произведения отражений:

$$S = \rho_{A_1} \dots \rho_{A_k}, \quad (26.114)$$

где $k \leq \dim \mathbb{V}$.

Доказательство. См., например, [199], теорема 3.20. $\square$

Таким образом, из теоремы Картана–Дьюдонé следует, что гомоморфизм

$$\widehat{\text{Ad}} : \mathbb{P}(p, q, \mathbb{F}) \rightarrow \mathbb{O}(p, q, \mathbb{F})$$

сюрьективен.

Теперь рассмотрим группу

$$\mathbb{P}_{\text{even}}(p, q, \mathbb{F}) := \mathbb{P}(p, q, \mathbb{F}) \cap \mathfrak{cl}_+(p, q, \mathbb{F}), \quad (26.115)$$

состоящую только из четных элементов.

Предложение 26.7.9. Гомоморфизм

$$\widehat{\text{Ad}} : \mathbb{P}_{\text{even}}(p, q, \mathbb{F}) \rightarrow \mathbb{SO}(p, q, \mathbb{F})$$

сюрьективен.

Доказательство. Сначала докажем, что $\det \rho_A = -1$. Для этого выберем ортогональный базис e^a такой, что первый вектор $e^1 = A$ и $Q(A, e^i) = 0$ для $i \geq 2$. Тогда из формулы (26.107) следуют равенства:

$$\rho_A e^1 = -e^1, \quad \rho_A e^i = e^i, \quad i \geq 2.$$

Поэтому $\det \rho_A = -1$. Поскольку матрица вращений имеет вид (26.114), и число элементов A четно, то $\det S = 1$. Сюръективность гомоморфизма следует из теоремы Картана–Дьюдоне. $\square$

Теперь рассмотрим группы PIN и SPIN . Любой элемент группы PIN представляет собой конечное произведение векторов (26.103) единичной длины, $Q(A) = \pm 1$. Согласно теореме Картана–Дьюдоне каждый элемент S группы вращений $\mathcal{O}(p, q, \mathbb{F})$ представляет конечное произведение матриц ρ_A (26.114). Из формулы (26.107) следует равенство

$$\rho_{tA} = \rho_A, \quad \forall t \in \mathbb{F}^\times.$$

Поэтому каждый вектор в произведении (26.114) можно отнормировать. Так как $Q(tA) = t^2 Q(A)$, то для нормировки на единицу уравнение $t^2 Q(A) = \pm 1$ должно быть разрешимо. Это всегда так для вещественных $\mathbb{R}$ и комплексных $\mathbb{C}$ чисел. Отсюда следует сюръективность отображений:

$$\begin{aligned} \widehat{\text{Ad}} : \text{PIN}(p, q, \mathbb{F}) &\rightarrow \mathcal{O}(p, q, \mathbb{F}), \\ \widehat{\text{Ad}} : \text{SPIN}(p, q, \mathbb{F}) &\rightarrow \text{SO}(p, q, \mathbb{F}). \end{aligned}$$

Теперь сформулируем и докажем основной результат данного раздела в вещественном случае.

Теорема 26.7.2. Пусть $\mathbb{V}$ – конечномерное векторное пространство над полем вещественных чисел $\mathbb{R}$ с невырожденной квадратичной формой сигнатуры (p, q) . Тогда существуют короткие точные последовательности:

$$1 \rightarrow \mathbb{Z}_2 \rightarrow \text{PIN}(p, q, \mathbb{R}) \rightarrow \mathcal{O}(p, q, \mathbb{R}) \rightarrow 1, \quad (26.116)$$

$$1 \rightarrow \mathbb{Z}_2 \rightarrow \text{SPIN}(p, q, \mathbb{R}) \rightarrow \text{SO}(p, q, \mathbb{R}) \rightarrow 1, \quad (26.117)$$

где $\mathbb{Z}_2 = \{1, -1\}$ – группа по умножению. Если $(p, q) \neq (1, 1)$, то эти двулистные накрытия нетривиальны (т.е. накрывающее пространство связно) для каждой компоненты связности группы вращений $\mathcal{O}(p, q, \mathbb{R})$. В частном случае, накрытие

$$1 \rightarrow \mathbb{Z}_2 \rightarrow \text{SPIN}(n, 0, \mathbb{R}) \rightarrow \text{SO}(n, 0, \mathbb{R}) \rightarrow 1, \quad (26.118)$$

является универсальным для всех $n \geq 3$.

Доказательство. Предположим, что элемент $a = A_1 \cdots A_k \in \text{PIN}(p, q, \mathbb{R})$ лежит в ядре отображения $\widehat{\text{Ad}}$. Тогда согласно предложению 26.7.7 этот элемент является отличным от нуля вещественным числом, $a \in \mathbb{R}^\times$. Поэтому

$$a^2 = N(a) = N(A_1) \cdots N(A_k) = \pm 1,$$

поскольку $\rho_A = \rho_{tA}$ для всех $t \in \mathbb{R}^\times$, и векторы A_i , $i = 1, \dots, k$, могут быть нормированы на единицу, $Q(A_i) = \pm 1$. Сюръективность отображения $\widehat{\text{Ad}}$ следует из теоремы Картана–Дьюдоне.

Для доказательства нетривиальности накрытий достаточно доказать, что элементы e и $-e$ можно соединить кривой, целиком лежащей в $\text{SPIN}(p, q)$. Выберем ортонормированные базисные векторы e^1 и e^2 такие, что выполнены условия $(e^1)^2 = (e^2)^2 = \pm 1$. Это всегда возможно, если $(p, q) \neq (1, 1)$. Тогда кривая

$$a(t) := \mp \cos(2t) + e^1 e^2 \sin(2t) = (e^1 \cos t + e^2 \sin t)(-e^1 \cos t + e^2 \sin t)$$

лежит в $\text{SPIN}(p, q)$ и соединяет точки $a(0) = \mp e$ и $a(\pi/2) = \pm e$. $\square$

В комплексном случае накрытия состоят из четырех листов. При этом ядром гомоморфизма является группа по умножению $\mathbb{Z}_4 := \{\pm 1, \pm i\}$ (см., например, [193], глава I, теорема 2.9).

Из теоремы 26.7.2 следует, что группы $\text{PIN}(p, q) := \text{PIN}(p, q, \mathbb{R})$ и $\text{SPIN}(p, q) := \text{SPIN}(p, q, \mathbb{R})$ два раза накрывают вещественные группы вращений:

$$\frac{\text{PIN}(p, q)}{\mathbb{Z}_2} \simeq \mathbb{O}(p, q), \quad \frac{\text{SPIN}(p, q)}{\mathbb{Z}_2} \simeq \mathbb{SO}(p, q). \quad (26.119)$$

Известно, что вещественные группы вращений $\mathbb{SO}(n) := \mathbb{SO}(n, 0)$, $n \geq 1$, являются связными, а группы $\mathbb{SO}(p, q)$ при $p, q \geq 1$ состоят из двух компонент связности (см., например, [64], глава IX, лемма 4.4). Известны также фундаментальные группы вещественных групп вращений:

$$\begin{aligned} \pi_1(\mathbb{SO}(n)) &\simeq \mathbb{Z}_2, & n \geq 3, \\ \pi_1(\mathbb{SO}_0(p, q)) &\simeq \pi_1(\mathbb{SO}(p)) \times \pi_1(\mathbb{SO}(q)), & p, q \geq 1, \end{aligned}$$

где $\mathbb{SO}_0(p, q)$ обозначает связную компоненту единицы. Следовательно,

$$\begin{aligned} \pi_1(\mathbb{SO}_0(1, q)) &\simeq \pi_1(\mathbb{SO}_0(p, 1)) \simeq \mathbb{Z}_2, \\ \pi_1(\mathbb{SO}_0(p, q)) &\simeq \mathbb{Z}_2 \times \mathbb{Z}_2, & \forall p, q \geq 3. \end{aligned}$$

Поэтому сужение отображения $\widehat{\text{Ad}}$ на связную компоненту единицы $\text{SPIN}_0(1, q)$ группы Лоренца $\text{SPIN}(1, q)$ дает универсальное накрытие

$$1 \rightarrow \mathbb{Z}_2 \rightarrow \text{SPIN}_0(p, 1) \xrightarrow{\widehat{\text{Ad}}} \mathbb{SO}_0(p, 1) \rightarrow 0, \quad p \geq 3. \quad (26.120)$$

Таким образом, вещественные группы $\text{PIN}(p, q)$, $\text{SPIN}(p, q)$ и $\text{SPIN}_0(p, q)$ дважды накрывают, соответственно, группы вращений $\mathbb{O}(p, q)$, $\mathbb{SO}(p, q)$ и $\mathbb{SO}_0(p, q)$. Для вращений евклидова пространства, $q = 0$, группы $\text{SPIN}(n)$ и $\text{SPIN}_0(n)$ совпадают.

В начале раздела мы показали, что группе Ли обратимых элементов $\mathbb{CL}(\mathbb{V}, Q)$ соответствует алгебра Ли, которая как множество совпадает со всей алгеброй Клиффорда $\mathfrak{cl}^\times(\mathbb{V}, Q)$. Построенные выше накрывающие групп вращений являются подгруппами в $\mathbb{CL}(\mathbb{V}, Q)$, и возникает вопрос: какое подмножество элементов в $\mathfrak{cl}^\times(\mathbb{V}, Q)$ образует алгебру Ли группы Липшица $\widehat{\mathbb{P}}(\mathbb{V}, Q)$ и ее подгрупп.

Прежде всего заметим, что группы Ли вещественных вращений $\mathbb{SO}_0(p, q)$, $\mathbb{SO}(p, q)$ и $\mathbb{O}(p, q)$ имеют одинаковую размерность и отличаются только количеством компонент связности. Поэтому они имеют одинаковые локальные группы Ли (окрестности единицы) и, следовательно, одинаковые алгебры Ли (см. раздел 8.1). Это же относится и к соответствующим накрывающим группам $\text{SPIN}_0(p, q)$, $\text{SPIN}(p, q)$ и $\text{PIN}(p, q)$. Кроме того, поскольку накрытия являются конечнолистными, то размерность накрывающих групп и групп вращений совпадают, и совпадают также их алгебры Ли.

Предложение 26.7.10. Алгебра Ли групп Ли $\mathrm{SPIN}_0(p, q)$, $\mathrm{SPIN}(p, q)$ и $\mathrm{PIN}(p, q)$ состоит из всех элементов алгебры Клиффорда второй степени $\mathfrak{cl}_2^\times(p, q)$.

Доказательство. Перестановки образующих показывают справедливость равенства

$$[e^a e^b, e^c e^d] = 2\eta^{ac} e^d e^b - 2\eta^{ad} e^c e^b + 2\eta^{bc} e^a e^d - 2\eta^{bd} e^a e^c. \quad (26.121)$$

Отсюда следует замкнутость элементов алгебры Клиффорда второй степени относительно коммутатора. То есть они образуют подалгебру Ли $\mathfrak{cl}_2^\times \subset \mathfrak{cl}^\times$. Теперь рассмотрим элементы

$$L^{ab} := \frac{1}{4}(e^a e^b - e^b e^a) = \frac{1}{2}\gamma^{ab}. \quad (26.122)$$

Тогда антисимметризация равенства (26.121) по индексам a, b и c, d приводит к коммутатору

$$[L^{ab}, L^{cd}] = -\eta^{ac} L^{bd} + \eta^{ad} L^{bc} + \eta^{bc} L^{ad} - \eta^{bd} L^{ac}. \quad (26.123)$$

Сравнение полученного равенства с коммутатором (1.165) доказывает изоморфизм алгебр Ли: $\mathfrak{cl}_2(p, q) \simeq \mathfrak{o}(p, q)$. $\square$

26.8 Накрытия в низших размерностях

Чтобы продемонстрировать как происходит накрытие групп вращений, описанное в предыдущем разделе, рассмотрим несколько простых примеров.

Пример 26.8.1. Пусть задана двумерная вещественная алгебра Клиффорда $\mathfrak{cl}(1, 0, \mathbb{R})$. В этом случае группа вращений состоит из двух точек $\mathbb{O}(1) = \{\pm 1\}$, а ее собственная подгруппа – из одной $\mathbb{SO}(1) = \{1\}$. Мы их рассматриваем как группы Ли нулевой размерности. Произвольный обратимый элемент алгебры Клиффорда первой степени (вектор) имеет вид $A_k = e^1 A_{1k}$, $A_{1k} \neq 0$. Группа $\mathbb{P}(1, 0, \mathbb{R})$ состоит из элементов вида

$$A_1 \dots A_r = \begin{cases} e \prod_{k=1}^r A_{1k}, & r \text{ четное,} \\ e^1 \prod_{k=1}^r A_{1k}, & r \text{ нечетное.} \end{cases}$$

Следовательно, она представляет собой объединение четных и нечетных обратимых элементов:

$$\mathbb{P}(1, 0, \mathbb{R}) = \mathbb{CL}_0(1, 0, \mathbb{R}) \cup \mathbb{CL}_1(1, 0, \mathbb{R}).$$

Обратимые элементы алгебры Клиффорда образуют группу $\mathbb{CL}(1, 0, \mathbb{R})$ и имеют вид

$$a = e a_0 + e^1 a_1, \quad (26.124)$$

где $a_0, a_1 \in \mathbb{R}$, с одним условием $a_0^2 - a_1^2 \neq 0$. Обратный элемент равен

$$a^{-1} = e \frac{a_0}{a_0^2 - a_1^2} - e^1 \frac{a_1}{a_0^2 - a_1^2}. \quad (26.125)$$

То есть группа Клиффорда $\mathbb{CL}(1, 0, \mathbb{R})$, как многообразие, представляет собой плоскость $a_0, a_1 \in \mathbb{R}^2$ из которой удалены две пересекающиеся прямые $a_1 = \pm a_0$. Она состоит из четырех компонент связности.

Нетрудно вычислить действие оператора $\widehat{\mathrm{Ad}}(a)$ на вектор $A = e^1 A_1$:

$$a A \hat{a}^{-1} = e A_1 \frac{2a_0 a_1}{a_0^2 - a_1^2} + e^1 A_1 \frac{a_0^2 + a_1^2}{a_0^2 - a_1^2}.$$

Для того, чтобы в результате получилось вращение, коэффициент при единице e должен обратиться в нуль. То есть группа Липшица состоит из четных ($a_1 = 0$) и нечетных ($a_0 = 0$) элементов. Следовательно, $\widehat{\mathbb{P}}(1, 0, \mathbb{R}) = \mathbb{P}(1, 0, \mathbb{R})$ в соответствии с предложением 26.7.6. Действие элементов группы Липшица имеет вид

$$\begin{aligned} a = ea_0 : \quad aA\hat{a}^{-1} &= A, \\ a = e^1a_1 : \quad aA\hat{a}^{-1} &= -A. \end{aligned} \quad (26.126)$$

Таким образом, четные элементы группы Липшица сохраняют векторы, а нечетные – меняют направление векторов.

Векторы единичной длины имеют вид $A_k = \pm e^1$. Поэтому группа $\text{PIN}(1, 0, \mathbb{R})$ состоит из четырех элементов:

$$\text{PIN}(1, 0, \mathbb{R}) = \{\pm e, \pm e^1\} \simeq \mathbb{Z}_2 \times \mathbb{Z}_2,$$

т.к. $e^1e^1 = e$. То есть группа $\text{PIN}(1, 0, \mathbb{R})$ представляет собой группу Клейна (см., раздел 1.7). Четные элементы этой группы образуют группу

$$\text{SPIN}(1, 0, \mathbb{R}) = \{\pm e\} \simeq \mathbb{Z}_2.$$

Накрытия

$$\begin{aligned} \frac{\mathbb{P}(1, 0, \mathbb{R})}{\mathbb{R}^\times} &= \mathbb{O}(1, 0, \mathbb{R}), \\ \frac{\text{PIN}(1, 0, \mathbb{R})}{\mathbb{Z}_2} &= \mathbb{O}(1, 0, \mathbb{R}), \\ \frac{\text{SPIN}(1, 0, \mathbb{R})}{\mathbb{Z}_2} &= \mathbb{SO}(1, 0, \mathbb{R}) \end{aligned}$$

задаются следующими отображениями:

$$\begin{aligned} \mathbb{P}(1, 0, \mathbb{R}) \ni \quad \pm eA_0 &\mapsto 1, \quad \pm e^1A_1 \mapsto -1 && \in \mathbb{O}(1, 0, \mathbb{R}), \\ \text{PIN}(1, 0, \mathbb{R}) \ni \quad \pm e &\mapsto 1, \quad \pm e^1 \mapsto -1 && \in \mathbb{O}(1, 0, \mathbb{R}), \\ \text{SPIN}(1, 0, \mathbb{R}) \ni \quad \pm e &\mapsto 1 && \in \mathbb{SO}(1, 0, \mathbb{R}). \end{aligned}$$

где $A_0, A_1 \in \mathbb{R}^\times$.

Теперь рассмотрим комплексную алгебру Клиффорда $\mathfrak{cl}(1, 0, \mathbb{C})$. Ей соответствует комплексная группа вращений $\mathbb{O}(1, 0, \mathbb{C}) = \{1, -1\}$ и собственная группа вращений $\mathbb{SO}(1, 0, \mathbb{C}) = \{1\}$, которые совпадают с группами вращений в вещественном случае. Группа $\mathbb{P}(1, 0, \mathbb{C}) = \mathbb{CL}_0(1, 0, \mathbb{C}) \cup \mathbb{CL}_1(1, 0, \mathbb{C})$ также представляет собой объединение четных и нечетных элементов.

Элементы группы $\mathbb{CL}(1, 0, \mathbb{C})$ и его обратный имеют вид (26.124) и (26.125), только теперь компоненты a_0 и a_1 – комплексные числа. То есть группа $\mathbb{CL}(1, 0, \mathbb{C})$, как многообразие, представляет собой четырехмерное многообразие $\mathbb{C}^2$, из которого удалено множество точек, определяемое уравнением $|a_0| = |a_1|$. $\square$

Пример 26.8.2. Рассмотрим двумерную вещественную алгебру Клиффорда $\mathfrak{cl}(0, 1, \mathbb{R})$. В примере 26.2.3 мы установили, что она изоморфна алгебре комплексных чисел, $\mathfrak{cl}(0, 1, \mathbb{R}) \simeq \mathbb{C}$. Так же как и для алгебры Клиффорда $\mathfrak{cl}(1, 0, \mathbb{R})$, группа $\mathbb{P}(0, 1, \mathbb{R})$ представляет собой объединение четных и нечетных элементов:

$$\mathbb{P}(0, 1, \mathbb{R}) \simeq \mathbb{CL}_0(1, 0, \mathbb{R}) \cup \mathbb{CL}_1(1, 0, \mathbb{R}).$$

Обратимые элементы алгебры Клиффорда $\mathfrak{cl}(0, 1, \mathbb{R})$ имеют вид

$$a = ea_0 + e^1a_1, \quad a_0^2 + a_1^2 \neq 0. \quad (26.127)$$

Обратный элемент равен

$$a^{-1} = e \frac{a_0}{a_0^2 + a_1^2} - e^1 \frac{a_1}{a_0^2 + a_1^2}. \quad (26.128)$$

Таким образом, группа Клиффорда $\mathbb{CL}(0, 1, \mathbb{R})$, как многообразие, представляет собой плоскость $a_0, a_1 \in \mathbb{R}^2$ с выколотым началом координат. То есть связным, но не односвязным многообразием. Здесь возникает отличие от предыдущего случая. Поскольку группа $\mathbb{P}(0, 1, \mathbb{R})$ является объединением четных и нечетных элементов (ось абсцисс и ординат), то она состоит из четырех компонент связности (четыре луча).

Действие оператора $\widehat{\text{Ad}}(a)$ на вектор имеет вид

$$aA\hat{a}^{-1} = -eA_1 \frac{2a_0a_1}{a_0^2 + a_1^2} + e^1A_1 \frac{a_0^2 - a_1^2}{a_0^2 + a_1^2}.$$

Как и в предыдущем примере, группа Лишица состоит либо из четных ($a_1 = 0$), либо из нечетных ($a_0 = 0$) элементов. Поэтому $\widehat{\mathbb{P}}(0, 1, \mathbb{R}) = \mathbb{P}(0, 1, \mathbb{R})$. Действие элементов группы Лишица на векторы имеет прежний вид (26.126). Поэтому четные элементы группы Лишица сохраняют векторы, а нечетные – меняют их направление.

Векторы единичной длины $A_k = \pm e^1$ генерируют группу $\mathbb{PIN}(0, 1, \mathbb{R})$, состоящую из четырех элементов $\{\pm e, \pm e^1\}$. Поскольку теперь $e^1e^1 = -e$, то таблица умножения меняется:

$$\mathbb{PIN}(0, 1, \mathbb{R}) = \{\pm e, \pm e^1\} \simeq \mathbb{Z}_4.$$

То есть группа $\mathbb{PIN}(0, 1, \mathbb{R})$ является циклической группой по умножению, которую можно представить на комплексной плоскости четырьмя точками: $\mathbb{Z}_4 = \{\pm 1, \pm i\} \in \mathbb{C}$. Четные элементы этой группы образуют группу

$$\mathbb{SPIN}(0, 1, \mathbb{R}) = \{\pm e\} \simeq \mathbb{Z}_2.$$

Отметим, что группы $\mathbb{PIN}(1, 0, \mathbb{R})$ и $\mathbb{SPIN}(0, 1, \mathbb{R})$ не изоморфны.

Накрытия

$$\begin{aligned} \frac{\mathbb{P}(0, 1, \mathbb{R})}{\mathbb{R}^\times} &= \mathbb{O}(0, 1, \mathbb{R}), \\ \frac{\mathbb{PIN}(0, 1, \mathbb{R})}{\mathbb{Z}_2} &= \mathbb{O}(0, 1, \mathbb{R}), \\ \frac{\mathbb{SPIN}(0, 1, \mathbb{R})}{\mathbb{Z}_2} &= \mathbb{SO}(0, 1, \mathbb{R}) \end{aligned}$$

задаются следующими отображениями:

$$\begin{aligned} \mathbb{P}(0, 1, \mathbb{R}) \ni \quad & \pm eA_0 \mapsto 1, & \pm e^1A_1 \mapsto -1 & \in \mathbb{O}(1, 0, \mathbb{R}), \\ \mathbb{PIN}(0, 1, \mathbb{R}) \ni \quad & \pm e \mapsto 1, & \pm e^1 \mapsto -1 & \in \mathbb{O}(0, 1, \mathbb{R}), \\ \mathbb{SPIN}(0, 1, \mathbb{R}) \ni \quad & \pm e \mapsto 1 & \in \mathbb{SO}(0, 1, \mathbb{R}). \end{aligned}$$

где $A_0, A_1 \in \mathbb{R}^\times$.

Комплексной алгебре Клиффорда $\mathbf{cl}(0, 1, \mathbb{C})$ соответствует комплексная группа вращений $\mathbb{O}(0, 1, \mathbb{C}) = \{1, -1\}$ и собственная группа вращений $\mathbb{SO}(0, 1, \mathbb{C}) = \{1\}$, которые совпадают с вещественными группами вращений. Группа

$$\mathbb{P}(0, 1, \mathbb{C}) = \mathbb{CL}_0(0, 1, \mathbb{C}) \cup \mathbb{CL}_1(0, 1, \mathbb{C})$$

представляет собой объединение четных и нечетных элементов.

Элементы группы $\mathbb{CL}(0, 1, \mathbb{C})$ имеют вид (26.127) и (26.128) однако теперь компоненты a_0 и a_1 – комплексные числа. Как многообразие группа $\mathbb{CL}(0, 1, \mathbb{C})$ представляет собой двумерную комплексную плоскость $a_0, a_1 \in \mathbb{C}^2$, из которой удалено подмножество точек, определяемое уравнением $|a_0| = |ia_1|$. Поскольку это уравнение эквивалентно уравнению $|a_0| = |a'_1|$, где $a'_1 := ia_1$, то $\mathbb{CL}(0, 1, \mathbb{C}) \simeq \mathbb{CL}(1, 0, \mathbb{C})$, что соответствует общему утверждению об изоморфизме комплексных алгебр Клиффорда одинаковой размерности.

Формулы действия оператора $\widehat{\text{Ad}}(a)$ на вектор A имеют тот же вид, что и в вещественном случае. Поэтому $\mathbb{P}(0, 1, \mathbb{C}) = \widehat{\mathbb{P}}(0, 1, \mathbb{C})$.

Комплексные векторы единичной длины имеют вид $A_k = \{\pm e^1, \pm ie^1\}$. Следовательно, группа $\mathbb{PIN}(0, 1, \mathbb{C})$ состоит из восьми элементов:

$$\mathbb{PIN}(0, 1, \mathbb{C}) = \{\pm e, \pm ie, \pm e^1, \pm ie^1\} = \mathbb{Z}_2 \times \mathbb{Z}_4.$$

Четные элементы образуют группу

$$\mathbb{SPIN}(0, 1, \mathbb{C}) = \{\pm e, \pm ie\} = \mathbb{Z}_2 \times \mathbb{Z}_2.$$

Накрытия

$$\begin{aligned} \frac{\mathbb{P}(0, 1, \mathbb{C})}{\mathbb{C}^\times} &= \mathbb{O}(0, 1, \mathbb{C}), \\ \frac{\mathbb{PIN}(0, 1, \mathbb{C})}{\mathbb{Z}_2 \times \mathbb{Z}_2} &= \mathbb{O}(0, 1, \mathbb{C}), \\ \frac{\mathbb{SPIN}(0, 1, \mathbb{C})}{\mathbb{Z}_2 \times \mathbb{Z}_2} &= \mathbb{SO}(0, 1, \mathbb{C}) \end{aligned}$$

задаются следующими отображениями:

$$\begin{aligned} \mathbb{P}(0, 1, \mathbb{C}) \ni \quad & \pm e A_0 \mapsto 1, & \pm e^1 A_1 \mapsto -1 & \in \mathbb{O}(1, 0, \mathbb{C}), \\ \mathbb{PIN}(0, 1, \mathbb{C}) \ni \quad & \pm e, \pm ie \mapsto 1, & \pm e^1, \pm ie^1 \mapsto -1 & \in \mathbb{O}(0, 1, \mathbb{C}), \\ \mathbb{SPIN}(0, 1, \mathbb{C}) \ni \quad & \pm e \pm ie \mapsto 1 & \in \mathbb{SO}(0, 1, \mathbb{C}). \end{aligned}$$

где $A_0, A_1 \in \mathbb{C}^\times$. □

Пример 26.8.3. Рассмотрим четырехмерную вещественную алгебру Клиффорда $\mathbf{cl}(2, 0, \mathbb{R})$ с двумя образующими e^a , $a = 1, 2$. В примере 26.2.4 доказано, что эта алгебра изоморфна алгебре вещественных 2×2 матриц, $\mathbf{cl}(2, 0, \mathbb{R}) \simeq \mathbf{mat}(2, \mathbb{R})$. Произвольный обратимый вектор имеет вид $A_k = e^a A_{ak}$, при этом для обратимости должно быть выполнено неравенство $Q(A_k) := A_k^2 := A_{1k}^2 + A_{2k}^2 \neq 0$. Группа $\mathbb{P}(2, 0, \mathbb{R})$ состоит из элементов вида

$$A_1 \dots A_r = \begin{cases} e B_r + e^{12} B_{12r}, & r \text{ четное,} \\ e^a C_{ar}, & r \text{ нечетное,} \end{cases}$$

где B_r , B_{12r} и C_{ar} – некоторые функции от A_{ak} . Следовательно, группа $\mathbb{P}(2, 0, \mathbb{R})$ представляет собой объединение четных и нечетных элементов:

$$\mathbb{P}(2, 0, \mathbb{R}) = [\mathbb{CL}_0(2, 0, \mathbb{R}) \oplus \mathbb{CL}_2(2, 0, \mathbb{R})] \cup \mathbb{CL}_1(2, 0, \mathbb{R}).$$

Нетрудно проверить, что обратимые элементы алгебры Клиффорда образуют четырехмерную группу Ли $\mathbb{CL}(2, 0, \mathbb{R})$ с элементами

$$a = ea_0 + e^1 a_1 + e^2 a_2 + e^{12} a_{12},$$

на коэффициенты которых наложено одно условие

$$a_0^2 - a_1^2 - a_2^2 + a_{12}^2 \neq 0. \quad (26.129)$$

То есть групповое многообразие представляет собой четырехмерное евклидово пространство $(a_0, a_1, a_2, a_{12}) \in \mathbb{R}^4$, из которого удалено подмножество точек, определяемых уравнением $a_0^2 - a_1^2 - a_2^2 + a_{12}^2 = 0$. При этом обратный элемент имеет вид

$$a^{-1} = \frac{ea_0 - e^1 a_1 - e^2 a_2 - e^{12} a_{12}}{a_0^2 - a_1^2 - a_2^2 + a_{12}^2}.$$

Несложные вычисления приводят к следующему действию оператора $\widehat{\text{Ad}}(a)$ на вектор $A = e^1 A_1 + e^2 A_2$:

$$\begin{aligned} aA\hat{a}^{-1} = \frac{1}{a_0^2 - a_1^2 - a_2^2 + a_{12}^2} & \left[2e[(a_0 a_1 - a_2 a_{12})A_1 + (a_0 a_2 + a_1 a_{12})A_2] + \right. \\ & + e^1[(a_0^2 + a_1^2 - a_2^2 - a_{12}^2)A_1 + 2(a_1 a_2 + a_0 a_{12})A_2] + \\ & \left. + e^2[2(a_1 a_2 - a_0 a_{12})A_1 + (a_0^2 - a_1^2 + a_2^2 - a_{12}^2)A_2] \right]. \end{aligned}$$

Для того, чтобы это действие было вращением, коэффициент при единице e должен обратиться в нуль для произвольных компонент A_1 и A_2 . Поэтому возможно только два случая:

$$\begin{aligned} 1) \quad a_1 = 0 \quad \text{и} \quad a_2 = 0, \quad \text{тогда} \quad a_+ &:= ea_0 + e^{12} a_{12} \in \mathbb{CL}_0 \oplus \mathbb{CL}_2, \\ 2) \quad a_0 = 0 \quad \text{и} \quad a_{12} = 0, \quad \text{тогда} \quad a_- &:= e^1 a_1 + e^2 a_2 \in \mathbb{CL}_1, \end{aligned} \quad (26.130)$$

т.е. элементы группы Липшица могут быть либо четными, либо нечетными. Отсюда следует изоморфизм группы Липшица $\widehat{\mathbb{P}}(2, 0, \mathbb{R}) = \mathbb{P}(2, 0, \mathbb{R})$, в соответствии с предложением 26.7.6. Действие элементов группы Липшица имеет вид

$$\begin{aligned} a_+ A \hat{a}_+^{-1} &= \frac{1}{a_0^2 + a_{12}^2} \left[e^1 [(a_0^2 - a_{12}^2)A_1 + 2a_0 a_{12} A_2] + e^2 [-2a_0 a_{12} A_1 + (a_0^2 - a_{12}^2)A_2] \right], \\ a_- A \hat{a}_-^{-1} &= -\frac{1}{a_1^2 + a_2^2} \left[e^1 [(a_1^2 - a_2^2)A_1 + 2a_1 a_2 A_2] + e^2 [2a_1 a_2 A_1 - (a_1^2 - a_2^2)A_2] \right]. \end{aligned} \quad (26.131)$$

Нетрудно видеть, что четные элементы сохраняют ориентацию векторов, а нечетные – меняют.

Группа $\mathbb{PIN}(2, 0, \mathbb{R})$ генерируется векторами единичной длины $A_k = e^1 A_{1k} + e^2 A_{2k}$, где $Q(A_k) = A_{1k}^2 + A_{2k}^2 = 1$. Поэтому их можно параметризовать углом:

$$A_k = e^1 \cos \varphi_k + e^2 \sin \varphi_k \in \mathbb{PIN}(2, 0, \mathbb{R}), \quad \varphi_k \in [0, 2\pi].$$

Воспользовавшись элементарными тригонометрическими формулами, легко вычислить два первых произведения:

$$\begin{aligned} A_1 A_2 &= e \cos(\varphi_2 - \varphi_1) + e^{12} \sin(\varphi_2 - \varphi_1), \\ A_1 A_2 A_3 &= e^1 \cos(\varphi_3 - \varphi_2 + \varphi_1) + e^2 \sin(\varphi_3 - \varphi_2 + \varphi_1). \end{aligned}$$

Продолжая эту процедуру, мы видим, что произвольные четные и нечетные элементы группы $\text{PIN}(2, 0, \mathbb{R})$ имеют вид

$$\begin{aligned} a_+ &= e^0 \cos \psi + e^{12} \sin \psi, \\ a_- &= e^1 \cos \varphi + e^2 \sin \varphi, \end{aligned}$$

для некоторых углов ψ и φ . Следовательно, группа $\text{PIN}(2, 0, \mathbb{R})$ состоит из двух компонент связности, каждая из которых представляет собой окружность $\mathbb{S}^1$:

$$\text{PIN}(2, 0, \mathbb{R}) = \{a_+, a_-\}, \quad \psi \in [0, 2\pi], \quad \varphi \in [0, 2\pi].$$

Вычислим действие четного элемента группы Липшица на вектор

$$a_+ A \hat{a}_+^{-1} = e^1 [\cos 2\psi A_1 + \sin 2\psi A_2] + e^2 [-\sin 2\psi A_1 + \cos 2\psi A_2].$$

Отсюда следует, что вектор вращается $A \mapsto SA$, где матрица вращений имеет вид

$$S_+ = \begin{pmatrix} \cos 2\psi & \sin 2\psi \\ -\sin 2\psi & \cos 2\psi \end{pmatrix} \in \text{SO}(2).$$

Это вращение сохраняет ориентацию и происходит на двойной угол 2ψ .

Нечетные элементы группы Липшица действуют по-другому:

$$a_- A \hat{a}_-^{-1} = -e^1 (\cos 2\varphi A_1 + \sin 2\varphi A_2) - e^2 (\sin 2\varphi A_1 - \cos 2\varphi A_2).$$

Соответствующая матрица имеет вид

$$S_- = - \begin{pmatrix} \cos 2\varphi & \sin 2\varphi \\ \sin 2\varphi & -\cos 2\varphi \end{pmatrix} \in \mathbb{O}(2).$$

Это вращение меняет ориентацию и также происходит на двойной угол 2φ .

Таким образом, группа PIN дважды накрывает группу вращений

$$\frac{\text{PIN}(2, 0, \mathbb{R})}{\mathbb{Z}_2} = \mathbb{O}(2).$$

Группа SPIN состоит только из четных элементов и, следовательно, дважды накрывает связную компоненту единицы

$$\frac{\text{SPIN}(2, 0, \mathbb{R})}{\mathbb{Z}_2} = \text{SO}(2).$$

Построенные накрытия нетривиальны, т.к. накрываемые пространства связны, но не являются универсальными, поскольку накрываемые пространства не являются односвязными. $\square$

Пример 26.8.4. Рассмотрим четырехмерную вещественную алгебру Клиффорда $\mathfrak{cl}(1, 1, \mathbb{R})$ с двумя образующими e^a , $a = 1, 2$. В примере 26.2.5 показано, что эта алгебра изоморфна алгебре Клиффорда, рассмотренной в предыдущем примере,

$$\mathfrak{cl}(1, 1, \mathbb{R}) \simeq \mathfrak{cl}(2, 0, \mathbb{R}) \simeq \mathfrak{mat}(2, \mathbb{R}). \quad (26.132)$$

Однако в данном случае группа $\mathbb{PIN}(1, 1, \mathbb{R})$ накрывает не группу вращений $\mathbb{O}(2)$, а группу Лоренца $\mathbb{O}(1, 1)$.

Произвольный обратимый вектор в рассматриваемом случае имеет вид $A_k = e^1 A_{1k} + e^2 A_{2k}$, причем для обратимости должно быть выполнено условие $Q(A_k) := A_{1k}^2 - A_{2k}^2 \neq 0$. Как и в предыдущем случае группа $\mathbb{P}(1, 1, \mathbb{R})$ состоит либо из четных, либо из нечетных элементов:

$$\mathbb{P}(1, 1, \mathbb{R}) = [\mathbb{CL}_0(1, 1, \mathbb{R}) \oplus \mathbb{CL}_2(1, 1, \mathbb{R})] \cup \mathbb{CL}_1(1, 1, \mathbb{R}).$$

Нетрудно проверить, что обратимые элементы алгебры Клиффорда образуют четырехмерную группу Ли $\mathbb{CL}(1, 1, \mathbb{R})$ с элементами

$$a = ea_0 + e^1 a_1 + e^2 a_2 + e^{12} a_{12},$$

на коэффициенты которых наложено одно условие

$$a_0^2 - a_1^2 + a_2^2 - a_{12}^2 \neq 0. \quad (26.133)$$

То есть групповое многообразие представляет собой четырехмерное евклидово пространство $(a_0, a_1, a_2, a_{12}) \in \mathbb{R}^4$, из которого удалено подмножество точек, определяемое уравнением $a_0^2 - a_1^2 + a_2^2 - a_{12}^2 = 0$. При этом обратный элемент группы имеет вид

$$a^{-1} = \frac{e - e^1 a_1 - e^2 a_2 - e^{12} a_{12}}{a_0^2 - a_1^2 + a_2^2 - a_{12}^2}.$$

Поскольку алгебры Клиффорда изоморфны (26.132), и условие обратимости отличается от условия (26.129) только перестановкой координат $a_2 \leftrightarrow a_{12}$, то группы также изоморфны:

$$\mathbb{CL}(1, 1, \mathbb{R}) \simeq \mathbb{CL}(2, 0, \mathbb{R}).$$

Несложные вычисления приводят к следующему действию оператора $\widehat{\text{Ad}}(a)$ на вектор:

$$\begin{aligned} aA\hat{a}^{-1} = \frac{1}{a_0^2 - a_1^2 + a_2^2 - a_{12}^2} & \left[2e[(a_0 a_1 + a_2 a_{12})A_1 - 2(a_0 a_2 + a_1 a_{12})A_2] + \right. \\ & + e^1[(a_0^2 + a_1^2 + a_2^2 + a_{12}^2)A_1 - 2(a_1 a_2 + a_0 a_{12})A_2] + \\ & \left. + e^2[2(a_1 a_2 - a_0 a_{12})A_1 + (a_0^2 - a_1^2 - a_2^2 + a_{12}^2)A_2] \right]. \end{aligned}$$

Для того, чтобы это действие было вращением, коэффициент при единице e должен обратиться в нуль для произвольных компонент A_1 и A_2 . С учетом условия обратимости (26.133) возможно только два случая (26.130) как и в предыдущем примере. Таким образом, элемент группы Лишица может быть только либо четным, либо нечетным. Следовательно, $\widehat{\mathbb{P}}(1, 1, \mathbb{R}) \simeq \mathbb{P}(1, 1, \mathbb{R})$. Действие четных и нечетных элементов группы Лишица имеет вид

$$\begin{aligned} a_+ A \hat{a}_+^{-1} &= \frac{1}{a_0^2 - a_{12}^2} \left[e^1 [(a_0^2 + a_{12}^2)A_1 - 2a_0 a_{12} A_2] + e^2 [-2a_0 a_{12} A_1 + (a_0^2 + a_{12}^2)A_2] \right], \\ a_- A \hat{a}_-^{-1} &= \frac{1}{a_2^2 - a_1^2} \left[e^1 [(a_1^2 + a_2^2)A_1 - 2a_1 a_2 A_2] + e^2 [2a_1 a_2 A_1 - (a_1^2 + a_2^2)A_2] \right]. \end{aligned} \quad (26.134)$$

Нетрудно видеть, что четные элементы сохраняют ориентацию векторов, а нечетные — меняют.

Группа $\mathbb{PIN}(1, 1, \mathbb{R})$ генерируется векторами единичной длины $A_k = e^1 A_{1k} + e^2 A_{2k}$, где $Q(A_k) := A_{1k}^2 - A_{2k}^2 = \pm 1$. Поэтому существует восемь возможностей, в каждой из которых вектор можно параметризовать гиперболическим углом:

$$\begin{aligned} A_k &= \pm(e^1 \operatorname{ch} \varphi_k + e^2 \operatorname{sh} \varphi_k), & \in \mathbb{PIN}(1, 1, \mathbb{R}), & \quad \varphi_k \in (-\infty, \infty). \\ B_k &= \pm(e^1 \operatorname{sh} \phi_k + e^2 \operatorname{ch} \phi_k), & \in \mathbb{PIN}(1, 1, \mathbb{R}), & \quad \phi_k \in (-\infty, \infty). \end{aligned}$$

В рассматриваемом случае лоренцевой метрики появилась дополнительная возможность, т.к. квадрат вектора может быть как положительным, так и отрицательным:

$$Q(A_k) = 1, \quad Q(B_k) = -1.$$

Произведения двух и трех векторов с точностью до общего знака имеют вид

$$\begin{aligned} A_1 A_2 &= e \operatorname{ch}(\varphi_2 - \varphi_1) + e^{12} \operatorname{sh}(\varphi_2 - \varphi_1), \\ B_1 B_2 &= -e \operatorname{ch}(\phi_2 - \phi_1) - e^{12} \operatorname{sh}(\phi_2 - \phi_1), \\ A_1 B_2 &= e \operatorname{sh}(\phi_2 - \varphi_1) + e^{12} \operatorname{ch}(\phi_2 - \varphi_1), \\ B_1 A_2 &= -e \operatorname{sh}(\varphi_2 - \phi_1) - e^{12} \operatorname{ch}(\varphi_2 - \phi_1), \\ A_1 A_2 A_3 &= e^1 \operatorname{ch}(\varphi_3 - \varphi_2 + \varphi_1) + e^2 \operatorname{sh}(\varphi_3 - \varphi_2 + \varphi_1), \\ B_1 B_2 B_3 &= -e^1 \operatorname{sh}(\phi_3 - \phi_2 + \phi_1) - e^2 \operatorname{ch}(\phi_3 - \phi_2 + \phi_1). \end{aligned}$$

Следовательно, произвольный элемент группы $\mathbb{PIN}(1, 1, \mathbb{R})$ имеет один из следующих восьми видов:

$$\begin{aligned} a_+ &= \pm(e^0 \operatorname{ch} \psi + e^{12} \operatorname{sh} \psi), \\ b_+ &= \pm(e^0 \operatorname{sh} \psi + e^{12} \operatorname{ch} \psi), \\ a_- &= \pm(e^1 \operatorname{ch} \varphi + e^2 \operatorname{sh} \varphi), \\ b_- &= \pm(e^1 \operatorname{sh} \varphi + e^2 \operatorname{ch} \varphi). \end{aligned} \tag{26.135}$$

для некоторых углов ψ и φ . Таким образом, группа $\mathbb{PIN}(1, 1, \mathbb{R})$ состоит из восьми компонент связности, каждая из которых диффеоморфна вещественной прямой $\mathbb{R}$.

Для того, чтобы построить накрытия групп вращений, необходимо рассмотреть действие элементов группы (26.135) на векторы. Из формулы (26.134) следует равенство

$$a_+ A \hat{a}_+^{-1} = e^1 (\operatorname{ch} 2\psi A_1 - \operatorname{sh} 2\psi A_2) + e^2 (-\operatorname{sh} 2\psi A_1 + \operatorname{ch} 2\psi A_2).$$

Следовательно, матрица вращений вектора имеет вид

$$S_{++} = \begin{pmatrix} \operatorname{ch} 2\psi & -\operatorname{sh} 2\psi \\ -\operatorname{sh} 2\psi & \operatorname{ch} 2\psi \end{pmatrix} \in \mathbb{SO}_0(1, 1)$$

и принадлежит собственной ортохронной группе Лоренца (связной компоненте единицы). Аналогично вычисляется действие других четных элементов:

$$b_+ A \hat{b}_+^{-1} = -e^1 (\operatorname{ch} 2\psi A_1 - \operatorname{sh} 2\psi A_2) - e^2 (-\operatorname{sh} 2\psi A_1 + \operatorname{ch} 2\psi A_2).$$

Соответствующая матрица вращений отличается общим знаком:

$$S_{+-} = - \begin{pmatrix} \operatorname{ch} 2\psi & -\operatorname{sh} 2\psi \\ -\operatorname{sh} 2\psi & \operatorname{ch} 2\psi \end{pmatrix}$$

Действие нечетных элементов также следует из формул (26.134):

$$\begin{aligned} a_- A \hat{a}_-^{-1} &= -e^1 (\operatorname{ch} 2\varphi A_1 - \operatorname{sh} 2\varphi A_2 - e^2 (\operatorname{sh} 2\varphi A_1 - \operatorname{ch} 2\varphi A_2)), \\ b_- A \hat{b}_-^{-1} &= e^1 (\operatorname{ch} 2\varphi A_1 - \operatorname{sh} 2\varphi A_2 e^2 (\operatorname{sh} 2\varphi A_1 - \operatorname{ch} 2\varphi A_2)). \end{aligned}$$

Следовательно, матрицы вращений имеют вид

$$S_{--} = \begin{pmatrix} -\operatorname{ch} 2\varphi & \operatorname{sh} 2\varphi \\ -\operatorname{sh} 2\varphi & \operatorname{ch} 2\varphi \end{pmatrix}, \quad S_{-+} = \begin{pmatrix} \operatorname{ch} 2\varphi & -\operatorname{sh} 2\varphi \\ \operatorname{sh} 2\varphi & -\operatorname{ch} 2\varphi \end{pmatrix}.$$

Тем самым мы получили в точности все четыре компоненты группы Лоренца $\mathbb{O}(1, 1)$, которые были подробно описаны в разделе 1.7.

Таким образом, мы построили накрытие

$$\frac{\operatorname{PIN}(1, 1, \mathbb{R})}{\mathbb{Z}_2} = \mathbb{O}(1, 1).$$

При этом группе $\mathbb{Z}_2$ соответствуют знаки $\pm$ в выражениях (26.135).

Группа $\operatorname{SPIN}(1, 1, \mathbb{R})$ состоит только из четных элементов, имеет четыре компоненты связности и накрывает собственную группу вращений:

$$\frac{\operatorname{SPIN}(1, 1, \mathbb{R})}{\mathbb{Z}_2} = \operatorname{SO}(1, 1, \mathbb{R}).$$

Связная компонента единицы $\operatorname{SPIN}_0(1, 1, \mathbb{R})$ состоит из двух компонент связности и накрывает собственную ортохронную группу Лоренца:

$$\frac{\operatorname{SPIN}_0(1, 1, \mathbb{R})}{\mathbb{Z}_2} = \operatorname{SO}_0(1, 1).$$

Все построенные в настоящем примере накрытия являются тривиальными, т.к. накрываемые многообразия не являются связными. Следовательно, они не являются универсальными. Так и должно было быть, поскольку собственная ортохронная группа Лоренца $\operatorname{SO}_0(1, 1)$ диффеоморфна прямой $\mathbb{R}$ и является связным и односвязным многообразием.

Отметим, что все накрываемые группы $\operatorname{SPIN}(p, q, \mathbb{R})$ при $p + q \geq 2$ являются связными за исключением группы $\operatorname{SPIN}(1, 1, \mathbb{R})$, которая рассмотрена в настоящем разделе. $\square$

Пример 26.8.5. Рассмотрим четырехмерную вещественную алгебру Клиффорда $\mathfrak{cl}(0, 2, \mathbb{R})$ с образующими e^a , $a = 1, 2$. В примере 26.2.6 было показано, что она изоморфна алгебре кватернионов, $\mathfrak{cl}(0, 2, \mathbb{R}) \simeq \mathbb{H}$. Как и ранее, элементы группы $\mathbb{P}(0, 2, \mathbb{R})$ являются либо четными, либо нечетными:

$$\mathbb{P}(0, 2, \mathbb{R}) = [\operatorname{CL}_0 \oplus \operatorname{CL}(0, 2, \mathbb{R})] \cup \operatorname{CL}_1(0, 2, \mathbb{R}).$$

Нетрудно проверить, что обратимые элементы алгебры Клиффорда образуют четырехмерную группу Ли $\operatorname{CL}(0, 2, \mathbb{R})$ с элементами

$$a = e^0 a_0 + e^1 a_1 + e^2 a_2 + e^{12} a_{12},$$

на коэффициенты которых наложено одно условие

$$a_0^2 + a_1^2 + a_2^2 + a_{12}^2 \neq 0.$$

То есть, групповое многообразие представляет собой четырехмерное евклидово пространство $(a_0, a_1, a_2, a_{12}) \in \mathbb{R}^4$, из которого удалено начало координат. Следовательно, групповое пространство связно и односвязно. Обратный элемент имеет вид

$$a^{-1} = \frac{ea_0 - e^1 a_1 - e^2 a_2 - e^{12} a_{12}}{a_0^2 + a_1^2 + a_2^2 + a_{12}^2}.$$

Несложные вычисления приводят к следующему действию оператора $\widehat{\text{Ad}}(a)$ на вектор $A = e^1 A_1 + e^2 A_2$:

$$\begin{aligned} aA\hat{a}^{-1} = \frac{1}{a_0^2 + a_1^2 + a_2^2 + a_{12}^2} & \left[-2e[(a_0 a_1 + a_2 a_{12})A_1 + (a_0 a_2 - a_1 a_{12})A_2] + \right. \\ & + e^1[(a_0^2 - a_1^2 + a_2^2 - a_{12}^2)A_1 - 2(a_1 a_2 + a_0 a_{12})A_2] + \\ & \left. + e^2[2(a_0 a_{12} - a_1 a_2)A_1 + (a_0^2 + a_1^2 - a_2^2 - a_{12}^2)A_2] \right]. \end{aligned}$$

Для того, чтобы это действие было вращением, коэффициент при единице e должен обратиться в нуль для произвольных компонент A_1 и A_2 . Поэтому возможно только два случая (26.130), как и в предыдущих примерах. Следовательно, $\widehat{\mathbb{P}}(0, 2, \mathbb{R}) \simeq \mathbb{P}(0, 2, \mathbb{R})$. Действие элементов группы Липшица имеет вид

$$\begin{aligned} a_+ A \hat{a}_+^{-1} &= \frac{1}{a_0^2 + a_{12}^2} \left[e^1[(a_0^2 - a_{12}^2)A_1 - 2a_0 a_{12} A_2] + e^2[2a_0 a_{12} A_1 + (a_0^2 - a_{12}^2)A_2] \right], \\ a_- A \hat{a}_-^{-1} &= \frac{1}{a_1^2 + a_2^2} \left[e^1[(-a_1^2 + a_2^2)A_1 - 2a_1 a_2 A_2] + e^2[-2a_1 a_2 A_1 + (a_1^2 - a_2^2)A_2] \right]. \end{aligned}$$

Дальнейший анализ совпадает с анализом, проведенным в примере 26.8.3, и мы получаем двулистные накрытия

$$\frac{\text{PIN}(0, 2, \mathbb{R})}{\mathbb{Z}_2} = \mathbb{O}(2), \quad \frac{\text{SPIN}(0, 2, \mathbb{R})}{\mathbb{Z}_2} = \mathbb{SO}(2).$$

Несмотря на то, что исходные алгебры Клиффорда не изоморфны, имеют место изоморфизмы их подгрупп:

$$\text{PIN}(2, 0, \mathbb{R}) \simeq \text{PIN}(0, 2, \mathbb{R}) \quad \text{и} \quad \text{SPIN}(2, 0, \mathbb{R}) \simeq \text{SPIN}(0, 2, \mathbb{R}). \quad \square$$

26.9 Алгебра Клиффорда $\mathfrak{cl}(p, q)$

В настоящем разделе мы более подробно рассмотрим вещественные алгебры Клиффорда $\mathfrak{cl}(p, q) := \mathfrak{cl}(p, q, \mathbb{R})$ с невырожденной квадратичной формой сигнатуры (p, q) , которые наиболее важны для приложений в математической физике. В этом случае векторное пространство можно отождествить с (псевдо)евклидовым пространством: $\mathbb{V} = \mathbb{R}^{p,q}$.

Для краткости, единицу алгебра Клиффорда e мы часто будем опускать.

Определение. Элементом объема в алгебре Клиффорда $\mathfrak{cl}(p, q)$ называется базисный элемент алгебры максимальной степени $n = p + q$

$$v := e^{1\dots n} := e^1 \cdots e^n = \frac{1}{n!} e^{a_1} \cdots e^{a_n} \varepsilon_{a_1 \dots a_n}, \quad (26.136)$$

где $\varepsilon_{a_1 \dots a_n}$ — полностью антисимметричный тензор n -того ранга, определенный в приложении 37.4. В последнем выражении производится суммирование по всем значениям индексов. $\square$

Нетрудно проверить, что при преобразовании базиса $e^a \mapsto e^b S_b^a$ элемент объема преобразуется как обычный элемент объема в дифференциальной геометрии

$$v \mapsto \det S v.$$

Элемент объема играет важную роль в теории представлений алгебр Клиффорда. В квантовой теории поля он хорошо известен как γ^5 матрица Дирака.

Предложение 26.9.1. *В алгебре Клиффорда $\mathfrak{cl}(p, q)$ элемент объема обладает следующими свойствами:*

$$v^2 = (-1)^{\frac{n(n-1)}{2}+q} e = (-1)^{\frac{n(n+1)}{2}+p} e, \quad (26.137)$$

$$Av = (-1)^{n-1} vA, \quad \forall A \in \mathbb{V}. \quad (26.138)$$

В частности, если размерность векторного пространства n нечетная, то элемент объема лежит в центре алгебры Клиффорда. Если размерность n четная, то

$$Uv = v\hat{U} \quad (26.139)$$

для всех $U \in \mathfrak{cl}(p, q)$.

Инволюции действуют на элемент объема следующим образом:

$$\hat{v} = (-1)^n v, \quad \tilde{v} = (-1)^{\frac{n(n-1)}{2}} v, \quad \bar{v} = (-1)^{\frac{n(n+1)}{2}} v.$$

Доказательство. Прямая проверка. □

Из формулы (26.137) следует, что элемент объема обратим, причем

$$v^{-1} = (-1)^{\frac{n(n+1)}{2}+p} v. \quad (26.140)$$

Умножение элемента алгебры Клиффорда на элемент объема соответствует применению оператора Ходжа. Действие оператора Ходжа на форму степени r имеет вид (3.52). Запишем элемент алгебры Клиффорда степени r в виде (26.27)

$$U_r = \frac{1}{r!} \gamma^{a_1 \dots a_r} u_{a_1 \dots a_r},$$

где элементы $\gamma^{a_1 \dots a_r}$ — это антисимметричная комбинация базисных векторов (26.20) и суммирование проводится по всем возможным значениям индексов. Тогда дуальный элемент алгебры Клиффорда имеет вид

$$*U_r = \frac{1}{(n-r)!} \gamma^{a_{r+1} \dots a_n} \frac{1}{r!} \varepsilon_{a_1 \dots a_r a_{r+1} \dots a_n} u^{a_1 \dots a_r}, \quad (26.141)$$

где $u^{a_1 \dots a_r} := \eta^{a_1 b_1} \dots \eta^{a_r b_r} u_{b_1 \dots b_r}$. Его степень равна $n-r$. С другой стороны, справедливо равенство

$$vU_r = e^{1 \dots n} (e^{1 \dots r} u_{1 \dots r} + \dots) = (-1)^{nr - \frac{r(r+1)}{2}} (e^{(k+1) \dots n} u^{1 \dots k} + \dots) = (-1)^{nr - \frac{r(r+1)}{2}} *U_r.$$

Отсюда следует выражение для дуального по Ходжу элемента:

$$*U_k = (-1)^{nr} v \bar{U}_r. \quad (26.142)$$

Для элемента алгебры Клиффорда $U \in \mathfrak{cl}$ общего вида полученное равенство можно переписать в виде

$$*U = \begin{cases} v\bar{U}, & n - \text{четное}, \\ v\tilde{U}, & n - \text{нечетное}. \end{cases} \quad (26.143)$$

Свойство объема (26.137) можно переписать следующим образом

$$v^2 = \begin{cases} (-1)^p, & n = 3 \text{ или } 4 \pmod{4}, \\ (-1)^{p+1}, & n = 1 \text{ или } 2 \pmod{4}. \end{cases} \quad (26.144)$$

Предложение 26.9.2. *Предположим, что элемент объема v в алгебре Клиффорда $\mathfrak{cl}(p, q)$ удовлетворяет равенству $v^2 = 1$. Тогда элементы*

$$\Pi_+ := \frac{1+v}{2}, \quad \Pi_- := \frac{1-v}{2} \quad (26.145)$$

являются проекционными операторами $\mathfrak{cl}(p, q) \rightarrow \mathfrak{cl}(p, q)$, т.е. удовлетворяют равенствам:

$$\Pi_+^2 = \Pi_+, \quad \Pi_-^2 = \Pi_-, \quad \Pi_+\Pi_- = \Pi_-\Pi_+ = 0, \quad \Pi_+ + \Pi_- = 1. \quad (26.146)$$

Доказательство. Прямая проверка. $\square$

Предложение 26.9.3. *Если размерность $n = p + q$ векторного пространства $\mathbb{V}$ нечетная, то центр алгебры Клиффорда $\mathfrak{cl}(p, q, \mathbb{F})$ состоит из всех элементов, натянутых на базисные векторы $\{e, v\}$. При четных n центр алгебры Клиффорда $\mathfrak{cl}(p, q, \mathbb{F})$ состоит только из элементов поля $\mathbb{F}$.*

Доказательство. Прямая проверка. $\square$

Из определения проекционных операторов следует представление для элемента объема

$$v = \Pi_+ - \Pi_-. \quad (26.147)$$

Понятие элемента объема упрощает доказательство ряда важных результатов.

Предложение 26.9.4. *Предположим, что элемент объема v в алгебре Клиффорда $\mathfrak{cl}(p, q)$ удовлетворяет равенству $v^2 = 1$ и размерность $p+q$ нечетная. Тогда алгебра Клиффорда представляет собой прямую сумму изоморфных подалгебр*

$$\mathfrak{cl}(p, q) = \mathfrak{cl}_+(p, q) \oplus \mathfrak{cl}_-(p, q), \quad (26.148)$$

где

$$\mathfrak{cl}_\pm(p, q) := \Pi_\pm \mathfrak{cl}(p, q) = \mathfrak{cl}(p, q) \Pi_\pm.$$

При этом выполнено равенство

$$\widehat{\mathfrak{cl}_\pm(p, q)} = \mathfrak{cl}_\mp(p, q). \quad (26.149)$$

Доказательство. Поскольку размерность нечетная, то согласно предложению 26.9.1 элемент объема лежит в центре алгебры Клиффорда. Следовательно, проекционные операторы также принадлежат центру. Поэтому из свойств проекционных операторов (26.146) вытекает, что разложение (26.148) является разложением в прямую сумму двух идеалов алгебры. Поскольку $\widehat{\Pi_\pm} = \Pi_\mp$, то выполнено равенство (26.149). Так как инволюция градуировки является автоморфизмом, то оба идеала изоморфны. $\square$

Предложение 26.9.5. *Предположим, что элемент объема v в алгебре Клиффорда $\mathfrak{cl}(p, q)$ удовлетворяет равенству $v^2 = 1$ и размерность $p + q$ четная. Пусть W – произвольный $\mathfrak{cl}(p, q)$ модуль (вещественное векторное пространство с линейным отображением $\mathfrak{cl}(p, q) : W \rightarrow W$). Тогда существует разложение в прямую сумму*

$$W = W_+ \oplus W_-, \quad (26.150)$$

где

$$W_+ := \Pi_+ W, \quad W_- := \Pi_- W.$$

Подпространства $W_\pm$ являются собственными подпространствами элемента объема,

$$vW_\pm = \pm W_\pm,$$

с собственными числами ± 1 . Любой обратимый вектор $V \in \mathbb{V}$ устанавливает изоморфизмы:

$$V : W_+ \rightarrow (VW)_-, \quad W_- \rightarrow (VW)_+. \quad (26.151)$$

Доказательство. Поскольку

$$v\Pi_\pm = \pm\Pi_\pm,$$

то разложение (26.150) следует из свойств проекционных операторов (26.146). Изоморфизм (26.151) является следствием равенств $V\Pi_\pm = \Pi_\mp V$. $\square$

Глава 27

Спиноры

Важнейшую роль в математической физике играют неабелевы группы (псевдо)вращений $\mathbb{O}(p, q)$, $p + q \geq 3$, и их представления. Связные компоненты единицы этих групп Ли не являются односвязными. Согласно общей теореме 11.3.3 для каждого связного многообразия существует универсальное накрытие с накрывающим пространством той же размерности. При этом для накрытий групп можно сохранить их групповую структуру.

Пример 27.0.1. В разделах 1.8, 1.9.4 и 1.9.5 мы доказали следующие изоморфизмы:

$$\mathrm{SO}(3) \simeq \frac{\mathrm{SU}(2)}{\mathbb{Z}_2}, \quad \mathrm{SO}_0(1, 2) \simeq \frac{\mathrm{SL}(2, \mathbb{R})}{\mathbb{Z}_2}, \quad \mathrm{SO}_0(1, 3) \simeq \frac{\mathrm{SL}(2, \mathbb{C})}{\mathbb{Z}_2}.$$

Поскольку группы $\mathrm{SU}(2)$, $\mathrm{SL}(2, \mathbb{R})$ и $\mathrm{SL}(2, \mathbb{C})$ связны и односвязны, то соответствующие накрытия:

$$\mathrm{SU}(2) \rightarrow \mathrm{SO}(3), \quad \mathrm{SL}(2, \mathbb{R}) \rightarrow \mathrm{SO}_0(1, 2), \quad \mathrm{SL}(2, \mathbb{C}) \rightarrow \mathrm{SO}_0(1, 3),$$

являются универсальными и двулиственными. $\square$

В главе 26 было установлено, что вещественные группы $\mathrm{PIN}(p, q)$, $\mathrm{SPIN}(p, q)$ и $\mathrm{SPIN}_0(p, q)$ дважды накрывают, соответственно, группы вещественных (псевдо)вращений $\mathbb{O}(p, q)$, $\mathrm{SO}(p, q)$ и $\mathrm{SO}_0(p, q)$. Группа $\mathrm{SPIN}_0(p, q)$ связна и односвязна. Группы $\mathrm{PIN}(p, q)$ и $\mathrm{SPIN}(p, q)$ при $pq \neq 0$ односвязны, но состоят, соответственно, из четырех и двух компонент связности, как и группы $\mathbb{O}(p, q)$ и $\mathrm{SO}(p, q)$. Поэтому данные двулистные накрытия являются универсальными.

Определение. Представления групп $\mathrm{PIN}(p, q)$, $\mathrm{SPIN}(p, q)$ и $\mathrm{SPIN}_0(p, q)$ называются *спинорными представлениями* групп (псевдо)вращений $\mathbb{O}(p, q)$, $\mathrm{SO}(p, q)$ и $\mathrm{SO}_0(p, q)$. Элемент векторного пространства, в котором действует спинорное представление группы (псевдо)вращений называется *спинором*. Если вектор представления зависит от точки некоторого многообразия, то он называется *спинорным полем*. $\square$

В настоящей главе мы рассмотрим спинорные поля в пространстве Минковского, которые лежат в основе описания частиц с полуцелым спином.

Поскольку накрытие является двулиственным, то каждому элементу группы (псевдо)вращений $\mathrm{SO}(p, q)$ соответствует два элемента универсальной накрывающей группы. Поэтому часто говорят, что спинорное представление является двузначным представлением группы (псевдо)вращений. Согласно общему определению отображения (см. раздел 1.4) это утверждение неверно, т.к. образом элемента при отображении может являться один и только один элемент. Тем не менее мы также будем употреблять широко распространенный термин “спинорное представление”.

27.1 Спинорное представление группы Лоренца $\mathbb{O}(1, 3)$

Спиноры играют исключительно важную роль в квантовой теории поля, т.к. с их помощью описываются частицы с полуцелым спином. Например, электроны, протоны, нейтроны, нейтрино и многие другие элементарные частицы имеют спин $1/2$. В квантовой теории поля частицы описываются полями (после вторичного квантования) $\varphi^A(x)$, $A = 1, 2, \dots$, в пространстве Минковского $x \in \mathbb{R}^{1,3}$. При этом математические модели строятся инвариантными относительно действия группы Пуанкаре в $\mathbb{R}^{1,3}$.

В настоящее время считается, что каждая элементарная частица описывается полем, преобразующимся по некоторому неприводимому представлению группы Пуанкаре. Если частица имеет массу, то неприводимые представления группы Пуанкаре параметризуются массой $m > 0$ и спином s , который может принимать только полуцелые значения $s = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$ (см., например, [5]). В безмассовом случае, $m = 0$, представления описываются спиральностью, которая так же, как и спин, принимает полуцелые значения.

Группа Пуанкаре не является полупростой, т.к. содержит инвариантную абелеву подгруппу сдвигов. Поэтому форма Киллинга–Картана для группы Пуанкаре вырождена, и возникает проблема с построением инвариантов из полей φ^A , если мы хотим работать с неприводимыми представлениями группы Пуанкаре.

Выход из этого положения следующий. Группа Лоренца является простой, ее форма Киллинга–Картана невырождена и никакой проблемы с построением инвариантов нет. Неприводимое представление группы Лоренца в общем случае распадается на несколько неприводимых представлений группы Пуанкаре. Например, массивное векторное поле Прокá (см. раздел 25.4) описывает векторную частицу со спином 1 (три компоненты поперечной части) и одну скалярную частицу со спином 0 (одна компонента продольной части). В квантовой теории поля модели строятся из многокомпонентных полей, образующих, возможно, приводимые представления группы Лоренца. Например, скалярное и векторное поля в каждой точке пространства-времени образуют неприводимые представления группы Лоренца, а симметричное векторное поле второго ранга является приводимым представлением. В зависимости от выбора действия для заданного поля, какие то его компоненты, соответствующие определенному представлению группы Пуанкаре могут быть физическими (распространяющимися, для которых может быть поставлена задача Коши), а какие то – нет.

Ниже будут рассмотрены спинорные представления полной группы Лоренца $\mathbb{O}(1, 3)$ и ее подгрупп: $\mathbb{SO}(1, 3)$ и $\mathbb{SO}_0(1, 3)$. При этом можно рассматривать как вещественные, так и комплексные представления.

27.1.1 γ -матрицы Дирака

Для описания частиц с полуцелым спином, в частности, со спином $1/2$, используются в общем случае комплексные 4×4 γ -матрицы Дирака. Они являются самостоятельным математическим объектом – единственным, с точностью до преобразования подобия, точным (взаимно однозначным) неприводимым представлением образующих алгебры Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$, где $\mathbb{F} = \mathbb{R}, \mathbb{C}$, – и заслуживают отдельного изучения. Хотя определение γ -матриц зависит от лоренцевой метрики, сами по себе они не являются образующими какой-либо алгебры Ли. Как было доказано в предложении

26.7.10, алгебру Ли $\mathfrak{o}(1, 3, \mathbb{F})$ образуют элементы алгебры Клиффорда второй степени $\mathfrak{cl}_2^\times(1, 3, \mathbb{F})$. Используя алгебры Клиффорда, удобно строить спинорные представления полной группы Лоренца $\mathbb{O}(1, 3)$ и ее подгрупп. Этим мы займемся позже, а сейчас изучим алгебру Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$ и ее представление.

В комплексном случае все алгебры Клиффорда с четырьмя образующими изоморфны. В частности, $\mathfrak{cl}(1, 3, \mathbb{C}) \simeq \mathfrak{cl}(4, 0, \mathbb{C})$. Тем не менее мы будем рассматривать сигнатуру $(1, 3)$ для облегчения перехода к вещественному случаю и для описания накрытия вещественной группы Лоренца $\mathbb{O}(1, 3)$, которая не изоморфна группе вращений $\mathbb{O}(4)$.

Рассмотрим алгебру Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$ над четырехмерным векторным пространством $\mathbb{F}^4$ с квадратичной формой $\eta^{ab} := \text{diag}(+---)$, $a, b = 0, 1, 2, 3$ (метрикой Лоренца) (см. раздел 26). Обозначим образующие алгебры Клиффорда через γ^a . По определению, они удовлетворяют следующим антикоммутиационным соотношениям:

$$\{\gamma^a, \gamma^b\} := \gamma^a \gamma^b + \gamma^b \gamma^a = 2\eta^{ab}e, \quad (27.1)$$

где фигурными скобками обозначен антикоммутатор образующих и e – единица алгебры Клиффорда. Это означает, что образующие алгебры Клиффорда с разными индексами антикоммутируют между собой, и квадрат каждой из образующих равен по модулю единице алгебры Клиффорда:

$$(\gamma^0)^2 = e, \quad (\gamma^i)^2 = -e, \quad i = 1, 2, 3.$$

Теперь построим представление алгебры Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$ в виде матриц.

Теорема 27.1.1. *Алгебра Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$, $\mathbb{F} = \mathbb{R}, \mathbb{C}$, обладает единственным (с точностью до преобразования подобия) неприводимым комплексным представлением; его размерность равна четырем.*

Доказательство. См. независимое доказательство, например, в [200], теорема Д.1. Четырехмерное комплексное представление было построено в примере 26.2.11. Его единственность вытекает из простоты матричных алгебр (см., раздел 26.4). $\square$

Замечание. Существование четырехмерного представления алгебры Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$ гарантирует существование четырехмерных представлений группы Ли обратимых элементов $\mathbb{CL}(1, 3, \mathbb{F})$ и ее подгрупп Ли $\text{PIN}(1, 3)$, $\text{SPIN}(1, 3)$ и $\text{SPIN}_0(1, 3)$. В дальнейшем мы увидим, что это представление неприводимо для групп $\text{PIN}(1, 3)$, $\text{SPIN}(1, 3)$ и приводимо для связной компоненты единицы $\text{SPIN}_0(1, 3) \simeq \text{SL}(2, \mathbb{C})$. $\square$

Определение. Четырехмерное точное неприводимое представление образующих γ^a алгебры Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$ называется γ -матрицами Дирака. $\square$

γ -матрицы Дирака определены не однозначно. А именно, справедлива

Лемма 27.1.1 (Паули). *Пусть γ^a и γ'^a – два набора матриц Дирака. Тогда существует единственная (с точностью до умножения на произвольную комплексную константу, отличную от нуля) невырожденная матрица S такая, что матрицы Дирака связаны преобразованием подобия*

$$\gamma'^a = S \gamma^a S^{-1}. \quad (27.2)$$

Доказательство. См. [201]. Мы также приведем доказательство, но несколько позже. $\square$

Лемма Паули – это доказательство единственности точного представления в теореме 27.1.1.

Класс допустимых γ -матриц необходимо сузить, если мы хотим удовлетворить важным физическим условиям при построении моделей математической физики. Поэтому наложим дополнительное условие эрмитового сопряжения для γ -матриц:

$$\gamma^{a\dagger} = \gamma^0 \gamma^a \gamma^0 \quad \Leftrightarrow \quad \gamma^0 \gamma^{a\dagger} \gamma^0 = \gamma^a. \quad (27.3)$$

В разделе 27.1.5 будет показано, что это условие является необходимым и достаточным условием эрмитовости гамильтониана для спинорного поля. Больше никаких дополнительных условий на γ -матрицы накладывать необязательно, и все физические следствия не зависят от выбора матриц Дирака, удовлетворяющих условию (27.3).

Условие эрмитового сопряжения (27.3) означает, что γ матрицы выбраны унитарными. Действительно, легко проверить выполнение равенств

$$\gamma^{a\dagger} \gamma^a = \gamma^a \gamma^{a\dagger} = e,$$

где e – единичная 4×4 матрица, и суммирование по индексу a отсутствует.

Условие эрмитового сопряжения (27.3) можно также записать в следующем виде

$$\gamma^{a\dagger} = \gamma_a := \gamma^b \eta_{ab}.$$

γ -матрицы, удовлетворяющие дополнительному условию (27.3), также определены неоднозначно.

Лемма 27.1.2. Пусть γ^a и γ'^a – два набора матриц Дирака, удовлетворяющих условию эрмитового сопряжения (27.3). Тогда существует единственная (с точностью до умножения на произвольный фазовый множитель) унитарная матрица $U \in \mathbb{U}(4)$ такая, что матрицы Дирака связаны преобразованием подобия

$$\gamma'^a = U \gamma^a U^{-1}.$$

Доказательство. Дано ниже как следствие доказательства леммы Паули [202]. $\square$

Физические следствия моделей математической физики, включающих спинорные поля, не зависят от выбора представления для γ -матриц. Поэтому в литературе используется несколько различных представлений, которые удобны для тех или иных задач. Для наших целей, а нас будут в основном интересовать трансформационные свойства, удобно воспользоваться следующим представлением

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}, \quad (27.4)$$

где $\mathbb{1}$ – единичная 2×2 матрица, $\sigma^i := -\sigma_i$ и σ_i , $i = 1, 2, 3$, – матрицы Паули (37.39). Для этого представления условие эрмитового сопряжения (27.3) выполнено.

Представление γ -матриц (27.4) называется *представлением Вейля* или *киральным представлением*. Оно особенно удобно для изучения трансформационных свойств спиноров.

Теперь построим базис алгебры Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$. Базис состоит из всех возможных произведений образующих γ^a . При этом каждая из образующих не может

входить в произведение дважды в силу определения (27.1). Поэтому алгебра Клиффорда шестнадцатимерна, и ее базис можно выбрать в виде

$$\gamma^I = \{ \mathbb{1}, \quad \gamma^a, \quad \gamma^{[a}\gamma^b], \quad \gamma^{[a}\gamma^b\gamma^c], \quad \gamma^{[a}\gamma^b\gamma^c\gamma^d] \}, \quad I = 1, 2, \dots, 16, \quad (27.5)$$

где $\mathbb{1}$ – единичная 4×4 матрица и квадратные скобки обозначают антисимметризацию всех индексов. Квадрат шести базисных элементов равен единице, а десяти – минус единице. В соответствии с этим свойством разместим базисные элементы в таблице 27.1.

	$(\gamma^I)^2 = \mathbb{1}$	$(\gamma^I)^2 = -\mathbb{1}$
$\mathbb{1}$	$\mathbb{1}$	
γ^a	γ^0	$\gamma^1 \quad \gamma^2 \quad \gamma^3$
$\gamma^{[a}\gamma^b]$	$\gamma^0\gamma^1 \quad \gamma^0\gamma^2 \quad \gamma^0\gamma^3$	$\gamma^1\gamma^2 \quad \gamma^1\gamma^3 \quad \gamma^2\gamma^3$
$\gamma^{[a}\gamma^b\gamma^c]$	$\gamma^1\gamma^2\gamma^3$	$\gamma^0\gamma^1\gamma^2 \quad \gamma^0\gamma^1\gamma^3 \quad \gamma^0\gamma^2\gamma^3$
$\gamma^{[a}\gamma^b\gamma^c\gamma^d]$		$\gamma^0\gamma^1\gamma^2\gamma^3$

Таблица 27.1: Базис алгебры Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$

Из всех базисных элементов γ^I только единичная матрица коммутирует со всеми остальными. Если элемент $\gamma^I \neq \mathbb{1}$, то он коммутирует с 8 из 16 матриц и антикоммутирует с 8 оставшимися. Последнее утверждение проверяется прямой проверкой.

Из определения γ -матриц (27.1) следует, что

$$\gamma_a := \eta_{ab}\gamma^b = (\gamma^a)^{-1}.$$

Отсюда вытекает, что мы получим обратную к матрице γ^I , если в ее выражении через γ^a изменим порядок их следования на противоположный (реверс) и каждую матрицу γ^a заменим на γ_a . Обозначим получившееся выражение для обратной матрицы через γ_I с нижним индексом:

$$\gamma_I \gamma^I = \gamma^I \gamma_I = \mathbb{1} \quad (\text{суммирования по } I \text{ нет}). \quad (27.6)$$

Установим некоторые свойства базиса алгебры Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$.

Предложение 27.1.1.

$$\text{tr } \gamma^I = \begin{cases} 4, & \text{если } \gamma^I = \mathbb{1}, \\ 0, & \text{если } \gamma^I \neq \mathbb{1}. \end{cases} \quad (27.7)$$

Доказательство. Первое равенство очевидно. Предположим теперь, что $\gamma^I \neq \mathbb{1}$ и пусть γ^J – одна из 8 матриц, антикоммутирующих с γ^I :

$$\gamma^I \gamma^J = -\gamma^J \gamma^I \quad \Leftrightarrow \quad \gamma^I = -\gamma^J \gamma^I \gamma_J \quad (\text{суммирования по } J \text{ нет}).$$

Тогда

$$\text{tr } \gamma^I = -\text{tr } (\gamma^J \gamma^I \gamma_J) = -\text{tr } (\gamma^I \gamma_J \gamma^J) = -\text{tr } \gamma^I.$$

Откуда следует искомое равенство $\text{tr } \gamma^I = 0$. □

В частном случае, след всех матриц Дирака в любом представлении равен нулю:

$$\mathrm{tr} \gamma^a = 0.$$

Поскольку квадрат каждой матрицы Дирака с точностью до знака равен единичной матрице, то ее определитель равен по модулю единице. С учетом условия эрмитова сопряжения (27.3) получаем равенство

$$\det \gamma^a = 1.$$

Отсюда сразу следует, что для всех базисных элементов алгебры Клиффорда $\mathbf{cl}(1, 3, \mathbb{F})$ выполнено равенство

$$\det \gamma^I = 1.$$

Следующее важное свойство базиса γ^I сформулируем в виде леммы.

Лемма 27.1.3 (о перестройке). *Если каждую из 16 матриц базиса γ^I умножить справа или слева на одну из них, то с точностью до знака и порядка получим те же 16 матриц, т.е. снова базис алгебры Клиффорда $\mathbf{cl}(1, 3, \mathbb{F})$.*

Доказательство. Прямая проверка. □

Следствие. Справедлива формула

$$\frac{1}{4} \mathrm{tr} (\gamma^I \gamma^J) = \delta_I^J. \quad (27.8)$$

Используя лемму о перестройке и свойства следа (27.7) легко доказать

Предложение 27.1.2. *Матрицы γ^I , $I = 1, \dots, 16$, линейно независимы.*

Доказательство. От противного. Допустим, что выполнено равенство

$$\sum_I u_I \gamma^I = u_0 \mathbb{1} + u_a \gamma^a + u_{ab} \gamma^{[a} \gamma^{b]} + u_{abc} \gamma^{[a} \gamma^b \gamma^{c]} + u_{abcd} \gamma^{[a} \gamma^b \gamma^c \gamma^{d]} = 0,$$

где $\{u_I\} = \{u_0, u_a, u_{ab}, u_{abc}, u_{abcd}\}$ – некоторые числа из поля $\mathbb{F}$. Взятие следа от этого равенства дает $u_0 = 0$. Умножая равенство последовательно на все γ^I и беря следы, получаем, что все коэффициенты разложения равны нулю. Следовательно, матрицы γ^I линейно независимы. □

Вообще говоря, сформулированное выше утверждение можно было бы не доказывать. Действительно, поскольку представление алгебры Клиффорда $\mathbf{cl}(1, 3, \mathbb{F})$ комплексными 4×4 матрицами является точным, то матрицы, соответствующие базису, являются линейно независимыми.

Матрицы γ^I представляют собой базис в алгебре всех комплексных 4×4 матриц $\mathbf{mat}(4, \mathbb{C})$, рассматриваемой как линейное комплексное пространство. Так, произвольную матрицу $A \in \mathbf{mat}(4, \mathbb{C})$ можно взаимно однозначно разложить по базису:

$$A = \sum_I u_I \gamma^I, \quad u_I = \frac{1}{4} \mathrm{tr} (\gamma_I A). \quad (27.9)$$

При этом число комплексных коэффициентов разложения u_I в общем случае совпадает с числом элементов матрицы.

Мы видим, что, как линейные пространства, алгебры $\mathbf{mat}(4, \mathbb{C})$ и $\mathbf{cl}(1, 3, \mathbb{C})$ совпадают. Соответственно, они имеют одинаковые базисы, просто в базисе γ^I , который мы выбрали, клиффордово умножение выглядит проще.

Предложение 27.1.3. Если матрица A коммутирует со всеми матрицами Дирака,

$$[A, \gamma^a] = 0, \quad (27.10)$$

то она кратна единичной матрице, $A \sim \mathbb{1}$.

Доказательство. Из условия (27.10) следует равенство нулю коммутатора $[A, \gamma^i] = 0$. Поскольку матрицы группы $\mathbb{GL}(4, \mathbb{C})$ являются линейными комбинациями матриц γ^i , то, по лемме Шура, $A \sim \mathbb{1}$. $\square$

Используя лемму о перестройке, нетрудно доказать лемму Паули 27.1.1.

Доказательство. Пусть γ^a и γ'^a два набора матриц Дирака, которые необязательно удовлетворяют условию эрмитового сопряжения (27.3). Им соответствуют два базиса алгебры Клиффорда γ^i и γ'^i . Взаимно однозначное соответствие матриц Дирака $\gamma^a \leftrightarrow \gamma'^a$, $a = 0, 1, 2, 3$, влечет за собой естественное взаимно однозначное соответствие всех матриц базиса $\gamma^i \leftrightarrow \gamma'^i$.

Возьмем произвольную квадратную 4×4 матрицу A и составим сумму

$$S := \sum_{i=1}^{16} \gamma'^i A \gamma_i.$$

Выберем одну из матриц γ^j . Ей соответствует обратная матрица γ_j и матрица γ'^j из другого набора. В силу леммы о перестройке, справедлива следующая цепочка равенств

$$\gamma'^j S \gamma_j = \sum_i \gamma'^j \gamma'^i A \gamma_i \gamma_j = \sum_i \gamma'^i A \gamma_i = S.$$

Следовательно,

$$\gamma'^j S = S \gamma^j. \quad (27.11)$$

Теперь для доказательства соотношения (27.2) достаточно доказать существование такой матрицы A , что матрица S невырождена. Построим матрицу

$$T := \sum_{i=1}^{16} \gamma^i B \gamma'_i,$$

где B – произвольная матрица. Рассуждая как и ранее, получаем равенство

$$\gamma^j T = T \gamma'^j.$$

Следовательно,

$$\gamma^j T S = T \gamma'^j S = T S \gamma^j$$

для любой матрицы γ^j . Так как матрица $T S$ коммутирует со всеми γ^j , то она пропорциональна единичной матрице: $T S = c \mathbb{1}$. Постоянная c (возможно нулевая) вычисляется по формуле

$$c = \frac{1}{4} \operatorname{tr} T S = \frac{1}{4} \sum_{ij} \operatorname{tr} (\gamma^i B \gamma'_i \gamma'^j A \gamma_j) = \frac{1}{4} \operatorname{tr} \left(B \sum_{ij} \gamma'_i \gamma'^j A \gamma_j \gamma^i \right) = 4 \operatorname{tr} B S,$$

где мы снова использовали лемму о перестройке (16 раз).

Матрицу A всегда можно выбрать так, чтобы по крайней мере один из матричных элементов S был отличен от нуля. В противном случае легко показать, что матрицы

γ^i не были бы линейно независимы. После этого можно выбрать B так, чтобы $\text{tr } BS = 1/4$. Следовательно, $c = 1$ и $TS = \mathbb{1}$. Таким образом, матрица S невырождена и, умножая равенство (27.11) на S^{-1} справа, получаем соотношение подобия

$$\gamma^i = S\gamma^i S^{-1},$$

которое влечет за собой (27.2).

Если существует другая матрица S' , для которой выполнены те же соотношения, то $S^{-1}S'$ коммутирует со всеми матрицами γ^i и, следовательно, $S^{-1}S' = c\mathbb{1}$. Верно и обратное: если для S выполнены соотношения (27.2), то они выполнены и для любой матрицы, пропорциональной S . Тем самым мы доказали, что несингулярная матрица S существует и определена с точностью до множителя. $\square$

Для матриц Дирака, удовлетворяющих соотношению (27.3), справедливо следующее правило эрмитова сопряжения

$$\gamma^{a\dagger} = \gamma^{-1} = \gamma_a.$$

Следовательно, аналогичные правила справедливы для всех матриц базиса γ^i . Поэтому они эрмитовы или антиэрмитовы в зависимости от того, равен ли квадрат $(\gamma^i)^2$ единичной матрице $\mathbb{1}$ или $-\mathbb{1}$.

Для доказательства леммы 27.1.2 достаточно взять эрмитово сопряжение от равенства (27.2) и воспользоваться (анти)эрмитовостью матриц γ^i .

Из свойства эрмитового сопряжения γ -матриц (27.3) вытекает правило эрмитового сопряжения генераторов лоренцевых вращений $\gamma^{ab} := \gamma^{[a}\gamma^{b]}$:

$$\gamma^{ab\dagger} = -\gamma^0 \gamma^{ab} \gamma^0 \quad \Leftrightarrow \quad \gamma^0 \gamma^{ab\dagger} \gamma^0 = -\gamma^{ab}. \quad (27.12)$$

В заключение раздела приведем еще одно представление матриц Дирака, в котором матрица γ^0 диагональна:

$$\gamma^0_{\text{D}} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^i_{\text{D}} = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}. \quad (27.13)$$

Это представление называется *представлением Дирака* или *стандартным представлением*. Оно связано с киральным представлением преобразованием подобия

$$\gamma^a_{\text{D}} = S\gamma^a S^{-1},$$

где унитарная матрица S имеет вид

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

В представлении Дирака

$$\gamma^5_{\text{D}} = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

То есть матрицы γ^0 и γ^5 меняются местами. Представление Дирака удобно использовать, например, при изучении нерелятивистского предела уравнения Дирака.

27.1.2 Алгебра матриц Дирака

В настоящем разделе мы приведем ряд тождеств для матриц Дирака, которые полезны при проведении вычислений и не зависят от представления.

Исходя из определения матриц Дирака (27.1), нетрудно доказать следующие тождества:

$$\begin{aligned}\gamma^e \gamma^a \gamma_e &= -2\gamma^a, \\ \gamma^e \gamma^a \gamma^b \gamma_e &= 4\eta^{ab}, \\ \gamma^e \gamma^a \gamma^b \gamma^c \gamma_e &= -2\gamma^c \gamma^b \gamma^a, \\ \gamma^e \gamma^a \gamma^b \gamma^c \gamma^d \gamma_e &= 2(\gamma^d \gamma^a \gamma^b \gamma^c + \gamma^c \gamma^b \gamma^a \gamma^d).\end{aligned}\tag{27.14}$$

В приложениях γ -матрицы часто свертываются с лоренцевыми ковекторами $X = e^a X_a \in \mathbb{R}^{1,3}$:

$$(X\gamma) = (\gamma X) := X_a \gamma^a \in \mathbf{mat}(4, \mathbb{C}).\tag{27.15}$$

Тогда из формул (27.14) вытекают следующие равенства:

$$\begin{aligned}\gamma^a (X\gamma) \gamma_a &= -2(X\gamma), \\ \gamma^a (X\gamma) (Y\gamma) \gamma_a &= 4(X, Y), \\ \gamma^a (X\gamma) (Y\gamma) (Z\gamma) \gamma_a &= -2(Z\gamma) (Y\gamma) (X\gamma), \\ \gamma^a (X\gamma) (Y\gamma) (Z\gamma) (W\gamma) \gamma_a &= 2[(W\gamma) (X\gamma) (Y\gamma) (Z\gamma) + (Z\gamma) (Y\gamma) (X\gamma) (W\gamma)],\end{aligned}\tag{27.16}$$

где $(X, Y) := X^a Y^b \eta_{ab}$ – лоренцево скалярное произведение векторов.

Введем матрицу (элемент объема)

$$\gamma^5 := -i\gamma^0 \gamma^1 \gamma^2 \gamma^3 = \frac{i}{24} \varepsilon^{abcd} \gamma_a \gamma_b \gamma_c \gamma_d,\tag{27.17}$$

где ε^{abcd} – полностью антисимметричный тензор четвертого ранга (см. приложение 37.4). Эта матрица в дальнейшем будет играть важную роль. Нетрудно проверить, что матрица γ^5 имеет следующие свойства:

$$(\gamma^5)^2 = 1, \quad \gamma^5 \gamma^a = -\gamma^a \gamma^5, \quad \gamma^{5\dagger} = \gamma^5.\tag{27.18}$$

В киральном представлении (27.4) она является блочно диагональной:

$$\gamma^5 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}.\tag{27.19}$$

Ранее мы построили базис алгебры Клиффорда $\mathfrak{cl}(1, 3, \mathbb{F})$, состоящий из шестнадцати матриц (27.5). Используя γ^5 -матрицу, базис алгебры Клиффорда можно выбрать в виде

$$\{\mathbb{1}, \gamma^a, \gamma^{ab}, \gamma^5 \gamma^a, \gamma^5\}.\tag{27.20}$$

Между матрицами $\gamma^{abc} := \gamma^{[a} \gamma^b \gamma^{c]}$ и $\gamma^5 \gamma^a$ существует взаимно однозначное соответствие:

$$\gamma^{012} = -i\gamma^5 \gamma^3, \quad \gamma^{123} = -i\gamma^5 \gamma^0, \quad \gamma^{230} = -i\gamma^5 \gamma^1, \quad \gamma^{301} = i\gamma^5 \gamma^2,\tag{27.21}$$

где равенства отличаются циклической перестановкой индексов $(0, 1, 2, 3)$ и одним знаком. Нетрудно проверить правило эрмитового сопряжения

$$(\gamma^5 \gamma^a)^\dagger = \gamma^0 \gamma^5 \gamma^a \gamma^0 \Leftrightarrow \gamma^0 (\gamma^5 \gamma^a)^\dagger \gamma^0 = \gamma^5 \gamma^a.$$

Мы уже выписали явный вид матриц γ^a, γ^{ab} и γ^5 . В киральном представлении оставшиеся матрицы имеют вид:

$$\gamma^5 \gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}, \quad \gamma^5 \gamma^i = -\begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}$$

В дальнейшем нам понадобятся свойства следов произведений γ -матриц:

$$\begin{aligned} \text{tr}(\gamma^a) &= 0, & \text{tr}(\gamma^5) &= 0, \\ \text{tr}(\gamma^a \gamma^b) &= 4\eta^{ab}, & \text{tr}(\gamma^5 \gamma^a) &= 0, \\ \text{tr}(\gamma^a \gamma^b \gamma^c \gamma^d) &= 4\eta^{ab}\eta^{cd} - 4\eta^{ac}\eta^{bd} + 4\eta^{ad}\eta^{bc}, \end{aligned} \quad (27.22)$$

которые проверяются прямой проверкой. В общем случае след произведения нечетного числа γ -матриц всегда равен нулю.

Кроме того, справедлива формула

$$\text{tr}(\gamma^5 \gamma^a \gamma^b \gamma^c \gamma^d) = 4i\epsilon^{abcd}. \quad (27.23)$$

Тождества Фирца

В приложениях часто необходимо совершать алгебраические преобразования с выражениями, которые являются полиномами четвертой или более высокой степени от компонент спиноров. В этом случае удобно воспользоваться тождествами для γ -матриц, известных как *тождества Фирца* [203].

Пусть заданы две матрицы γ^K и γ^L из набора (27.5). Рассмотрим их тензорное произведение

$$\gamma_{AB}^K \gamma_{CD}^L.$$

Это – тензор четвертого ранга по спинорным индексам A, B, C и D. Воспользуемся полнотой набора γ -матриц (27.5) и разложим полученное выражение с другим порядком индексов:

$$\gamma_{AB}^K \gamma_{CD}^L = C^{KL}_{IJ} \gamma_{AD}^I \gamma_{CB}^J, \quad (27.24)$$

где C^{KL}_{IJ} – некоторые коэффициенты разложения. Из ортогональности набора матриц (27.8) следует выражение для коэффициентов разложения:

$$C^{KL}_{IJ} = \frac{1}{16} \text{tr}(\gamma_I \gamma^K \gamma_J \gamma^L). \quad (27.25)$$

Отсюда вытекает симметрия полученных коэффициентов относительно перестановки индексов:

$$C^{KL}_{IJ} = C^{LK}_{JI}.$$

Приведем несколько примеров использования тождеств Фирца, опустив промежуточные алгебраические выкладки. Можно доказать следующие тождества:

$$\begin{aligned} \delta_{AB} \delta_{CD} &= \frac{1}{4} \left(\delta_{AD} \delta_{CB} + \gamma_{AD}^a \gamma_{aCB} + \frac{1}{2} \gamma_{AD}^{ab} \gamma_{abCB} - (\gamma^5 \gamma^a)_{AD} (\gamma^5 \gamma_a)_{CB} + \gamma_{AD}^5 \gamma_{CB}^5 \right), \\ \gamma_{AB}^a \gamma_{aCD} &= \delta_{AD} \delta_{CB} - \frac{1}{2} \gamma_{AD}^a \gamma_{aCB} - \frac{1}{2} (\gamma^5 \gamma^a)_{AD} (\gamma^5 \gamma_a)_{CB} - \gamma_{AD}^5 \gamma_{CB}^5, \\ \gamma_{AB}^{ab} \gamma_{abCD} &= 3\delta_{AD} \delta_{CB} - \frac{1}{2} \gamma_{AD}^{ab} \gamma_{abCB} + 3\gamma_{AD}^5 \gamma_{CB}^5, \\ (\gamma^5 \gamma^a)_{AB} (\gamma^5 \gamma_a)_{CD} &= -\delta_{AD} \delta_{CB} - \frac{1}{2} \gamma_{AD}^a \gamma_{aCB} - \frac{1}{2} (\gamma^5 \gamma^a)_{AD} (\gamma^5 \gamma_a)_{CB} + \gamma_{AD}^5 \gamma_{CB}^5, \\ \gamma_{AB}^5 \gamma_{CD}^5 &= \frac{1}{4} \left(\delta_{AD} \delta_{CB} - \gamma_{AD}^a \gamma_{aCB} + \frac{1}{2} \gamma_{AD}^{ab} \gamma_{abCB} + (\gamma^5 \gamma^a)_{AD} (\gamma^5 \gamma_a)_{CB} + \gamma_{AD}^5 \gamma_{CB}^5 \right). \end{aligned} \quad (27.26)$$

Приведенные тождества Фирца используются для перестановки спиноров в произведениях. Например, для четырех спиноров из второго тождества (27.26) вытекает следующее равенство

$$(\bar{\psi}_1 \gamma^a \psi_2)(\bar{\psi}_3 \gamma_a \psi_4) = (\bar{\psi}_1 \psi_4)(\bar{\psi}_3 \psi_2) - (\bar{\psi}_1 \gamma^5 \psi_4)(\bar{\psi}_3 \gamma^5 \psi_2) - \frac{1}{2}(\bar{\psi}_1 \gamma^a \psi_4)(\bar{\psi}_3 \gamma_a \psi_2) - \frac{1}{2}(\bar{\psi}_1 \gamma^5 \gamma^a \psi_4)(\bar{\psi}_3 \gamma^5 \gamma_a \psi_2). \quad (27.27)$$

27.1.3 Трансформационные свойства спиноров

В настоящем разделе мы рассматриваем комплексное точное неприводимое спинорное представление вещественной группы $\mathbb{PIN}(1, 3)$, которая дважды накрывает вещественную группу Лоренца $\mathbb{O}(1, 3)$. Это накрытие является универсальным.

Определение γ -матриц (27.1) можно переписать в виде

$$\gamma^a \gamma^b = \eta^{ab} + \gamma^{ab},$$

где

$$\gamma^{ab} := \frac{1}{2}[\gamma^a, \gamma^b] := \frac{1}{2}(\gamma^a \gamma^b - \gamma^b \gamma^a).$$

Рассмотрим матрицы

$$L_{ab} := \frac{1}{2}\gamma_{ab} := \frac{1}{2}\eta_{ac}\eta_{bd}\gamma^{cd}. \quad (27.28)$$

Пользуясь определением (27.1), нетрудно проверить, что матрицы L_{ab} являются представлением генераторов группы Лоренца $\mathbb{SO}_0(1, 3)$ с коммутационными соотношениями (1.165). Конечно, алгебра матриц L_{ab} не зависит от представления.

Для кирального представления γ -матриц (27.4) представление генераторов группы Лоренца $\mathbb{SO}_0(1, 3)$ имеет блочно диагональный вид:

$$L_{0i} = \frac{1}{2} \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix}, \quad L_{ij} = -\frac{1}{4} \begin{pmatrix} [\sigma_i, \sigma_j] & 0 \\ 0 & [\sigma_i, \sigma_j] \end{pmatrix} = \frac{1}{2} i \varepsilon_{ijk} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}, \quad (27.29)$$

где ε_{ijk} — полностью антисимметричный тензор третьего ранга и $\sigma^i := -\sigma_i$.

Генераторы собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, 3)$ или ее универсальной накрывающей $\mathbb{SL}(2, \mathbb{C})$ (27.29) (алгебры Ли у них одинаковые) представлены в четырехмерном комплексном векторном пространстве $\mathbb{C}^4$. Будем считать, что в этом пространстве действует соответствующая группа Ли $\mathbb{SPIN}_0(1, 3)$, которая изоморфна $\mathbb{SL}(2, \mathbb{C})$. Поскольку представление генераторов (27.29) с помощью γ -матриц является блочно диагональным, то представление группы $\mathbb{SPIN}_0(1, 3) \simeq \mathbb{SL}(2, \mathbb{C})$ является приводимым. Именно блочно диагональный вид генераторов группы Лоренца L_{ab} определяет удобство кирального представления γ -матриц (27.4) при изучении трансформационных свойств.

В киральном представлении матрицы L_{0i} эрмитовы, а L_{ij} — антиэрмитовы:

$$L_{0i}^\dagger = L_{0i}, \quad L_{ij}^\dagger = -L_{ij}.$$

Это значит, что представление пространственных вращений, которые генерируются матрицами L_{ij} , унитарно, а лоренцевых бустов с генераторами L_{0i} неунитарно. Это

свойство находится в согласии с известным утверждением о том, что некомпактная группа Ли не имеет унитарных конечномерных представлений.

Теперь определим трансформационные свойства. Обозначим элемент комплексного векторного пространства $\mathbb{C}^4$ через $\psi \in \mathbb{C}^4$. Он имеет четыре комплексных компоненты ψ_A , $A = 1, 2, 3, 4$, которые мы будем записывать в виде столбца. Будем считать, что в $\mathbb{C}^4$ действует группа $\mathbb{PIN}(1, 3)$:

$$\psi \mapsto \psi' = U\psi, \quad (27.30)$$

где $U \in \mathbb{PIN}(1, 3)$ – 4×4 комплексная матрица. В окрестности единицы матрица U задается элементами алгебры $\omega^{ab}L_{ab} \in \mathfrak{pin}(1, 3) = \mathfrak{o}(1, 3)$, где $\omega^{ab} = -\omega^{ba}$ – “углы вращения”, с помощью экспоненты от матрицы:

$$U_{AB} = \left[\exp \left(\frac{1}{2} \omega^{ab} L_{ab} \right) \right]_{AB}, \quad (27.31)$$

где генераторы группы Лоренца определены соотношениями (27.28), и множитель $1/2$ возник из-за суммирования по всем значениям индексов, а не только по упорядоченным парам $a < b$. Поскольку мы рассматриваем вещественную группу $\mathbb{PIN}(1, 3)$, то параметры ω^{ab} вещественны.

В разделе 26.7 мы установили, что группа Ли $\mathbb{PIN}(1, 3)$ два раза накрывает полную группу Лоренца $\mathbb{O}(1, 3)$. Поэтому каждому лоренцеву вращению соответствует две матрицы из накрывающей группы $\mathbb{PIN}(1, 3)$.

Мы не будем различать верхние и нижние спинорные индексы $A, B, \dots$, потому что это приводит к путанице при определении операции отражения времени и зарядового сопряжения, которые будут рассмотрены в следующем разделе. Для определенности, будем писать их снизу.

Теперь рассмотрим комплексное векторное (по отношению к действию группы $\mathbb{PIN}(1, 3)$) поле $\psi(x) = \{\psi_A(x)\} \in \mathbb{C}^4$ (спинорное поле) в пространстве-времени Минковского с декартовой системой координат $x = \{x^a\} \in \mathbb{R}^{1,3}$, $a = 0, 1, 2, 3$. Мы предполагаем, что при каждом лоренцевом вращении пространства-времени

$$x^a \mapsto x'^a = x^b S_b^a, \quad S \in \mathbb{O}(1, 3) \quad (27.32)$$

компоненты спинорного поля преобразуются по правилу

$$\psi_A(x) \mapsto \psi'_A(x') = U_{AB} \psi_B(x), \quad U \in \mathbb{PIN}(1, 3). \quad (27.33)$$

Поскольку группа $\mathbb{PIN}(1, 3)$ два раза накрывает группу Лоренца $\mathbb{O}(1, 3)$, то нельзя говорить, что каждому лоренцеву вращению пространства Минковского соответствует некоторое фиксированное преобразование поля ψ . Наоборот, каждому преобразованию спинорного поля ψ соответствует определенное лоренцево вращение. При этом двум матрицам $\pm U \in \mathbb{PIN}(1, 3)$, которые отличаются знаком, соответствует один и тот же лоренцев поворот.

В физике говорят, что векторное поле ψ преобразуется по двузначному (спинорному) представлению группы Лоренца.

Напомним, что бесконечно малые преобразования координат и полей мы записываем следующим образом

$$\begin{aligned} x^a &\mapsto x'^a = x^a + \delta x^a, \\ \psi_A(x) &\mapsto \psi'_A(x') = \psi_A(x) + \bar{\delta} \psi_A(x). \end{aligned} \quad (27.34)$$

Тогда преобразования спиноров (27.31) в инфинитезимальной форме имеют вид

$$\begin{aligned}\delta x^a &= -x^b \omega_b^a, & \omega^{ab} &= -\omega^{ba} \in \mathbb{R}, \\ \bar{\delta} \psi_A &= \frac{1}{2} \omega^{ab} L_{abAB} \psi_B.\end{aligned}\quad (27.35)$$

Определение. Комплексное векторное поле $\psi(x) = \{\psi_A(x)\} \in \mathbb{C}^4$, с трансформационными свойствами (27.33) называется *спинором Дирака*. $\square$

Введем проекционные операторы

$$\Pi_{\pm} := \frac{1 \pm \gamma^5}{2}. \quad (27.36)$$

В киральном представлении они имеют вид

$$\Pi_+ = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}, \quad \Pi_- = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1} \end{pmatrix}.$$

То есть

$$\psi_+ := \Pi_+ \psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ 0 \\ 0 \end{pmatrix}, \quad \psi_- := \Pi_- \psi = \begin{pmatrix} 0 \\ 0 \\ \psi_3 \\ \psi_4 \end{pmatrix}.$$

Введем распространенные обозначения для спиноров Дирака в случае кирального представления γ -матриц:

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \xi_a \\ \eta_{\dot{a}} \end{pmatrix}, \quad a = 1, 2 \quad \text{и} \quad \dot{a} = 1, 2.$$

где введены двухкомпонентные спиноры

$$\xi := (\xi_a) = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \eta := (\eta_{\dot{a}}) = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}.$$

Двухкомпонентный спинор η называется *пунктирным*, а ξ — *непунктирным*.

При преобразованиях из собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, 3)$ (связной компоненты единицы) спиноры ψ_+ и ψ_- преобразуются независимо по неприводимым представлениям группы $\mathbb{SL}(2, \mathbb{C}) \simeq \mathbb{SPIN}_0(1, 3)$. Рассмотрим сначала лоренцевы бусты с параметром $v^k := \omega^{0k}$. В этом случае матрица преобразования компонент спинора имеет вид

$$U = \begin{pmatrix} \exp\left(\frac{1}{2}v^k \sigma_k\right) & 0 \\ 0 & \exp\left(-\frac{1}{2}v^k \sigma_k\right) \end{pmatrix} = \begin{pmatrix} \operatorname{ch} \frac{v}{2} + \frac{v^k \sigma_k}{v} \operatorname{sh} \frac{v}{2} & 0 \\ 0 & \operatorname{ch} \frac{v}{2} - \frac{v^k \sigma_k}{v} \operatorname{sh} \frac{v}{2} \end{pmatrix}, \quad (27.37)$$

где $v := \sqrt{(v^1)^2 + (v^2)^2 + (v^3)^2}$. Для получения этой формулы можно использовать формулу для конечных $\mathbb{SU}(2)$ преобразований (1.144). Мы видим, что если двухкомпонентный непунктирный спинор ξ_a преобразуется по правилу

$$\xi_a \mapsto V_{ab} \xi_b,$$

где

$$V = \operatorname{ch} \frac{v}{2} + \frac{v^k \sigma_k}{v} \operatorname{sh} \frac{v}{2}, \quad (27.38)$$

то пунктирный спинор $\eta_{\dot{a}}$ преобразуется по представлению с помощью обратной матрицы:

$$\eta_{\dot{a}} \mapsto V_{\dot{a}\dot{b}}^{-1} \eta_{\dot{b}},$$

которая отличается от матрицы преобразований (27.38) знаком перед вторым слагаемым.

Матрица лоренцевых бустов (27.37), как легко видеть, эрмитова, $U^\dagger = U$.

Матрицу лоренцевых бустов для непунктирных спиноров можно записать в виде

$$V = \cos i \frac{v}{2} - i \frac{v^k \sigma_k}{v} \sin i \frac{v}{2}.$$

Следовательно, обратная матрица равна

$$V^{-1} = \cos i \frac{v}{2} + i \frac{v^k \sigma_k}{v} \sin i \frac{v}{2}.$$

Рассмотрим теперь подгруппу пространственных вращений $\mathbb{SO}(3) \subset \mathbb{SO}_0(1, 3)$, которая параметризуется вектором трехмерных вращений с компонентами $\omega^k := \omega_{ij} \varepsilon^{ijk}$, где ε^{ijk} — полностью антисимметричный тензор третьего ранга. Используя явный вид генераторов вращений (27.29), получаем матрицу преобразования для спиноров

$$U = \begin{pmatrix} \exp\left(\frac{i}{2} \omega^k \sigma_k\right) & 0 \\ 0 & \exp\left(\frac{i}{2} \omega^k \sigma_k\right) \end{pmatrix} = \begin{pmatrix} \cos \frac{\omega}{2} + i \frac{\omega^k \sigma_k}{\omega} \sin \frac{\omega}{2} & 0 \\ 0 & \cos \frac{\omega}{2} + i \frac{\omega^k \sigma_k}{\omega} \sin \frac{\omega}{2} \end{pmatrix}, \quad (27.39)$$

где $\omega := \sqrt{(\omega^1)^2 + (\omega^2)^2 + (\omega^3)^2}$. Мы видим, что и непунктирные и пунктирные спиноры при пространственных вращениях преобразуются одинаково.

Матрица преобразования спиноров, соответствующая пространственным вращением (27.39), является унитарной, $U^\dagger = U^{-1}$.

Лоренцевы бусты и пространственные вращения можно объединить, если ввести комплексный параметр преобразования

$$z^k := \omega^k - i v^k.$$

Тогда матрицу лоренцевых преобразований для непунктирных спиноров можно записать в виде

$$V(z) = \exp\left(\frac{i}{2} z^k \sigma_k\right). \quad (27.40)$$

Эта параметризация комплексной группы $\mathbb{SL}(2, \mathbb{C}) \simeq \mathbb{SPIN}_0(1, 3)$ уже встречалась ранее (1.206). Следовательно, матрица лоренцевых преобразований пунктирного спинора имеет вид $V(z^*)$, где $z^{k*} := \omega^k + i v^k$ обозначает комплексно сопряженный параметр преобразования.

Покажем, что представления матриц преобразования непунктирных и пунктирных спиноров не эквивалентны, т.е. что не существует такой невырожденной матрицы M , для которой выполнено равенство

$$V(z) = M V(z^*) M^{-1}.$$

Действительно, след этого равенства приводит к соотношению

$$\cos \frac{z}{2} = \cos \frac{z^*}{2},$$

которое в общем случае не выполнено.

Таким образом, при преобразованиях спинора Дирака из группы $\text{SPIN}_0(1, 3)$, которые соответствуют преобразованиям из собственной ортохронной группы Лоренца $\mathbb{SO}_0(1, 3)$, спинор Дирака является приводимым. Он разлагается на два неприводимых представления: непунктирные и пунктирные спиноры, которые преобразуются по неэквивалентным представлениям группы $\mathbb{SL}(2, \mathbb{C})$, связанным между собой комплексным сопряжением. Напомним, что группа $\mathbb{SL}(2, \mathbb{C})$ два раза накрывает связную компоненту единицы группы Лоренца $\mathbb{SO}_0(1, 3)$ (см. раздел 1.9.5).

Определение. Спиноры

$$\psi_{\pm} := \frac{1 \pm \gamma^5}{2} \psi \quad (27.41)$$

называются *спинорами Вейля*. $\square$

Определение дано безотносительно к какому бы то ни было представлению γ -матриц. В киральном представлении, как было показано выше, математический смысл спиноров Вейля прозрачен. В других представлениях приводимость представления группы $\text{SPIN}_0(1, 3)$ также имеет место, но выглядит сложнее.

Вейлевские спиноры $\psi_{\pm}$ называют также *левым* и *правым*, т.к. они являются собственными векторами γ^5 -матрицы:

$$\gamma^5 \psi_{\pm} = \pm \psi_{\pm}.$$

Для них часто применяют следующее обозначение

$$\psi_L := \psi_+, \quad \psi_R := \psi_-.$$

Заметим, что пространственное вращение в пространстве Минковского на угол 2π является тождественным преобразованием. При этом из формулы преобразования (27.39) следует, что спинор Дирака меняет знак: $\psi \mapsto -\psi$. В квантовой теории поля предполагается, что сами спиноры ненаблюдаемы, а наблюдаемыми являются только их квадратичные формы. В этом случае изменение знака спинора на противоположный никак не сказывается на физических наблюдаемых.

Проанализируем квадратичные формы спиноров $\psi, \phi \in \mathbb{C}^4$. Поскольку базис алгебры Клиффорда состоит из шестнадцати матриц (27.20), то существует столько же независимых квадратичных форм, которые запишем в виде

$$(\bar{\psi}\phi), \quad (\bar{\psi}\gamma^a\phi), \quad (\bar{\psi}\gamma^{ab}\phi), \quad (\bar{\psi}\gamma^5\gamma^a\phi), \quad (\bar{\psi}\gamma^5\phi), \quad (27.42)$$

где мы ввели *сопряженный спинор*

$$\bar{\psi} := \psi^\dagger \gamma^0. \quad (27.43)$$

Этот спинор записывается в виде строки и называют также *сопряженным по Дираку*. В отличие от эрмитова сопряжения, обозначенного символом $\dagger$, сопряжение по Дираку содержит дополнительное умножение справа на матрицу γ^0 , которое необходимо для правильных трансформационных свойств.

Рассмотрим трансформационные свойства квадратичных форм (27.42). При бесконечно малых преобразованиях компонент спинора (27.35) преобразование сопряженного спинора имеет вид

$$\delta \bar{\psi} = -\frac{1}{2} \omega^{ab} (\bar{\psi} L_{ab}), \quad (27.44)$$

где мы воспользовались правилом эрмитового сопряжения (27.12). Теперь нетрудно проверить, что квадратичные формы (27.42) при лоренцевых вращениях ведут себя следующим образом

$$\bar{\delta}(\bar{\psi}\phi) = 0, \quad (27.45)$$

$$\bar{\delta}(\bar{\psi}\gamma^a\phi) = -(\bar{\psi}\gamma^b\phi)\omega_b^a, \quad (27.46)$$

$$\bar{\delta}(\bar{\psi}\gamma^{ab}\phi) = -(\bar{\psi}\gamma^{cb}\phi)\omega_c^a - (\bar{\psi}\gamma^{ac}\phi)\omega_c^b, \quad (27.47)$$

$$\bar{\delta}(\bar{\psi}\gamma^5\gamma^a\phi) = -(\bar{\psi}\gamma^5\gamma^b\phi)\omega_b^a, \quad (27.48)$$

$$\bar{\delta}(\bar{\psi}\gamma^5\phi) = 0. \quad (27.49)$$

То есть квадратичные формы $(\bar{\psi}\phi)$ и $(\bar{\psi}\gamma^5\phi)$ являются скалярными полями относительно собственных ортохронных преобразований Лоренца, $(\bar{\psi}\gamma^a\phi)$ и $(\bar{\psi}\gamma^5\gamma^a\phi)$ – векторными, а квадратичная форма $(\bar{\psi}\gamma^{ab}\phi)$ – антисимметричным тензорным полем второго ранга.

Квадратичная форма $Q(\psi, \phi) := (\bar{\psi}\phi)$ определяет в комплексном векторном пространстве $\mathbb{C}^4$ эрмитову квадратичную форму, которая линейна по второму аргументу и антилинейна по первому. Эта форма не является положительно определенной и инвариантна относительно действия группы $\text{SPIN}_0(1, 3)$. Таким образом, матрица γ^0 определяет инвариантную эрмитову форму (метрику) в четырехмерном комплексном пространстве $\mathbb{C}^4$. Заметим, что в представлении Дирака (27.13) матрица γ^0 диагональна: $\gamma^0 = \text{diag}(+, +, -, -)$. Это значит, что в комплексном векторном пространстве $\mathbb{C}^4$ задана эрмитова полуторалинейная форма, инвариантная относительно унитарной группы $\mathbb{U}(2, 2)$.

Отметим свойства квадратичных форм при комплексном сопряжении, которые следуют из правила эрмитового сопряжения γ -матриц (27.3):

$$\begin{aligned} (\bar{\psi}\phi)^\dagger &= (\bar{\phi}\psi), \\ (\bar{\psi}\gamma^a\phi)^\dagger &= (\bar{\phi}\gamma^a\psi), \\ (\bar{\psi}\gamma^{ab}\phi)^\dagger &= -(\bar{\phi}\gamma^{ab}\psi), \\ (\bar{\psi}\gamma^5\gamma^a\phi)^\dagger &= (\bar{\phi}\gamma^5\gamma^a\psi), \\ (\bar{\psi}\gamma^5\phi)^\dagger &= -(\bar{\phi}\gamma^5\psi). \end{aligned} \quad (27.50)$$

Следовательно, компоненты антисимметричного тензора $(\bar{\psi}\gamma^{ab}\psi)$ и псевдоскаляра $(\bar{\psi}\gamma^5\psi)$ являются чисто мнимыми.

В заключение раздела выпишем формулы для конечных преобразований. Спинор ψ преобразуется по правилу (27.30):

$$\psi \mapsto \psi' = U\psi, \quad U = \exp\left(\frac{1}{2}\omega^{ab}L_{ab}\right) \in \text{SPIN}_0(1, 3). \quad (27.51)$$

В интегральной форме преобразование сопряженного спинора (27.44) имеет вид

$$\bar{\psi} \mapsto \bar{\psi}' = \bar{\psi}U^{-1}, \quad (27.52)$$

где обратное преобразование

$$U^{-1} = \exp\left(-\frac{1}{2}\omega^{ab}L_{ab}\right)$$

отличается от преобразования (27.51) только знаком вещественных параметров ω^{ab} .

Инфинитезимальное преобразование (27.46) в интегральной форме принимает вид

$$(\bar{\psi}U^{-1}\gamma^a U\phi) = (\bar{\psi}\gamma^b\phi)S_b^{-1a}, \quad S \in \mathbb{SO}_0(1, 3). \quad (27.53)$$

Поскольку эта формула верна для всех спиноров ψ и ϕ , то γ -матрицы удовлетворяют равенству

$$U^{-1}\gamma^a U = \gamma^b S_b^{-1a}. \quad (27.54)$$

Отсюда следует, что γ -матрицы ковариантны относительно действия группы $\text{SPIN}_0(1, 3)$ в спинорном пространстве. Кроме того, из последнего равенства следует, что двум матрицам $\pm U$ соответствует одно лоренцево вращение.

27.1.4 Дискретные преобразования P , C и T

В квантовой теории поля важную роль играют дискретные симметрии: пространственная инверсия P (преобразование четности), обращение времени T и зарядовое сопряжение C , которые мы рассмотрим в настоящем разделе.

Введем в спинорном пространстве $\mathbb{C}^4$ оператор *пространственной инверсии (четности)*

$$P : \quad \psi \mapsto \psi_p = P\psi, \quad (27.55)$$

где P – некоторая комплексная 4×4 матрица. Отсюда следует, что сопряженный спинор преобразуется по правилу

$$P : \quad \bar{\psi} \mapsto \bar{\psi}_p = \bar{\psi}\gamma^0 P^\dagger \gamma^0.$$

Поскольку квадратичная форма $(\bar{\psi}\gamma^a\phi)$ – вектор, то для пространственных отражений необходимо потребовать выполнения следующих равенств:

$$\begin{aligned} (\bar{\psi}\gamma^0\phi) &\mapsto (\bar{\psi}_p\gamma^0\phi_p) = (\bar{\psi}\gamma^0 P^\dagger P\phi) = (\bar{\psi}\gamma^0\phi), \\ (\bar{\psi}\gamma^i\phi) &\mapsto (\bar{\psi}_p\gamma^i\phi_p) = (\bar{\psi}\gamma^0 P^\dagger \gamma^0 \gamma^i P\phi) = -(\bar{\psi}\gamma^i\phi), \end{aligned} \quad (27.56)$$

где ψ и ϕ – произвольные спиноры. Из первого равенства следует условие унитарности

$$P^\dagger P = 1.$$

Второе равенство имеет решение

$$P = \epsilon_p \gamma^0, \quad (27.57)$$

где $\epsilon_p \in \mathbb{C}$ – произвольный фазовый множитель, $\epsilon_p^\dagger \epsilon_p = 1$. В пространстве Минковского пространственная инверсия, примененная дважды, дает тождественное преобразование. В спинорном пространстве спинор при двойной инверсии преобразуется по правилу

$$\psi_{pp} = \epsilon_p^2 \psi.$$

При этом он может изменить знак. Поэтому для фазового множителя возможны четыре значения: $\epsilon_p = \pm 1$ и $\epsilon_p = \pm i$.

Таким образом, преобразование четности действует на спиноры следующим образом

$$P : \quad \psi(t, \mathbf{x}) \mapsto \epsilon_p \gamma^0 \psi(t, -\mathbf{x}), \quad \bar{\psi}(t, \mathbf{x}) \mapsto \epsilon_p^* \bar{\psi}(t, -\mathbf{x}) \gamma^0,$$

где $x := (t, \mathbf{x})$ и мы учли также, что при преобразовании четности меняется направление пространственных осей, $\mathbf{x} \mapsto -\mathbf{x}$. В этом случае квадратичные формы преобразуются по следующим правилам:

$$\begin{aligned} (\bar{\psi}\phi) &\mapsto (\bar{\psi}\phi), & (\bar{\psi}\gamma^{0i}\phi) &\mapsto -(\bar{\psi}\gamma^{0i}\phi), \\ (\bar{\psi}\gamma^5\phi) &\mapsto -(\bar{\psi}\gamma^5\phi), & (\bar{\psi}\gamma^{ij}\phi) &\mapsto (\bar{\psi}\gamma^{ij}\phi), \\ (\bar{\psi}\gamma^0\phi) &\mapsto (\bar{\psi}\gamma^0\phi), & (\bar{\psi}\gamma^5\gamma^0\phi) &\mapsto -(\bar{\psi}\gamma^5\gamma^0\phi), \\ (\bar{\psi}\gamma^i\phi) &\mapsto -(\bar{\psi}\gamma^i\phi), & (\bar{\psi}\gamma^5\gamma^i\phi) &\mapsto (\bar{\psi}\gamma^5\gamma^i\phi), \end{aligned} \quad (27.58)$$

которые не зависят от выбора фазового множителя. Поэтому мы говорим, что $(\bar{\psi}\gamma^5\gamma^a\phi)$ – это псевдовектор (аксиальный вектор), а $(\bar{\psi}\gamma^5\phi)$ – псевдоскаляр.

В киральном представлении преобразование четности имеет вид

$$P \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \epsilon_p \begin{pmatrix} \eta \\ \xi \end{pmatrix},$$

т.е. меняет непунктирный и пунктирный спиноры местами. Отсюда следует, что спинор Дирака преобразуется по неприводимому представлению группы $\mathbb{PIN}_0(1, 3)$, которая включает преобразование четности P и два раза накрывает группу Лоренца $\mathbb{O}_0(1, 3)$.

Теперь рассмотрим другие дискретные преобразования. Согласно общим принципам квантовой теории поля операции зарядового сопряжения и обращения времени должны быть антиунитарны ([200]). Поэтому они включают в себя комплексное сопряжение спинорных компонент.

Рассмотрим операцию *зарядового сопряжения*:

$$C : \quad \psi \mapsto \psi_c := C\psi^*,$$

где C – некоторая матрица и звездочка $*$ обозначает комплексное сопряжение компонент спинора без транспонирования. Тогда сопряженный спинор преобразуется по правилу

$$C : \quad \bar{\psi} \mapsto \bar{\psi}_c = \psi^T C^\dagger \gamma^0.$$

Зарядовое сопряжение не действует на пространственные координаты $x \in \mathbb{R}^{1,3}$. Поэтому для его определения будем использовать вид действия (27.90) для спиноров, взаимодействующих с электромагнитным полем. Потребуем, чтобы квадратичные формы преобразовывались следующим образом

$$\begin{aligned} (\bar{\psi}\psi) &= (\bar{\psi}_c\psi_c), \\ (\bar{\psi}\gamma^a\psi) &= -(\bar{\psi}_c\gamma^a\psi_c). \end{aligned} \quad (27.59)$$

Второе условие означает изменение знака электромагнитного тока, что соответствует изменению знака заряда, как будет показано в следующем разделе.

Будем искать матрицу C для кирального представления γ -матриц, в котором $\gamma^{0T} = \gamma^0$. Тогда из первого условия (27.59) следует равенство

$$(\psi^T C^\dagger \gamma^0 C \psi^*) = (\psi^\dagger C^T \gamma^0 C^* \psi) = (\psi^\dagger \gamma^0 \psi),$$

где C^* обозначает комплексно сопряженную матрицу. Отсюда вытекает соотношение

$$C^T \gamma^0 C^* = \gamma^0,$$

которое после транспонирования дает равенство

$$C^\dagger = \gamma^0 C^{-1} \gamma^0. \quad (27.60)$$

После этого второе равенство (27.59) приводит к следующему уравнению

$$M^{-1} \gamma^a M = \gamma^{a\tau}, \quad (27.61)$$

где $M := C \gamma^0$. В киральном представлении (27.4) γ -матрицы удовлетворяют следующим правилам транспонирования:

$$\gamma^{0\tau} = \gamma^0, \quad \gamma^{1\tau} = -\gamma^1, \quad \gamma^{2\tau} = \gamma^2, \quad \gamma^{3\tau} = -\gamma^3. \quad (27.62)$$

Уравнения (27.60) и (27.61) определяют матрицу зарядового сопряжения с точностью до произвольного фазового множителя ϵ_c . Нетрудно проверить, что матрица

$$M = i\epsilon_c \gamma^2 \gamma^0 \Leftrightarrow M^{-1} = i\epsilon_c^* \gamma^0 \gamma^2.$$

удовлетворяет уравнению (27.61). Тогда матрица зарядового сопряжения примет вид

$$C = i\epsilon_c \gamma^2. \quad (27.63)$$

Легко проверить, что эта матрица унитарна

$$C^\dagger = C^{-1}.$$

При зарядовом сопряжении декартовы координаты пространства Минковского не меняются. Поэтому зарядовое сопряжение спинора принимает вид

$$C : \quad \psi(x) \mapsto \psi_c(x) = i\epsilon_c \gamma^2 \psi^*(x). \quad (27.64)$$

Если зарядовое сопряжение применить дважды, то спинор не изменится:

$$\psi_{cc} = \epsilon_c \epsilon_c^* \gamma^2 \gamma^{2*} \psi = \psi.$$

То есть на фазовый множитель ϵ_c при этом никаких ограничений не возникает.

В киральном представлении (27.4) операция зарядового сопряжения принимает вид

$$C \begin{pmatrix} \xi \\ \eta \end{pmatrix} = i\epsilon_c \begin{pmatrix} -\sigma^2 \eta^* \\ \sigma^2 \xi^* \end{pmatrix}, \quad (27.65)$$

и перемешивает пунктирные и непунктирные спиноры.

В других представлениях γ -матриц матрица зарядового сопряжения будет связана с матрицей (27.63) преобразованием подобия. Конечно, свойство унитарности матрицы зарядового сопряжения при этом сохранится.

Аналогично вводится операция *обращения времени*, которая также включает комплексное сопряжение,

$$T : \quad \psi \mapsto \psi_t := T \psi^*, \quad (27.66)$$

где T – некоторая матрица. Отсюда вытекает правило преобразования сопряженного спинора

$$T : \quad \bar{\psi} \mapsto \bar{\psi}_t = \psi^\tau T^\dagger \gamma^0.$$

Мы требуем, чтобы при обращении времени были выполнены следующие условия:

$$\begin{aligned}(\bar{\psi}\psi) &= (\bar{\psi}_t\psi_t), \\ (\bar{\psi}\gamma^0\psi) &= (\bar{\psi}_t\gamma^0\psi_t), \\ (\bar{\psi}\gamma^i\psi) &= -(\bar{\psi}_t\gamma^i\psi_t).\end{aligned}$$

Второе и третье равенство означают изменение спинорного тока, что и должно происходить при обращении времени. В отличие от преобразования четности обращение времени включает в себя комплексное сопряжение.

Как и в случае зарядового сопряжения из первого условия вытекает равенство

$$T^\dagger = \gamma^0 T^{-1} \gamma^0 \quad (27.67)$$

для кирального представления γ -матриц. После этого второе и третье условия приводят к равенствам

$$M^{-1}\gamma^0 M = \gamma^{0\tau}, \quad M^{-1}\gamma^i M = -\gamma^{i\tau}, \quad (27.68)$$

где $M := T\gamma^0$. Нетрудно проверить, что уравнения (27.67) и (27.68) имеют решение

$$T = i\epsilon_t \gamma^1 \gamma^3, \quad (27.69)$$

которое определено с точностью до произвольного фазового множителя ϵ_t . С учетом преобразования координат операция обращения времени принимает вид

$$T : \quad \psi(t, \mathbf{x}) \mapsto \psi_t(t, \mathbf{x}) = i\epsilon_t \gamma^1 \gamma^3 \psi^*(-t, \mathbf{x}). \quad (27.70)$$

При двойном применении операции обращения времени спинор меняет свой знак:

$$\psi_{tt} = \epsilon_t \epsilon_t^* \gamma^1 \gamma^3 \gamma^{1*} \gamma^{3*} \psi = -\psi.$$

Как и в случае зарядового сопряжения никаких ограничений на фазовый множитель при этом не возникает.

В киральном представлении операция обращения времени записывается в виде

$$T \begin{pmatrix} \xi \\ \eta \end{pmatrix} = -\epsilon_t \begin{pmatrix} \sigma^2 \xi^* \\ \sigma^2 \eta^* \end{pmatrix},$$

и не перемешивает пунктирные и непунктирные спиноры.

Если применить все три рассмотренные дискретные симметрии вместе, то получим следующее преобразование спинора

$$PCT : \quad \psi(t, \mathbf{x}) \mapsto -i\epsilon_P \epsilon_C \epsilon_t^* \gamma^5 \psi(-t, -\mathbf{x}). \quad (27.71)$$

По построению, действие (27.78) для свободного поля Дирака инвариантно относительно каждого из преобразований P , C и T по-отдельности.

В модели Глэшоу–Вайнберга–Салама электрослабых взаимодействий, которая является общепринятой в настоящее время и находится в хорошем согласии с экспериментальными данными, инвариантность относительно пространственной инверсии P и зарядового сопряжения C явно нарушена. Инвариантность относительно комбинированных PC -преобразований нарушается только в редких экспериментах. Нарушение PCT -преобразований на сегодняшний день не обнаружено. Поэтому полная PCT -симметрия считается точной.

В аксиоматической квантовой теории поля доказана *PCT-теорема* [204]. Эта теорема утверждает, что модели квантовой теории поля при достаточно общих условиях, включающих локальность и лоренц инвариантность, инвариантны относительно совместных PCT -преобразований. Для формулировки этой теоремы требуется введение многих новых понятий, что выходит за рамки настоящей монографии.

27.1.5 Уравнение Дирака

В настоящем разделе мы введем уравнение Дирака и изучим его простейшие свойства безотносительно к трансформационным свойствам спиноров, которые были рассмотрены в предыдущих разделах. Это сделано для того, чтобы подчеркнуть те математические и физические аспекты теории, которые часто тонут в обсуждении относительно сложных трансформационных свойств.

Рассмотрим пространство-время Минковского $\mathbb{R}^{1,3}$ с декартовыми координатами $\{x^a\} = \{x^0 := t, \mathbf{x}\}$ и метрикой $\eta_{ab} = \text{diag}(+---)$. Пусть задан дираковский спинор $\psi(x)$:

$$\mathbb{R}^{1,3} \ni x \mapsto \psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} \in \mathbb{C}^4$$

с трансформационными свойствами, определенными в разделе 27.1.3. На данном этапе – это просто столбец из четырех комплекснозначных функций. Уравнением Дирака называется система линейных дифференциальных уравнений первого порядка для компонент спинора Дирака

$$i\gamma^a \partial_a \psi - m\psi = 0, \quad (27.72)$$

где γ^a – матрицы Дирака и $m \in \mathbb{R}$ – масса спинора. Будем считать, что масса неотрицательна $m \geq 0$.

Замечание. Уравнение Дирака играет чрезвычайно важную роль в квантовой теории поля. После вторичного квантования оно описывает две частицы: частицу со спином $1/2$ и ее античастицу, имеющую все те же характеристики, но противоположный электрический заряд. Первоначально Дирак предложил это уравнение для описания электрона [205, 206]. Позже выяснилось, что уравнение (27.72) описывает не только электрон, но и его античастицу с противоположным зарядом, которая в то время не была известна. Вскоре эту частицу обнаружили экспериментально и назвали позитроном. Это открытие стало триумфом для уравнения Дирака. В настоящее время оно используется для описания всех частиц со спином $1/2$: электронов, протонов, нейтронов, нейтрино, кварков и т.д. $\square$

Опишем некоторые свойства уравнения Дирака.

Предложение 27.1.4. Уравнение Дирака ковариантно относительно собственных ортохронных преобразований Лоренца $\mathbb{SO}_0(1, 3)$.

Доказательство. Координаты и компоненты спинора преобразуются по правилам (27.32) и (27.33):

$$x'^a = x^b S_b^{-1a}, \quad \psi'(x') = U\psi(x).$$

Учитывая свойство γ -матриц (27.54), мы видим, что после преобразования уравнение Дирака примет вид

$$i\gamma^a \partial'_a \psi' - m\psi' = U(i\gamma^a \partial_a \psi - m\psi) = 0,$$

т.е. умножится на некоторую невырожденную матрицу $U \in \mathbb{SPIN}_0(1, 3)$. Это и означает ковариантность. $\square$

Предложение 27.1.5. Пусть спинор ψ удовлетворяет уравнению Дирака (27.72). Тогда каждая его компонента удовлетворяет уравнению Клейна–Гордона–Фока (25.34).

Доказательство. Умножим уравнение Дирака (27.72) слева на дифференциальный оператор $i\gamma^a\partial_a + m$ и воспользуемся определением γ -матриц (27.1)

$$(i\gamma^a\partial_a + m)(i\gamma^b\partial_b - m)\psi = -(\square + m^2)\psi = 0,$$

где $\square := \eta^{ab}\partial_a\partial_b$ – оператор Даламбера, и мы воспользовались перестановочностью частных производных, $\partial_a\partial_b = \partial_b\partial_a$. $\square$

Обратное утверждение, вообще говоря, не имеет места. Не каждый набор четырех комплексных функций, удовлетворяющий уравнению Клейна–Гордона–Фока, будет удовлетворять также системе уравнений (27.72). Трансформационные свойства при этом не играют никакой роли.

Нетрудно проверить, что дифференциальные операторы $i\gamma^a\partial_a + m$ и $i\gamma^b\partial_b - m$ перестановочны между собой. Поэтому знак массы m в уравнении Дирака не является существенным.

Умножим уравнение Дирака слева на γ^0 и перепишем его в виде уравнения Шредингера

$$i\partial_0\psi = H\psi,$$

где гамильтониан имеет вид

$$H := -i\gamma^0\gamma^k\partial_k + m\gamma^0, \quad k = 1, 2, 3. \quad (27.73)$$

Мы предполагаем, что спинор $\psi(t, \mathbf{x})$ в каждый момент времени является элементом гильбертова пространства $\mathbb{H} = \mathcal{L}_2(\mathbb{R}^3)$ с положительно определенным скалярным произведением

$$(\varphi, \psi) = \int d\mathbf{x}(\varphi_1^\dagger\psi_1 + \varphi_2^\dagger\psi_2 + \varphi_3^\dagger\psi_3 + \varphi_4^\dagger\psi_4), \quad \varphi, \psi \in \mathbb{H}. \quad (27.74)$$

Тогда, для сохранения нормы $\|\psi\| := \sqrt{(\psi, \psi)}$ во времени гамильтониан должен быть эрмитов $H^\dagger = H$. Из этого условия следуют равенства:

$$\gamma^{0\dagger} = \gamma^0, \quad \gamma^{k\dagger}\gamma^{0\dagger} = \gamma^0\gamma^k.$$

Напомним, что при эрмитовом сопряжении производная меняет знак, $\partial_i \mapsto -\partial_i$, т.к. необходимо интегрирование по частям. Эти четыре условия можно переписать в виде равенства (27.3). Таким образом, условие эрмитового сопряжения γ -матриц является необходимым и достаточным условием эрмитовости гамильтониана (27.73) для уравнения Дирака.

Возьмем эрмитово сопряжение от уравнения Дирака

$$-i\partial_a\psi^\dagger\gamma^{a\dagger} - m\psi^\dagger = 0,$$

где $\psi^\dagger := (\psi_1^\dagger, \psi_2^\dagger, \psi_3^\dagger, \psi_4^\dagger)$ – матрица-строка с комплексно сопряженными компонентами. Умножим его справа на γ^0 и запишем в виде

$$-i\partial_a\bar{\psi}\gamma^a - m\bar{\psi} = 0, \quad (27.75)$$

где мы воспользовались правилом эрмитового сопряжения γ -матриц (27.3) и определением сопряженного спинора (27.43). Теперь введем *спинорный ток*

$$j^a := \bar{\psi}\gamma^a\psi. \quad (27.76)$$

Каждая из четырех компонент тока является вещественным числом в силу условия эрмитового сопряжения γ -матриц. Нетрудно проверить, что если выполнено уравнение Дирака (27.72) и, следовательно, сопряженное уравнение (27.75), то ток сохраняется:

$$\partial_a j^a = 0. \quad (27.77)$$

В разделе 27.1.3 было показано, что при преобразованиях из группы Лоренца $\mathbb{SO}_0(1, 3)$ компоненты квадратичной формы преобразуются, как компоненты вектора. Поэтому после нормировки ток имеет прямую квантово-механическую интерпретацию: это плотность j^0 и поток j^k вероятности найти частицу в данной точке пространства-времени. При этом закон сохранения тока (27.77) приводит к сохранению полной вероятности:

$$\partial_0 \int d\mathbf{x} j^0 = 0, \quad j^0 = \psi_1^\dagger \psi_1 + \psi_2^\dagger \psi_2 + \psi_3^\dagger \psi_3 + \psi_4^\dagger \psi_4.$$

если мы рассматриваем решения уравнения Дирака, достаточно быстро убывающие на пространственной бесконечности. Важно, что такая интерпретация допустима, т.к. плотность вероятности

$$j^0 := \bar{\psi} \gamma^0 \psi = \psi^\dagger \psi$$

положительно определена.

Таким образом, уравнение Дирака (27.72), определение скалярного произведения в гильбертовом пространстве (27.74) и спинорного тока (27.76) допускают вероятностную квантово механическую интерпретацию.

Законы сохранения

Уравнение Дирака следует из принципа наименьшего действия для функционала

$$S = \int dx L = \int dx \bar{\psi} (i\gamma^a \partial_a \psi - m\psi). \quad (27.78)$$

Мы предполагаем, конечно, что этот интеграл сходится. Приведенное действие зависит от восьми вещественных функций $\text{re } \psi_A$ и $\text{im } \psi_A$, по которым проводится варьирование. Однако для вычислений удобнее выбрать в качестве независимых функций восемь комплекснозначных функций ψ_A и $\bar{\psi}_A$. Тогда вариационные производные примут вид

$$\frac{\delta S}{\delta \bar{\psi}} = i\gamma^a \partial_a \psi - m\psi = 0, \quad (27.79)$$

$$\frac{\delta S}{\delta \psi} = -i\partial_a \bar{\psi} \gamma^a - m\bar{\psi} = 0, \quad (27.80)$$

где мы, для краткости, опустили матричные индексы. Кроме этого мы пренебрегли всеми граничными слагаемыми.

Для безразмерности действия (27.78) мы приписываем полям и массе следующие размерности:

$$[\psi] = [\bar{\psi}] = l^{-3/2}, \quad [m] = l^{-1}. \quad (27.81)$$

Как видим, масса имеет ту же размерность, что и для скалярного поля.

Мы уже показали, что уравнения (27.79) и (27.80) эквивалентны. Кроме того, лагранжиан и действие обращаются в нуль на любом решении уравнений движения.

Уравнение (27.80) часто записывают в виде

$$-\bar{\psi}(i\gamma^a \overleftarrow{\partial}_a + m) = 0,$$

где стрелка $\leftarrow$ означает, что производная действует на функцию, стоящую слева.

Действие (27.78) принимает вещественные значения. Действительно, при комплексном сопряжении слагаемые в лагранжиане ведут себя следующим образом:

$$(\bar{\psi}\gamma^a\partial_a\psi)^\dagger = (\partial_a\bar{\psi}\gamma^a\psi), \quad (\bar{\psi}\psi)^\dagger = (\bar{\psi}\psi).$$

После интегрирования по частям и с учетом равенства $i^\dagger = -i$ действие принимает прежний вид. В то же время лагранжиан в действии (27.78) является комплексно-значным. Ситуацию можно подправить, записав действие в следующем виде

$$S = \int dx \left[\frac{i}{2}(\bar{\psi}\gamma^a\partial_a\psi) - \frac{i}{2}(\partial_a\bar{\psi}\gamma^a\psi) - m(\bar{\psi}\psi) \right]. \quad (27.82)$$

Теперь лагранжиан принимает вещественные значения.

Действие в виде (27.78) или (27.82) инвариантно относительно трансляций и преобразований Лоренца. Согласно первой теореме Нетер (см. раздел 20.2) этой глобальной инвариантности соответствует закон сохранения тензора энергии-импульса и момента количества движения. Для получения вещественных выражений будем использовать действие в виде (27.82).

Тензор энергии-импульса (20.43) для действия (27.82) принимает вид

$$T_a{}^b = \frac{i}{2}\bar{\psi}\gamma^b\partial_a\psi - \frac{i}{2}\partial_a\bar{\psi}\gamma^b\psi - \delta_a^b \left[\frac{i}{2}(\bar{\psi}\gamma^c\partial_c\psi) - \frac{i}{2}(\partial_c\bar{\psi}\gamma^c\psi) - m(\bar{\psi}\psi) \right]. \quad (27.83)$$

Поскольку на любом решении уравнений движения лагранжиан равен нулю, то, не ограничивая общности, тензор энергии-импульса можно выбрать в более простом виде

$$T_a{}^b = \frac{i}{2}\bar{\psi}\gamma^b\partial_a\psi - \frac{i}{2}\partial_a\bar{\psi}\gamma^b\psi. \quad (27.84)$$

Согласно первой теореме Нетер тензор энергии-импульса сохраняется, $\partial_b T_a{}^b = 0$, для любого решения уравнений Эйлера–Лагранжа.

При инфинитезимальных лоренцевых вращениях спиноры преобразуются по правилу (27.35). Поэтому тензор спинового момента количества движения (20.53) в рассматриваемом случае имеет вид

$$S_{ab}{}^c = -\frac{1}{4}i\bar{\psi}(\gamma_{ab}\gamma^c + \gamma^c\gamma_{ab})\psi. \quad (27.85)$$

При этом из первой теоремы Нетер следует сохранение полного момента количества движения $\partial_c J_{ab}{}^c = 0$, где

$$J_{ab}{}^c = M_{ab}{}^c + S_{ab}{}^c, \quad (27.86)$$

и орбитальный момент $M_{ab}{}^c$ определяется тензором энергии-импульса:

$$M_{ab}{}^c := x_b T_a{}^c - x_a T_b{}^c.$$

Вид спинового момента количества движения можно упростить.

Предложение 27.1.6. *Справедлива формула*

$$\gamma^a \gamma_{bc} + \gamma_{bc} \gamma^a = 2i \varepsilon^a{}_{bcd} \gamma^5 \gamma^d,$$

где ε^{abcd} – полностью антисимметричный тензор четвертого ранга.

Доказательство. Из полноты набора матриц (27.20) следует существование разложения

$$\gamma^a \gamma_{bc} + \gamma_{bc} \gamma^a = A^a{}_{bc} \mathbb{1} + B^a{}_{bcd} \gamma^d + C^a{}_{bcde} \gamma^{de} + D^a{}_{bcd} \gamma^5 \gamma^d + E^a{}_{bc} \gamma^5,$$

где коэффициенты разложения A, B, C, D, E определяются по формуле (27.9). Поскольку след произведения нечетного числа γ -матриц равен нулю, то

$$A^a{}_{bc} = 0, \quad C^a{}_{bcde} = 0, \quad E^a{}_{bc} = 0.$$

Коэффициенты B также равны нулю:

$$4B^a{}_{bcd} = \text{tr}(\gamma_d \gamma^a \gamma_{bc} + \gamma_d \gamma_{bc} \gamma^a) = \text{tr}(\gamma_{bc} \gamma_d \gamma^a + \gamma_{bc} \gamma^a \gamma_d) = 2\delta_d^a \text{tr} \gamma_{bc} = 0.$$

Наконец,

$$4D^a{}_{bcd} = \text{tr}(\gamma_d \gamma^5 \gamma^a \gamma_{bc} + \gamma_d \gamma^5 \gamma_{bc} \gamma^a) = \text{tr}(-\gamma^5 \gamma_{bc} \gamma_d \gamma^a + \gamma^5 \gamma_{bc} \gamma^a \gamma_d) = 8i \varepsilon^a{}_{bcd},$$

где мы воспользовались формулой (27.23). $\square$

Поэтому спинорный момент количества движения можно переписать в виде

$$S_{ab}{}^c = \frac{1}{2} \varepsilon^c{}_{abd} \bar{\psi} \gamma^5 \gamma^d \psi. \quad (27.87)$$

Отсюда, в частности, вытекает, что спинорный момент количества движения для спинорного поля со всеми опущенными индексами S_{abc} антисимметричен по всем трем индексам.

Закон сохранения момента количества движения с учетом сохранения тензора энергии-импульса принимает вид

$$\partial_c J_{ab}{}^c = T_{ab} - T_{ba} + \partial_c S_{ab}{}^c = 0. \quad (27.88)$$

Мы видим, что тензор энергии-импульса для спинорного поля в общем случае несимметричен относительно перестановки индексов.

Наконец, заметим, что действие (27.82) инвариантно относительно изменения фазы спинорного поля

$$\psi \mapsto e^{iq\phi} \psi, \quad (27.89)$$

где $\phi \in \mathbb{R}$ – фаза и $q \in \mathbb{R}$ – заряд поля. Тогда сопряженный спинор преобразуется по правилу

$$\bar{\psi} \mapsto e^{-iq\phi} \bar{\psi}$$

и сохраняющийся ток имеет вид

$$J^a = \bar{\psi} \gamma^a \psi.$$

Сохранение этого тока (27.76), которое вытекает из теоремы Нетер, мы уже обсудили.

При переходе к квантовой электродинамике мы предполагаем, что преобразование (27.89) становится локальным (калибровочным), т.е. фаза $\phi \mapsto \phi(x) \in \mathcal{C}^2(\mathbb{R}^{1,3})$

становится произвольной функцией от точки пространства-времени. Для того, чтобы сохранить калибровочную инвариантность действия вводится калибровочное поле (электромагнитный потенциал) $A_a(x)$ и частные производные заменяются на ковариантные

$$\partial_a \psi \mapsto \nabla_a \psi := \partial_a \psi - iq A_a \psi.$$

Тогда действие примет вид

$$\begin{aligned} S &= \int dx \left[\frac{i}{2} (\bar{\psi} \gamma^a \nabla_a \psi) - \frac{i}{2} (\nabla_a \bar{\psi} \gamma^a \psi) - m(\bar{\psi} \psi) \right] = \\ &= \int dx \left[\frac{i}{2} (\bar{\psi} \gamma^a \partial_a \psi) - \frac{i}{2} (\partial_a \bar{\psi} \gamma^a \psi) + q A_a (\bar{\psi} \gamma^a \psi) - m(\bar{\psi} \psi) \right]. \end{aligned} \quad (27.90)$$

Нетрудно убедиться, что это действие инвариантно относительно калибровочных преобразований

$$\psi \mapsto e^{iq\phi} \psi, \quad A_a \mapsto A_a + \partial_a \phi$$

с локальным параметром преобразования $\phi(x)$.

Действие (27.90) уже было использовано для определения операции зарядового сопряжения в разделе (27.1.4). А именно, изменение знака спинорного тока эквивалентно изменению знака заряда q . При этом все остальные слагаемые в лагранжиане не меняются при зарядовом сопряжении.

Безмассовые спиноры

Важный случай представляют безмассовые спиноры, $m = 0$. Ранее считалось, что нейтрино – это безмассовые частицы со спином $1/2$. Существующие в настоящее время экспериментальные данные указывают на то, что нейтрино имеет хотя и небольшую, но отличную от нуля массу. Ниже мы рассмотрим безмассовые спиноры, которые представляют интерес также с математической точки зрения.

Безмассовые спиноры описываются действием

$$S = \int dx \left[\frac{i}{2} (\bar{\psi} \gamma^a \partial_a \psi) - \frac{i}{2} (\partial_a \bar{\psi} \gamma^a \psi) \right]. \quad (27.91)$$

Все, сказанное выше относительно массивного случая, справедливо и для безмассовых спиноров. Однако в безмассовом случае возникают интересные дополнительные свойства.

Во-первых, действие (27.91) приобретает дополнительную *киральную* симметрию:

$$\psi \mapsto e^{i\gamma^5 \phi} \psi \quad (27.92)$$

с произвольной постоянной фазой $\phi \in \mathbb{R}$. Из свойств γ^5 -матрицы (27.18) следует правило преобразования сопряженного спинора

$$\bar{\psi} \mapsto \bar{\psi} e^{i\gamma^5 \phi}.$$

Поскольку

$$e^{i\gamma^5 \phi} \gamma^a e^{i\gamma^5 \phi} = \gamma^a,$$

то действие (27.91) инвариантно относительно киральных преобразований.

Заметим, что отсутствие массового члена в действии существенно, т.к. он нарушает киральную инвариантность:

$$(\bar{\psi} \psi) \mapsto (\bar{\psi} e^{2i\gamma^5 \phi} \psi).$$

Так как действие (27.91) инвариантно относительно киральных преобразований, то из первой теоремы Нетер следует закон сохранения

$$\partial_a j^{5a} = 0,$$

где

$$j^{5a} := \bar{\psi} \gamma^a \gamma^5 \psi \quad (27.93)$$

– *киральный ток*.

Поскольку справедливо равенство

$$\gamma^a \frac{1 \pm \gamma^5}{2} = \frac{1 \mp \gamma^5}{2} \gamma^a,$$

то в безмассовом случае уравнение Дирака для спиноров Вейля (27.41) расщепляется:

$$i\gamma^a \partial_a \psi_{\pm} = 0. \quad (27.94)$$

Если для γ -матриц выбрано киральное представление (27.4), то уравнения (27.94) можно переписать для двухкомпонентных спиноров

$$\begin{aligned} i(\partial_0 - \sigma^i \partial_i) \xi &= 0, \\ i(\partial_0 + \sigma^i \partial_i) \eta &= 0. \end{aligned}$$

Эти уравнения называются *уравнениями Вейля*.

Майорановские спиноры

Дираковский спинор описывает заряженные частицы со спином $1/2$. Если частица с полуцелым спином является нейтральной, то она описывается спинором Дирака с дополнительным условием [207].

Определение. Спинор Дирака называется *майорановским*, если он совпадает со своим зарядово сопряженным спинором:

$$\psi_c = \psi. \quad \square \quad (27.95)$$

Условие (27.95) означает, что частица совпадает со своей античастицей. Из известных частиц в настоящее время такими частицами являются нейтрон, нейтрино и др.

В киральном представлении дираковский спинор имеет вид

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix},$$

где ξ и η являются непунктирным и пунктирным спинорами, соответственно, которые совпадают со спинорами Вейля. Из формулы (27.65) следует, что условие (27.95) для майорановских спиноров устанавливает взаимно однозначное соответствие между пунктирным и непунктирным спинором:

$$\xi = -i\epsilon_c \sigma^2 \eta^* \quad \Leftrightarrow \quad \eta = i\epsilon_c \sigma^2 \xi^*. \quad (27.96)$$

То есть майорановский спинор в киральном представлении γ -матриц имеет вид

$$\psi = \begin{pmatrix} \xi \\ i\epsilon_c \sigma^2 \xi^* \end{pmatrix} = \begin{pmatrix} -i\epsilon_c \sigma^2 \eta^* \\ \eta \end{pmatrix}. \quad (27.97)$$

Существует чисто мнимое представление γ -матриц:

$$\begin{aligned}\gamma^0 &= \begin{pmatrix} 0 & -\sigma^2 \\ -\sigma^2 & 0 \end{pmatrix}, & \gamma^1 &= \begin{pmatrix} -i\sigma^3 & 0 \\ 0 & -i\sigma^3 \end{pmatrix}, \\ \gamma^2 &= \begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix}, & \gamma^3 &= \begin{pmatrix} i\sigma^1 & 0 \\ 0 & i\sigma^1 \end{pmatrix},\end{aligned}\quad (27.98)$$

которое называется *представлением Майораны*. В этом представлении γ^5 -матрица также является чисто мнимой

$$\gamma^5 = \begin{pmatrix} \sigma^2 & 0 \\ 0 & -\sigma^2 \end{pmatrix}. \quad (27.99)$$

Допустим, что непунктирный ξ и пунктирный η спиноры вещественны. Тогда, если положить фазовый множитель зарядового сопряжения равным единице, $\epsilon_c = 1$, то майорановский спинор (27.97) также будет вещественным. В майорановском представлении γ -матриц уравнение Дирака (27.72) будет вещественным, благодаря наличию множителя i в кинетическом слагаемом. Тем самым можно ограничиться вещественными четырехкомпонентными спинорами, удовлетворяющими вещественной системе уравнений Дирака. Однако подчеркнем, что это верно только в некоторой фиксированной системе координат, потому что при преобразованиях Лоренца двухкомпонентный спинор ξ преобразуется по фундаментальному представлению группы $\mathbb{SL}(2, \mathbb{C})$, которое комплексно.

Замечание. В отличие от вейлевского спинора, майорановский спинор может быть массивным. $\square$

Каноническая формулировка

Перейдем к канонической формулировке теории дираковских спиноров. Из действия (27.82) следуют выражения для обобщенных импульсов:

$$\bar{p} := \frac{\partial L}{\partial \dot{\bar{\psi}}} = \frac{i}{2} \bar{\psi} \gamma^0, \quad p := \frac{\partial L}{\partial \dot{\psi}} = -\frac{i}{2} \gamma^0 \psi. \quad (27.100)$$

Одновременные скобки Пуассона, которые отличны от нуля, имеют обычный вид:

$$[\psi_A, \bar{p}'^B] = \delta_A^B \delta(\mathbf{x} - \mathbf{x}'), \quad [\bar{\psi}^A, p'_B] = \delta_B^A \delta(\mathbf{x} - \mathbf{x}'),$$

где в правой части стоят трехмерные δ -функции.

Поскольку действие (27.78) зависит только от первых производных по времени, то в канонической формулировке возникает восемь первичных связей:

$$\bar{\Phi} := \bar{p} - \frac{i}{2} \bar{\psi} \gamma^0 = 0, \quad \Psi := p + \frac{i}{2} \gamma^0 \psi = 0. \quad (27.101)$$

Следовательно, канонический гамильтониан имеет вид

$$\begin{aligned}\mathcal{H} &= \int d\mathbf{x} (\bar{p} \dot{\bar{\psi}} + \dot{\bar{\psi}} p - L + \bar{\Phi} \lambda + \bar{\mu} \Psi) = \\ &= \int d\mathbf{x} \left[-\frac{i}{2} (\bar{\psi} \gamma^i \partial_i \psi) + \frac{i}{2} (\partial_i \bar{\psi} \gamma^i \psi) + m(\bar{\psi} \psi) + \bar{\Phi} \lambda + \bar{\mu} \Psi \right],\end{aligned}\quad (27.102)$$

где $\lambda = \{\lambda_A\}$ и $\bar{\mu} = \{\bar{\mu}_A\}$ – неопределенные множители Лагранжа, и индекс $i = 1, 2, 3$ пробегает только пространственные значения.

Условие сохранения связей во времени приводит к следующим уравнениям:

$$\begin{aligned}\dot{\bar{\Phi}} &= [\bar{\Phi}, \mathcal{H}] = -\bar{\psi}(i\gamma^i \overleftarrow{\partial}_i + m) - i\bar{\mu}\gamma^0 = 0, \\ \dot{\Psi} &= [\Psi, \mathcal{H}] = (i\gamma^i \partial_i - m)\psi + i\gamma^0\lambda = 0.\end{aligned}$$

Из этих уравнений однозначно определяются множители Лагранжа:

$$\begin{aligned}\bar{\mu} &= i\bar{\psi}(i\gamma^i \overleftarrow{\partial}_i + m)\gamma^0, \\ \lambda &= i\gamma^0(i\gamma^i \partial_i - m)\psi.\end{aligned}$$

Следовательно, связи (27.101) являются связями второго рода.

Ясно, что других связей в теории не возникает.

Согласно общим правилам (см. раздел 22), в канонической формулировке спинорного поля мы имеем $4 + 4 + 4 + 4 = 16$ канонически сопряженных переменных $(\psi_A, \bar{p}^A)$ и $(\bar{\psi}^A, p_A)$, на которые наложены 8 связей второго рода. Следовательно, имеется только $4 + 4 = 8$ физических степеней свободы. Чтобы найти гамильтониан для физических степеней свободы можно решить связи (27.101) относительно сопряженных переменных $\bar{\psi}$ и p :

$$\bar{\psi} = -2i\bar{p}\gamma^0, \quad p = -\frac{i}{2}\gamma^0\psi.$$

Подстановка решений связей в гамильтониан (27.102) приводит к следующей гамильтоновой плотности для физических степеней свободы

$$H_{\text{ph}} = -(\bar{p}\gamma^0\gamma^i\partial_i\psi) + (\partial_i\bar{p}\gamma^0\gamma^i\psi) - 2im(\bar{p}\gamma^0\psi). \quad (27.103)$$

Если в полученное выражение подставить определение обобщенных импульсов (27.100), то получим выражение гамильтониана для физических степеней свободы:

$$H_{\text{ph}} = -\frac{i}{2}(\bar{\psi}\gamma^i\partial_i\psi) + \frac{i}{2}(\partial_i\bar{\psi}\gamma^i\psi) + m(\bar{\psi}\psi). \quad (27.104)$$

Учитывая правила комплексного сопряжения квадратичных форм (27.50), этот гамильтониан принимает вещественные значения.

Гамильтониан (27.104) совпадает с компонентой T_0^0 тензора энергии-импульса в форме (27.83).

Кроме этого гамильтониан для физических степеней свободы с точностью до дивергенции равен

$$H_{\text{ph}} = \psi^\dagger H \psi,$$

где H – гамильтониан для уравнения Дирака, записанного в форме уравнения Шредингера (27.73).

27.1.6 Спиноры в геометрии Римана–Картана

При рассмотрении спинорных полей в различных моделях гравитации первое, что необходимо сделать, это перейти от пространства-времени Минковского $\mathbb{R}^{1,3}$ к четырехмерному многообразию $\mathbb{M}$ с метрикой g лоренцевой сигнатуры. После этого группа Лоренца, действующая в пространстве-времени Минковского $\mathbb{R}^{1,3}$ и в касательном

пространстве, заменяется на группу общих преобразований координат, которая действует в $\mathbb{M}$, и соответствующую ей группу $\mathbb{GL}(4, \mathbb{R})$ (матрицы Якоби), действующую в касательном пространстве.

Поскольку метрика $g = (g_{\alpha\beta})$, $\alpha, \beta = 0, 1, 2, 3$, по предположению, имеет лоренцеву сигнатуру, $\text{sign } g = (+ - - -)$, то в касательном пространстве к каждой точке можно ввести ортонормальный репер e^α_a :

$$e^\alpha_a e^\beta_b g_{\alpha\beta} = \eta_{ab}, \quad \eta_{ab} = \text{diag}(+ - - -).$$

Этот репер определен с точностью до локальных лоренцевых вращений, действующих в касательном пространстве. Поэтому для описания тензорных полей в искривленном пространстве-времени (многообразии) $\mathbb{M}$ мы рассматриваем главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{O}(1, 3))$ со структурной группой Лоренца $\mathbb{O}(1, 3)$ (см. главу 12), а также ассоциированные с $\mathbb{P}$ векторные расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{W}, \mathbb{O}(1, 3), \mathbb{P})$, типичным слоем которых является векторное пространство $\mathbb{W}$. Сечения ассоциированных расслоений, по предположению, описывают поля материи.

В случае спиноров ситуация другая. Мы рассматриваем главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{PIN}(1, 3))$ с вещественной структурной группой $\mathbb{PIN}(1, 3)$, дважды накрывающей группу Лоренца $\mathbb{O}(1, 3)$, и ассоциированные расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{PIN}(1, 3), \mathbb{P})$. Для дираковских спиноров типичным слоем является четырехмерное комплексное векторное пространство $\mathbb{V} = \mathbb{C}^4$, в котором действует точное представление группы $\mathbb{PIN}(1, 3)$.

Поскольку алгебры Ли групп $\mathbb{O}(1, 3)$ и $\mathbb{PIN}(1, 3)$ совпадают, а связности на главных расслоениях определяются 1-формами со значениями в алгебре Ли структурной группы, то возникает возможность отождествить семейство локальных форм связности на главных расслоениях $\mathbb{P}(\mathbb{M}, \pi, \mathbb{O}(1, 3))$ и $\mathbb{P}(\mathbb{M}, \pi, \mathbb{PIN}(1, 3))$; что мы и сделаем. Соответствующую связность мы будем называть *лоренцевой* или *спиновой связностью*.

Лоренцева связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{PIN}(1, 3))$ определяется семейством локальных форм связности. То есть задается координатное покрытие базы $\mathbb{M}$ и в каждой карте $\mathbb{U} \subset \mathbb{M}$ – 1-форма со значениями в алгебре Ли

$$\omega = \frac{1}{2} dx^\alpha \omega_\alpha^{ab} L_{ab}, \quad (27.105)$$

где $\omega_\alpha^{ab}(x)$ – компоненты локальной формы связности, $L_{ab} = -L_{ba}$ – базис алгебры Ли. Множитель $1/2$ возник, потому что мы предполагаем суммирование по всем значениям индексов, а не только по упорядоченным парам $a < b$ (это удобнее при проведении вычислений). Кроме этого необходимо также предположить, что в областях пересечения карт локальные формы связности (27.105) преобразуются соответствующим образом.

Семейство локальных форм связности (27.105) определяет связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{PIN}(1, 3))$, которая в свою очередь определяет связности на всех ассоциированных расслоениях. В ассоциированном расслоении с типичным слоем $\mathbb{V} = \mathbb{C}^4$ базис алгебры Ли имеет представление (27.28), поэтому локальная форма связности принимает вид

$$\omega = dx^\alpha \omega_\alpha = \frac{1}{4} dx^\alpha \omega_\alpha^{ab} \gamma_{ab} \in \text{mat}(4, \mathbb{C}). \quad (27.106)$$

Пусть в области пересечения двух карт $\mathbb{U} \cap \mathbb{U}' \neq \emptyset$ преобразование координат задается соответствующим набором функций $x^{\alpha'} = x^{\alpha'}(x)$. Тогда в области пересечения этих

карт локальная форма связности преобразуется по следующему правилу:

$$\omega'_{\alpha'} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} (U \omega_\alpha U^{-1} + \partial_\alpha U U^{-1}),$$

где $U(x) \in \mathbb{PIN}(1, 3)$ – некоторая матрица преобразования, которая может зависеть от точки $x \in \mathbb{U} \cap \mathbb{U}'$.

В дальнейшем мы ограничимся топологически тривиальными многообразиями, $\mathbb{M} \simeq \mathbb{R}^4$, которые покрываются одной картой. При этом локальные вращения в касательном пространстве без изменения координат в $\mathbb{M}$ примут вид

$$\omega'_\alpha = U \omega_\alpha U^{-1} + \partial_\alpha U U^{-1}, \quad U(x) \in \mathbb{PIN}(1, 3). \quad (27.107)$$

Это – хорошо знакомые калибровочные преобразования локальной формы связности.

Если в пространстве-времени задан репер e_α^a и лоренцева связность $\omega_\alpha^{ab} = -\omega_\alpha^{ba}$, то это означает, что на многообразии $\mathbb{M}$ задана геометрия Римана–Картана, а переменные e, ω есть переменные Картана (см. раздел 6.2).

Замечание. Если геометрия Римана–Картана задается метрикой и кручением, то главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{PIN}(1, 3))$ не определено. Поэтому для описания спинорных полей необходимо введение переменных Картана. $\square$

Ковариантная производная от дираковского спинора, по определению, имеет вид

$$\nabla_\alpha \psi = \partial_\alpha \psi - \omega_\alpha \psi = \partial_\alpha \psi - \frac{1}{4} \omega_\alpha^{ab} \gamma_{ab} \psi. \quad (27.108)$$

Нетрудно проверить, что при локальном лоренцевом вращении (точнее, $\mathbb{PIN}$ вращении),

$$\psi \mapsto \psi' = U \psi, \quad (27.109)$$

ковариантная производная преобразуется по тензорному закону

$$\nabla_\alpha \psi \mapsto \nabla'_\alpha \psi' = U \nabla_\alpha \psi, \quad (27.110)$$

где ∇'_α обозначает ковариантную производную с лоренцевой связностью (27.107).

Из правила эрмитова сопряжения для базиса алгебры Ли (27.12) следует вид ковариантной производной для сопряженного спинора

$$\nabla_\alpha \bar{\psi} = \partial_\alpha \bar{\psi} + \bar{\psi} \omega_\alpha = \partial_\alpha \bar{\psi} + \frac{1}{4} \omega_\alpha^{ab} \bar{\psi} \gamma_{ab}. \quad (27.111)$$

Отсюда следует, что квадратичная форма $\bar{\psi} \psi$ – скаляр, и ее ковариантная производная совпадает с обычной:

$$\nabla_\alpha (\bar{\psi} \psi) = \partial_\alpha (\bar{\psi} \psi).$$

Нетрудно проверить, что при локальном лоренцевом вращении (27.109) сопряженный спинор преобразуется с помощью обратной матрицы:

$$\bar{\psi} \mapsto \bar{\psi}' = \bar{\psi} U^{-1}. \quad (27.112)$$

Отсюда вытекает, что скалярная квадратичная форма при лоренцевых вращениях не преобразуется:

$$\bar{\psi} \psi \mapsto \bar{\psi}' \psi' = \bar{\psi} \psi.$$

Из свойства преобразования γ -матриц (27.54) вытекает правило преобразования векторной квадратичной формы:

$$\bar{\psi}\gamma^a\psi \mapsto \bar{\psi}'\gamma^a\psi' = (\bar{\psi}\gamma^b\psi)S_b^{-1a}, \quad (27.113)$$

где $S(x) \in \mathbb{O}(1, 3)$ – матрица локальных лоренцевых вращений. Учитывая правило преобразования репера при локальных лоренцевых вращениях,

$$e^\alpha_a \mapsto e'^\alpha_a = S_a^b e^\alpha_b,$$

получаем равенство

$$\bar{\psi}\gamma^\alpha\psi \mapsto \bar{\psi}'\gamma^\alpha\psi' = \bar{\psi}\gamma^\alpha\psi,$$

где $\gamma^\alpha(x) := e^\alpha_a(x)\gamma^a$ – четыре матрицы, зависящие от точки $x \in \mathbb{M}$, которые не преобразуются при локальных лоренцевых вращениях и преобразуются как контравариантный вектор при замене координат на $\mathbb{M}$.

Используя явный вид ковариантных производных от спиноров (27.108), (27.111) и свойства γ -матриц, нетрудно проверить, что ковариантная производная от векторной квадратичной формы имеет вид

$$\nabla_\alpha(\bar{\psi}\gamma^a\psi) = \partial_\alpha(\bar{\psi}\gamma^a\psi) - \omega_\alpha^{ab}(\bar{\psi}\gamma_b\psi).$$

Теперь все готово для построения инвариантного действия для спинорного поля. В соответствии общим правилом минимальной подстановки действие для спинорного поля в геометрии Римана–Картана принимает вид

$$S_s = \int dx L_s = \int dx \sqrt{|g|} \left[\frac{i}{2} e^\alpha_a (\bar{\psi}\gamma^a \nabla_\alpha \psi) - \frac{i}{2} e^\alpha_a (\nabla_\alpha \bar{\psi} \gamma^a \psi) - m(\bar{\psi}\psi) \right]. \quad (27.114)$$

Для его получения мы в выражение (27.82) ввели элемент объема $\sqrt{|g|} = \det e_\alpha^a$, заменили частные производные от спинорного поля на ковариантные и использовали репер для инвариантной свертки γ -матриц с ковариантной производной.

В геометрии Римана–Картана мы рассматриваем поля ψ , $\bar{\psi}$, e_α^a и ω_α^{ab} как независимые переменные, по которым проводится варьирование. Конечно, если мы хотим, чтобы переменные Картана e и ω удовлетворяли дифференциальным уравнениям, например, второго порядка (т.е. были бы динамическими переменными), то к действию (27.114) необходимо добавить гравитационную часть, квадратичную по тензорам кручения и кривизны.

В (псевдо)римановой геометрии лоренцева связность ω заменятся на лоренцеву связность (6.113)

$$\tilde{\omega}_{abc} = \frac{1}{2}(c_{abc} - c_{bca} + c_{cab}), \quad (27.115)$$

которую мы помечаем знаком тильды, где

$$c_{\alpha\beta}^c := -\partial_\alpha e_\beta^c + \partial_\beta e_\alpha^c$$

– коэффициенты неголономности. (Переход от греческих индексов к латинским и наоборот всегда осуществляется с помощью репера.) В этом случае лоренцева связность выражается через репер и его первые производные, и независимыми переменными в действии являются поля ψ , $\bar{\psi}$ и e_α^a . В таком виде описание спинорного поля в общей теории относительности было впервые предложено Фоком [208] и Вейлем [209].

Из вида действия для спинорного поля (27.114) следуют размерности полей. По определению, координаты, метрика, репер и лоренцева связность имеют следующие размерности:

$$[x^\alpha] = l, \quad [g_{\alpha\beta}] = [e_\alpha^a] = 1, \quad [\omega_\alpha^{ab}] = l^{-1}.$$

Тогда тензоры кручения и кривизны имеют размерности

$$[T_{abc}] = l^{-1}, \quad [R_{abcd}] = l^{-2}.$$

Из безразмерности действия следуют размерности спинорного поля и его массы:

$$[\psi_A] = l^{-3/2}, \quad [m] = l^{-1}. \quad (27.116)$$

По построению, действие (27.114) инвариантно относительно общих преобразований координат x^α в пространстве-времени и локальных лоренцевых (точнее, $\mathbb{PIN}$) вращений.

Из правила комплексного сопряжения спинорного тока (27.50) следует, что лагранжиан в действии (27.114) принимает вещественные значения.

Для варьирования действия (27.114) по спинорному полю, заметим, что выполнено равенство

$$\omega_\alpha \gamma^a = \gamma^a \omega_\alpha - \omega_\alpha^{ab} \gamma_b.$$

Поэтому интегрирование второго слагаемого в действии (27.114) по частям с учетом формулы для дифференцирования элемента объема (6.52) приводит к равенству

$$\int dx \sqrt{|g|} e^\alpha_a (\nabla_\alpha \bar{\psi} \gamma^a \psi) = - \int dx \sqrt{|g|} [e^\alpha_a (\bar{\psi} \gamma^a \nabla_\alpha \psi) + T_a (\bar{\psi} \gamma^a \psi)],$$

где $T_a := T_{\beta\alpha}^\beta e^\alpha_a$ – след тензора кручения. Поэтому уравнение Дирака и его сопряженное в геометрии Римана–Картана принимают вид

$$\frac{1}{\sqrt{|g|}} \frac{\delta S_s}{\delta \bar{\psi}} : \quad ie^\alpha_a \gamma^a \nabla_\alpha \psi - \frac{i}{2} T_a \gamma^a \psi - m \psi = 0, \quad (27.117)$$

$$\frac{1}{\sqrt{|g|}} \frac{\delta S_s}{\delta \psi} : \quad -ie^\alpha_a \nabla_\alpha \bar{\psi} \gamma^a + \frac{i}{2} T_a \bar{\psi} \gamma^a - m \bar{\psi} = 0. \quad (27.118)$$

Нетрудно проверить, что уравнение (27.118) получается из уравнения (27.117) сопряжением, т.е. эквивалентно уравнению (27.117).

Как и в плоском пространстве-времени действие для спинорного поля (27.114) равно нулю на уравнениях движения (27.117).

Рассмотрим подробнее член взаимодействия спиноров с лоренцевой связностью, которая содержится в ковариантных производных в действии (27.114),

$$- \sqrt{|g|} \frac{i}{8} \omega_a^{bc} [\bar{\psi} (\gamma^a \gamma_{bc} + \gamma_{bc} \gamma^a) \psi]. \quad (27.119)$$

Из этого выражения следует, что вариационная производная действия для спиноров в геометрии Римана–Картана по лоренцевой связности имеет вид

$$S_{s, \alpha}{}^{bc} := \frac{\delta S_s}{\delta \omega_\alpha^{bc}} = - \sqrt{|g|} \frac{i}{8} e^\alpha_a [\bar{\psi} (\gamma^a \gamma_{bc} + \gamma_{bc} \gamma^a) \psi] = \frac{1}{2} \sqrt{|g|} S_{bc}{}^\alpha, \quad (27.120)$$

где $S_{bc}{}^\alpha = e^\alpha_a S_{bc}{}^a$ – ковариантное обобщение спинового момента количества движения (27.85).

Если к действию для спинорного поля добавить гравитационную часть $S_G(e, \omega)$, зависящую от репера и лоренцевой связности, то уравнения движения для лоренцевой связности примут вид

$$\frac{\delta S_G}{\delta \omega_{\alpha}{}^{bc}} = -\frac{1}{2}\sqrt{|g|}S_{bc}{}^{\alpha}. \quad (27.121)$$

Таким образом, в случае спинорного поля спиновый момент количества движения является источником для лоренцевой связности.

Предложение (27.1.6) позволяет упростить вид действия для спинорного поля в геометрии Римана–Картана:

$$S_s = \int dx \sqrt{|g|} \left[\frac{i}{2} e^{\alpha}{}_a (\bar{\psi} \gamma^a \partial_{\alpha} \psi) - \frac{i}{2} e^{\alpha}{}_a (\partial_{\alpha} \bar{\psi} \gamma^a \psi) - \frac{3}{2} B_a (\bar{\psi} \gamma^5 \gamma^a \psi) - m (\bar{\psi} \psi) \right], \quad (27.122)$$

где

$$B_a := -\frac{1}{6} \omega^{bcd} \varepsilon_{bcda} \quad (27.123)$$

– полностью антисимметричная часть лоренцевой связности (псевдовектор). Мы видим, что в геометрии Римана–Картана спиноры взаимодействуют только с одной из неприводимых компонент лоренцевой связности: полностью антисимметричной частью.

Вариация действия для спинорного поля (27.114) по реперу имеет вид

$$S_{s, \alpha}{}^a := \frac{\delta S_s}{\delta e_{\alpha}{}^a} = -\sqrt{|g|} T^{\alpha}{}_a,$$

где

$$T^{\alpha}{}_a := \frac{i}{2} (\bar{\psi} \gamma^{\alpha} \nabla_a \psi) - \frac{i}{2} (\nabla_a \bar{\psi} \gamma^{\alpha} \psi) - \frac{1}{\sqrt{|g|}} e^{\alpha}{}_a L_s$$

– ковариантное обобщение канонического тензора энергии-импульса (27.83), и

$$\gamma^{\alpha} := e^{\alpha}{}_a \gamma^a, \quad \nabla_a := e^{\alpha}{}_a \nabla_{\alpha}.$$

Если к действию спинорного поля добавлено действие для репера и лоренцевой связности $S_G(e, \omega)$, то уравнения движения для репера примут вид

$$\frac{\delta S_G}{\delta e_{\alpha}{}^a} = \sqrt{|g|} T^{\alpha}{}_a.$$

Таким образом, в геометрии Римана–Картана источником в уравнениях движения для репера является ковариантное обобщение канонического тензора энергии-импульса спинорного поля.

При рассмотрении спинорного поля в аффинной геометрии возникает возможность добавления к лагранжиану неминимальных членов взаимодействия спинорного поля с репером и лоренцевой связностью. Инвариантные слагаемые неминимального взаимодействия высших размерностей квадратичны по спинорному полю. Следовательно, они представляют собой линейные комбинации ковариантных слагаемых вида (27.42). Существует два инварианта, линейных по тензору кручения,

$$\alpha_1 T_a (\bar{\psi} \gamma^a \psi), \quad \alpha_2 T^{*a} (\bar{\psi} \gamma^5 \gamma_a \psi), \quad (27.124)$$

где $T_a := T_{ba}{}^b$ – след и

$$T^{*a} := \frac{1}{6} \varepsilon^{abcd} T_{bcd}$$

– псевдослед тензора кручения (псевдовектор). При этом константы связи безразмерны:

$$[\alpha_1] = [\alpha_2] = 1.$$

Кроме того, к лагранжиану можно добавить инвариантные слагаемые с тензором кривизны:

$$\beta_1 R(\bar{\psi}\psi), \quad \beta_2 R^*(\bar{\psi}\gamma^5\psi), \quad \beta_3 R_{ab}(\bar{\psi}\gamma^{ab}\psi),$$

где $R_{ab} := R_{acb}{}^c$ – тензор Риччи, $R := R_a{}^a$ – скалярная кривизна и

$$R^* := -\frac{1}{24}\varepsilon^{abcd}R_{abcd}$$

– полностью антисимметричная часть тензора кривизны (псевдоскаляр). Напомним, что в геометрии Римана–Картана тензор Риччи имеет нетривиальную антисимметричную часть. Константы связи имеют размерность длины:

$$[\beta_1] = [\beta_2] = [\beta_3] = l.$$

В настоящей монографии модели с неминимальным взаимодействием не рассматриваются.

Зависимость уравнений движения

Поскольку действие для спинорного поля в геометрии Римана–Картана инвариантно относительно локальных лоренцевых вращений и общих преобразований координат, то согласно второй теореме Нетер между уравнениями движения существует линейная зависимость.

Из общих правил (5.10) и (5.12) следует, что при локальных лоренцевых вращениях репер и лоренцева связность ведут себя следующим образом:

$$e'^\alpha{}_a = S_a{}^b e^\alpha{}_b, \quad (27.125)$$

$$\omega'_{\alpha a}{}^b = S_a{}^c \omega_{\alpha c}{}^d S_d{}^{-1b} + \partial_\alpha S_a{}^c S_c{}^{-1b}, \quad (27.126)$$

где $S(x) \in \mathbb{O}(1, 3)$ – матрица локальных вращений. В инфинитезимальной форме

$$S_a{}^b = \delta_a^b + \omega_a{}^b, \quad S_a{}^{-1b} = \delta_a^b - \omega_a{}^b, \quad (27.127)$$

где $\omega_{ab} = -\omega_{ba}$ – угол поворота. Поэтому при локальных инфинитезимальных лоренцевых вращениях репер и лоренцева связность преобразуются по правилам

$$\delta e_\alpha{}^a = -e_\alpha{}^b \omega_b{}^a, \quad \delta e^\alpha{}_a = \omega_a{}^b e^\alpha{}_b, \quad (27.128)$$

$$\delta \omega_{\alpha a}{}^b = \omega_a{}^c \omega_{\alpha c}{}^b - \omega_{\alpha a}{}^c \omega_c{}^b + \partial_\alpha \omega_a{}^b. \quad (27.129)$$

Правило преобразования спиноров при локальных лоренцевых вращениях следуют из уравнений (27.35) и (27.44):

$$\delta \psi_\Lambda = \frac{1}{4} \omega^{ab} (\gamma_{ab} \psi)_\Lambda, \quad \delta \bar{\psi}_\Lambda = -\frac{1}{4} \omega^{ab} (\bar{\psi} \gamma_{ab})_\Lambda. \quad (27.130)$$

В этих выражениях черта над знаком вариации отсутствует, поскольку при локальных лоренцевых вращениях координаты не преобразуются.

Согласно второй теореме Нетер из инвариантности полного действия $S := S_G + S_S$ относительно локальных преобразований следует зависимость уравнений движения.

Чтобы ее записать в компактном виде, введем обозначения для вариационных производных:

$$S,^{\alpha}_{ab} := \frac{\delta S}{\delta \omega_{\alpha}^{ab}}, \quad S,^{\alpha}_a := \frac{\delta S}{\delta e_{\alpha}^a}. \quad (27.131)$$

Тогда из инвариантности действия относительно локальных лоренцевых вращений (27.128), (27.129), следует зависимость уравнений движения:

$$\sqrt{|g|} \nabla_{\alpha} \left(\frac{1}{\sqrt{|g|}} S,^{\alpha}_{ab} \right) - T_{\alpha} S,^{\alpha}_{ab} + \frac{1}{2} (S,_{ab} - S,_{ba}) - \frac{1}{4} \frac{\delta S}{\delta \psi} \gamma_{ab} \psi + \frac{1}{4} \bar{\psi} \gamma_{ab} \frac{\delta S}{\delta \bar{\psi}} = 0, \quad (27.132)$$

где опущены спинорные индексы и введена ковариантная производная

$$\nabla_{\alpha} \left(\frac{1}{\sqrt{|g|}} S,^{\alpha}_{ab} \right) := \partial_{\alpha} \left(\frac{1}{\sqrt{|g|}} S,^{\alpha}_{ab} \right) + \frac{1}{\sqrt{|g|}} (\Gamma_{\beta\alpha}^{\beta} S,^{\alpha}_{ab} - \omega_{\alpha a}^c S,^{\alpha}_{cb} - \omega_{\alpha b}^c S,^{\alpha}_{ac}).$$

В приведенные равенства введен элемент объема $\sqrt{|g|}$, чтобы из тензорных плотностей (27.131) получить тензоры, и мы воспользовались равенствами (6.52) и (6.57).

При инфинитезимальных преобразованиях координат репер и лоренцева связность преобразуются как ковекторы по правилам (2.113), (2.114). Спиноры преобразуются как скаляры:

$$\delta \psi = -u^{\alpha} \partial_{\alpha} \psi, \quad \delta \bar{\psi} = -u^{\alpha} \partial_{\alpha} \bar{\psi},$$

где $u^{\alpha}(x)$ – параметр бесконечно малого преобразования координат:

$$x^{\alpha} \mapsto x'^{\alpha} = x^{\alpha} + u^{\alpha}.$$

Из инвариантности полного действия относительно общих преобразований координат, согласно второй теореме Нетер, следует зависимость уравнений движения:

$$\begin{aligned} \sqrt{|g|} \nabla_{\beta} \left(\frac{1}{\sqrt{|g|}} S,^{\beta}_{ab} \right) \omega_{\beta}^{ab} - T_{\beta} S,^{\beta}_{ab} \omega_{\alpha}^{ab} - S,^{\beta}_{ab} R_{\alpha\beta}^{ab} + \\ + \sqrt{|g|} \nabla_{\beta} \left(\frac{1}{\sqrt{|g|}} S,^{\beta}_a \right) e_{\alpha}^a - T_{\beta} S,^{\beta}_a e_{\alpha}^a - S,^{\beta}_a T_{\alpha\beta}^a + S,^{\beta}_a e_{\beta}^b \omega_{\alpha b}^a - \\ - \frac{\delta S}{\delta \psi} \partial_{\alpha} \psi - \partial_{\alpha} \bar{\psi} \frac{\delta S}{\delta \bar{\psi}} = 0, \end{aligned}$$

где $R_{\alpha\beta}^{ab}$ и $T_{\alpha\beta}^a$ – тензоры кривизны и кручения в переменных Картана (5.57), (5.56). Полученное равенство можно упростить и записать в явно ковариантном виде, если использовать равенство (27.132):

$$\begin{aligned} \sqrt{|g|} \nabla_{\beta} \left(\frac{1}{\sqrt{|g|}} S,^{\beta}_a \right) e_{\alpha}^a - T_{\beta} S,^{\beta}_a e_{\alpha}^a - S,^{\beta}_a T_{\alpha\beta}^a - S,^{\beta}_{ab} R_{\alpha\beta}^{ab} - \\ - \frac{\delta S}{\delta \psi} \nabla_{\alpha} \psi - \nabla_{\alpha} \bar{\psi} \frac{\delta S}{\delta \bar{\psi}} = 0. \quad (27.133) \end{aligned}$$

Таким образом, из инвариантности действия относительно локальных лоренцевых вращений согласно второй теореме Нетер вытекает шесть линейных соотношений между уравнениями движения (27.132), и четыре соотношения (27.133) являются следствием инвариантности относительно общих преобразований координат.

Глава 28

Геометрия поверхностей

Геометрия поверхностей является классическим разделом дифференциальной геометрии. Она относительно проста, и многие вопросы можно решить аналитически, что позволяет глубже понять проблемы, возникающие в более высоких размерностях. Кроме этого геометрия поверхностей в последние годы приобрела особую актуальность в связи с большим числом физических приложений и, в первую очередь, в связи с интересом к теории струн, мировая поверхность которых представляет собой двумерное многообразие с метрикой лоренцевой сигнатуры.

28.1 Геометрия Римана–Картана на поверхностях

Минимальная размерность многообразия, на котором возможно существование нетривиального кручения и кривизны, равна двум. Настоящая глава посвящена именно этому случаю.

Пусть на поверхности задана геометрия Римана–Картана, т.е. задана метрика и кручение, а тензор неметричности тождественно равен нулю, $Q_{\alpha\beta\gamma} = 0$. Мы предполагаем, что на поверхности задана метрика g либо лоренцевой, либо евклидовой сигнатуры.

В двумерном случае многие формулы дифференциальной геометрии значительно упрощаются, благодаря наличию антисимметричного (псевдо)тензора второго ранга

$$\varepsilon_{ab} = -\varepsilon_{ba}, \quad \varepsilon_{01} = \varepsilon_{12} = 1. \quad (28.1)$$

Приставка “псевдо” означает, что при пространственном отражении или обращении времени ε_{ab} меняет знак. В евклидовом случае, конечно, есть только пространственное отражение одной из координат. Напомним, что в псевдоевклидовом и евклидовом случаях латинские индексы принимают значения $\{0, 1\}$ и $\{1, 2\}$, соответственно. Контравариантные компоненты этого тензора зависят от сигнатуры метрики и отличаются знаком

$$\varepsilon^{01} = -1, \quad \varepsilon^{12} = 1. \quad (28.2)$$

Свойства полностью антисимметричного тензора второго ранга приведены в приложении 37.4. Существование антисимметричного тензора второго ранга означает, что произвольное антисимметричное тензорное поле второго ранга, $X_{ab} = -X_{ba}$, взаимно однозначно определяется псевдоскалярным полем X^* :

$$X_{ab} = \bigoplus \frac{1}{2} \varepsilon_{ab} X^*, \quad \text{где} \quad X^* = \varepsilon^{ab} X_{ab}. \quad (28.3)$$

Здесь и далее верхний знак внутри окружности $\oplus$ или $\otimes$ относится к римановой поверхности, а нижний – к лоренцевой.

Благодаря антисимметрии, тензор кривизны в геометрии Римана–Картана всегда можно представить в виде

$$R_{abcd} = \oplus \frac{1}{2} \varepsilon_{ab} \varepsilon_{cd} R = \frac{1}{2} (\eta_{ac} \eta_{bd} - \eta_{ad} \eta_{bc}) R, \quad (28.4)$$

где R – скалярная кривизна. Здесь и далее в этом разделе римановы и лоренцевы поверхности будут рассматриваться параллельно. Поэтому символ η_{ab} будет обозначать и символ Кронекера, и метрику Минковского. Из формулы (28.4) следует, что у тензора кривизны только одна неприводимая компонента – скалярная кривизна – отлична от нуля. Соотношение (28.4) справедливо как в римановой геометрии, так и в геометрии Римана–Картана при нетривиальном кручении. В частности, тензор Риччи всегда симметричен

$$R_{ab} = \frac{1}{2} \eta_{ab} R. \quad (28.5)$$

Отсюда следует, что вакуумные уравнения Эйнштейна (23.2) при нулевой космологической постоянной в двумерном пространстве-времени удовлетворяются тождественно.

Все квадратичные инварианты тензора кривизны выражаются через квадрат скалярной кривизны. Например,

$$R_{abcd} R^{abcd} = R^2, \quad (28.6)$$

$$R_{ab} R^{ab} = \frac{1}{2} R^2. \quad (28.7)$$

В двумерном случае упрощается также общий вид тензора кручения. Нетрудно проверить, что он взаимно однозначно представляется либо своим следом:

$$T_{abc} = \eta_{ac} T_b - \eta_{bc} T_a, \quad T_b := T_{ab}{}^a, \quad (28.8)$$

либо псевдоследом:

$$T_{abc} = \oplus \frac{1}{2} \varepsilon_{ab} T_c^*, \quad T_c^* := T_{abc} \varepsilon^{ab}. \quad (28.9)$$

Псевдовектор T_c^* и след T_a тензора кручения взаимно однозначно связаны между собой простыми соотношениями:

$$T_a^* = 2 \varepsilon_{ab} T^b, \quad T_a = \frac{1}{2} \varepsilon_{ab} T^{*b}.$$

Квадратичный инвариант тензора кручения имеет вид

$$T_{abc} T^{abc} = 2 T_a T^a = \oplus \frac{1}{2} T_a^* T^{*a}. \quad (28.10)$$

При проведении вычислений часто бывает удобно переходить от ковариантной производной $\tilde{\nabla}_\alpha$, определяемой символами Кристоффеля и возникающей при интегрировании по частям, к ковариантной производной ∇_α , определяемой полной метрической связностью с кручением. Эта связь осуществляется с помощью соотношения

$$\Gamma_{\alpha\beta\gamma} = \tilde{\Gamma}_{\alpha\beta\gamma} + g_{\alpha\gamma} T_\beta - g_{\alpha\beta} T_\gamma,$$

которое следует из (6.21), (28.8).

Поскольку тождества Бианки антисимметричны по трем индексам, то они автоматически удовлетворяются в двух измерениях и не накладывают никаких ограничений на тензор кривизны и кручения.

Теперь обсудим переменные Картана: репер e_α^a и лоренцеву связность ω_α^{ab} . Наличие антисимметричного (псевдо)тензора второго ранга позволяет воспользоваться следующей параметризацией лоренцевой связности в геометрии Римана–Картана:

$$\omega_\alpha^{ab} = \omega_\alpha \varepsilon^{ab}. \quad (28.11)$$

1-форму ω_α также будем называть лоренцевой связностью. В евклидовом случае вместо термина “лоренцева связность” следует говорить “ $\mathbb{SO}(2)$ связность”. Для определенности, мы будем говорить лоренцева связность, имея в виду оба случая. При пространственных отражениях или обращении времени эта форма, так же, как и антисимметричный тензор ε^{ab} , меняет знак.

Из определения (28.11) следует, что

$$\omega_a = \bigoplus \frac{1}{2} \omega_{abc} \varepsilon^{bc} = \bigoplus \varepsilon_a^b (c_b + T_b), \quad (28.12)$$

где c_b – след коэффициентов неголономности (6.103).

Закон преобразования связности ω_α при локальных лоренцевых вращениях неоднороден (5.12). Прямые вычисления показывают, что при вращении на угол $\omega(x)$, лоренцева связность преобразуется так же, как и электромагнитное поле

$$\omega_\alpha \mapsto \omega'_\alpha = \omega_\alpha + \partial_\alpha \omega, \quad (28.13)$$

т.к. группа $\mathbb{SO}_0(1, 1)$ абелева.

Тензор кривизны поверхности в переменных Картана представляется в виде

$$R_{\alpha\beta}^{ab} = F_{\alpha\beta} \varepsilon^{ab}, \quad F_{\alpha\beta} := \partial_\alpha \omega_\beta - \partial_\beta \omega_\alpha. \quad (28.14)$$

Отсутствие квадратичных слагаемых в тензоре кривизны связано с тем, что в двумерном случае группы Лоренца $\mathbb{SO}_0(1, 1)$ и вращений $\mathbb{SO}(2)$ являются абелевыми. При этом скалярная кривизна имеет вид

$$R = 2\varepsilon^{\alpha\beta} \partial_\alpha \omega_\beta = \varepsilon^{\alpha\beta} F_{\alpha\beta}, \quad F_{\alpha\beta} = \bigoplus \frac{1}{2} \varepsilon_{\alpha\beta} R. \quad (28.15)$$

где введен антисимметричный тензор

$$\varepsilon^{\alpha\beta} = e^\alpha_a e^\beta_b \varepsilon^{ab}, \quad \nabla_\alpha \varepsilon^{\beta\gamma} = 0. \quad (28.16)$$

Выше введен репер e_α^a , который в двумерном случае называется *диадой*, и определяется соотношением

$$g_{\alpha\beta} = e_\alpha^a e_\beta^b \eta_{ab}$$

с точностью до локальных лоренцевых вращений.

Приведем также явное выражение для псевдоследа тензора кручения в картановых переменных, которое часто используется в приложениях,

$$T^{*a} = 2\varepsilon^{\alpha\beta} (\partial_\alpha e_\beta^a + \omega_\alpha e_\beta^b \varepsilon_b^a). \quad (28.17)$$

В римановом пространстве кручение равно нулю, и из уравнения (28.17) следует выражение лоренцевой связности через производные от диады:

$$\tilde{\omega}_\alpha = -e_\alpha^a \varepsilon^{\beta\gamma} \partial_\beta e_{\gamma a}. \quad (28.18)$$

В приложениях часто используется также полностью антисимметричная тензорная плотность единичного веса

$$\hat{\varepsilon}_{\alpha\beta} = \frac{1}{\sqrt{|g|}} \varepsilon_{\alpha\beta}, \quad \hat{\varepsilon}^{\alpha\beta} = \sqrt{|g|} \varepsilon^{\alpha\beta}. \quad (28.19)$$

Ее компоненты во всех системах координат равны по модулю единице, и частная производная обращается в нуль:

$$\partial_\alpha \hat{\varepsilon}_{\beta\gamma} = \partial_\alpha \hat{\varepsilon}^{\beta\gamma} = 0. \quad (28.20)$$

Определитель диады можно записать в различных видах:

$$\sqrt{|g|} = \det e_\alpha{}^a = e_0{}^a e_1{}^b \varepsilon_{ab} = \bigoplus \hat{\varepsilon}^{\alpha\beta} e_\alpha{}^0 e_\beta{}^1 = \bigoplus \frac{1}{2} \hat{\varepsilon}^{\alpha\beta} e_\alpha{}^a e_\beta{}^b \varepsilon_{ab}. \quad (28.21)$$

Для вычислений, проводимых с мономами третьего и более высокого порядка по лоренцевым векторам, полезны следующие тождества:

$$\varepsilon_{ab} \varepsilon_{cd} + \varepsilon_{bc} \varepsilon_{ad} + \varepsilon_{ca} \varepsilon_{bd} = 0, \quad (28.22)$$

$$\eta_{ab} \varepsilon_{cd} + \eta_{da} \varepsilon_{bc} + \eta_{cd} \varepsilon_{ab} + \eta_{bc} \varepsilon_{da} = 0. \quad (28.23)$$

Тождество (28.22), в котором слагаемые отличаются циклической перестановкой первых трех индексов, следует из того, что в двумерном пространстве не существует тензора, антисимметричного по трем индексам. Тождество (28.23) аналогично тождествам Фирца для γ матриц Дирака и проверяется прямой подстановкой.

Тождество, связывающее скалярную кривизну в (псевдо)римановой геометрии со скалярной кривизной в геометрии Римана–Картана (6.96), в двумерном случае существенно упрощается:

$$R - \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} \varepsilon^\alpha{}_\beta T^{*\beta}) = \tilde{R}. \quad (28.24)$$

То есть скалярные кривизны отличаются на дивергенцию.

28.2 Неголономный базис в двух измерениях

В двумерном пространстве-времени Римана–Картана удобно использовать неголономный базис в касательном пространстве, определяемый диадой:

$$e_a := \partial_a := e^\alpha{}_a \partial_\alpha. \quad (28.25)$$

Многие геометрические объекты выражаются через коэффициенты неголономности $c_{ab}{}^c$, которые определяются коммутатором базисных векторных полей,

$$[e_a, e_b] := c_{ab}{}^c e_c. \quad (28.26)$$

Из выражения (28.25) следует, что коэффициенты неголономности имеют вид

$$\begin{aligned} c_{ab}{}^c &= (e^\alpha{}_a \partial_\alpha e^\beta{}_b - e^\alpha{}_b \partial_\alpha e^\beta{}_a) e_\beta{}^c \\ &= -e^\alpha{}_a e^\beta{}_b (\partial_\alpha e_\beta{}^c - \partial_\beta e_\alpha{}^c). \end{aligned} \quad (28.27)$$

В двух измерениях они взаимно однозначно определяются следом

$$c_{ab}{}^c = \delta_a^c c_b - \delta_b^c c_a, \quad c_b := c_{ab}{}^a. \quad (28.28)$$

Из уравнения (28.27) вытекает, что коэффициенты неголономности инвариантны относительно общих преобразований координат. При локальном лоренцевом повороте на угол $\omega(x)$ они преобразуются, как лоренцева связность:

$$\delta c_b = -\omega \varepsilon_b{}^c c_c - \varepsilon_b{}^c \partial_c \omega. \quad (28.29)$$

Приведем два полезных тождества:

$$\varepsilon^{ab} \partial_a e_b{}^\alpha = \varepsilon^{ab} c_b, \quad \varepsilon^{\alpha\beta} \partial_\alpha e_\beta{}^a = -\varepsilon^{ab} c_b. \quad (28.30)$$

Лоренцева связность в неголономном базисе определяется соотношением

$$\omega_a := e_a{}^\alpha \omega_\alpha. \quad (28.31)$$

Скалярная кривизна R и след тензора кручения $T_b := T_{ab}{}^a$ в неголономном базисе имеют вид

$$R = 2\varepsilon^{ab}(\partial_a \omega_b + c_a \omega_b), \quad (28.32)$$

$$T_b = -c_b + \varepsilon_b{}^c \omega_c. \quad (28.33)$$

В геометрии Римана–Картана диада и лоренцева связность являются независимыми переменными. В то же время по заданной диаде всегда можно построить вторую лоренцеву связность и другие геометрические объекты, которые соответствуют нулевому кручению. В этом случае, т.е. в (псевдо)римановой геометрии, лоренцева связность взаимно однозначно определяется коэффициентами неголономности:

$$\tilde{\omega}_a = \varepsilon_a{}^b c_b. \quad (28.34)$$

Разность между двумя лоренцевыми связностями определяется тензором кручения:

$$\omega_a = \tilde{\omega}_a + \epsilon_a{}^b T_b. \quad (28.35)$$

Из уравнения (28.32) следует, что скалярная кривизна в римановой геометрии определяется коэффициентами неголономности следующей простой формулой

$$\tilde{R} = 2\partial_a c^a + 2c_a c^a \quad (28.36)$$

28.3 Выбор системы координат

В двумерном пространстве метрика содержит три независимые компоненты, при этом две из них можно фиксировать за счет выбора системы координат (фиксирования калибровки). Таким образом, после фиксирования координат многие геометрические характеристики выражаются через одну функцию, что существенно упрощает анализ двумерных моделей. В настоящем разделе приведен ряд формул в различных системах координат, которые часто встречаются в приложениях.

28.3.1 Диагональная калибровка

В *диагональной калибровке*, которую используют в вычислениях, метрика и ее обратная имеют диагональный вид:

$$g_{\alpha\beta} = \begin{pmatrix} g_{00} & 0 \\ 0 & g_{11} \end{pmatrix}, \quad g^{\alpha\beta} = \begin{pmatrix} 1/g_{00} & 0 \\ 0 & 1/g_{11} \end{pmatrix}. \quad (28.37)$$

Эта калибровка фиксирует только одну из двух произвольных функций, параметризующих преобразования координат. Оставшийся функциональный произвол можно использовать для фиксирования других переменных, входящих в модель. Символы Кристоффеля (6.24) для метрики (28.37) имеют следующий вид:

$$\begin{aligned} \tilde{\Gamma}_{00}^0 &= \frac{1}{2} \frac{\dot{g}_{00}}{g_{00}}, & \tilde{\Gamma}_{01}^0 &= \tilde{\Gamma}_{10}^0 = \frac{1}{2} \frac{g'_{00}}{g_{00}}, & \tilde{\Gamma}_{11}^0 &= -\frac{1}{2} \frac{\dot{g}_{11}}{g_{00}}, \\ \tilde{\Gamma}_{00}^1 &= -\frac{1}{2} \frac{g'_{00}}{g_{11}}, & \tilde{\Gamma}_{01}^1 &= \tilde{\Gamma}_{10}^1 = \frac{1}{2} \frac{\dot{g}_{11}}{g_{11}}, & \tilde{\Gamma}_{11}^1 &= \frac{1}{2} \frac{g'_{11}}{g_{11}}, \end{aligned} \quad (28.38)$$

Где точка и штрих обозначают дифференцирование по x^0 и x^1 , соответственно. У тензора кривизны со всеми опущенными индексами имеется только одна нетривиальная линейно независимая компонента

$$R_{0101} = \frac{1}{2}(\ddot{g}_{11} + g''_{00}) - \frac{1}{4} \left(\frac{\dot{g}_{11}^2}{g_{11}} + \frac{g_{00}^2}{g_{00}} + \frac{\dot{g}_{00}\dot{g}_{11}}{g_{00}} + \frac{g'_{00}g'_{11}}{g_{11}} \right). \quad (28.39)$$

Остальные отличные от нуля компоненты получаются простой перестановкой индексов. Тензор Риччи в рассматриваемой калибровке диагонален:

$$R_{00} = \frac{1}{g_{11}} R_{0101}, \quad R_{11} = \frac{1}{g_{00}} R_{0101}. \quad (28.40)$$

Скалярная кривизна равна следующему выражению

$$R = \frac{\dot{g}_{11} + g''_{00}}{g_{00}g_{11}} - \frac{1}{2g_{00}g_{11}} \left(\frac{\dot{g}_{11}^2}{g_{11}} + \frac{g_{00}^2}{g_{00}} + \frac{\dot{g}_{00}\dot{g}_{11}}{g_{00}} + \frac{g'_{00}g'_{11}}{g_{11}} \right). \quad (28.41)$$

28.3.2 Конформная калибровка для римановых поверхностей

Для двумерной римановой метрики часто используют *конформную калибровку*, т.е. такую систему координат, в которой метрика является вейлевски евклидовой:

$$ds^2 = \Phi(dx^2 + dy^2), \quad (28.42)$$

где $\Phi = \Phi(x, y)$ – некоторая функция от координат $x = x^1$ и $y = x^2$. *Конформный множитель* Φ предполагается всюду отличным от нуля, иначе метрика будет вырождена, а соответствующая скалярная кривизна может обратиться в бесконечность (это зависит от того, как функция Φ стремится к нулю). Координаты в которых метрика имеет вид (28.42) называются *изотермическими*. Для положительно определенной метрики известна локальная теорема существования изотермических координат.

Теорема 28.3.1. *Если метрика принадлежит классу $\mathcal{C}^3(\mathbb{M})$, тогда для любой точки $x \in \mathbb{M}$ существует окрестность, в которой можно выбрать изотермические координаты.*

Доказательство. См., например, [54], теорема 2.5.14. $\square$

Заметим, что конформный множитель не является скалярным полем, т.к. в других системах координат метрика в общем случае не будет иметь вейлевски евклидова вида.

Если метрика имеет вид (28.42), то удобно ввести комплексные координаты (1.87). Тогда интервал запишется в виде

$$ds^2 = \Phi dz d\bar{z}.$$

После конформного преобразования $z \rightarrow \zeta(z)$ интервал сохраняет вейлевски евклидову форму

$$ds^2 = \Phi \frac{dz}{d\zeta} \frac{d\bar{z}}{d\bar{\zeta}} d\zeta d\bar{\zeta}.$$

То есть изотермические координаты определены с точностью до конформных преобразований. По этой причине калибровка (28.42) называется конформной. В изотермических координатах отличны от нуля только два символа Кристоффеля:

$$\tilde{\Gamma}_{zz}^z = \frac{\partial_z \Phi}{\Phi}, \quad \tilde{\Gamma}_{\bar{z}\bar{z}}^{\bar{z}} = \frac{\partial_{\bar{z}} \Phi}{\Phi}.$$

Отличные от нуля линейно независимые компоненты тензора кривизны имеют вид

$$\tilde{R}_{z\bar{z}\bar{z}}^{\bar{z}} = \tilde{R}_{\bar{z}zz}^z = \frac{\Phi \partial_{z\bar{z}}^2 \Phi - \partial_z \Phi \partial_{\bar{z}} \Phi}{\Phi^2} \quad (28.43)$$

У тензора Риччи отличны от нуля только две компоненты:

$$\tilde{R}_{z\bar{z}} = \tilde{R}_{\bar{z}z} = \frac{\Phi \partial_{z\bar{z}}^2 \Phi - \partial_z \Phi \partial_{\bar{z}} \Phi}{\Phi^2}. \quad (28.44)$$

Это приводит к следующему выражению для скалярной кривизны

$$\tilde{R} = 4 \frac{\Phi \partial_{z\bar{z}}^2 \Phi - \partial_z \Phi \partial_{\bar{z}} \Phi}{\Phi^3} = \frac{4}{\Phi} \partial_{z\bar{z}}^2 \ln |\Phi| \quad (28.45)$$

Часто для конформного множителя используют следующую параметризацию

$$\Phi = e^{2\phi}. \quad (28.46)$$

Тогда

$$\tilde{R} = 2 e^{-2\phi} \Delta \phi, \quad (28.47)$$

где $\Delta \phi := 4 \partial_{z\bar{z}}^2 \phi$ – обычный лапласиан.

Для поверхностей постоянной кривизны (см. раздел 29) скалярная кривизна постоянна,

$$\tilde{R} = -2K = \text{const}, \quad (28.48)$$

где K – гауссова кривизна. Это уравнение в изотермических координатах сводится к уравнению Лиувилля

$$\Delta \phi + K e^{2\phi} = 0. \quad (28.49)$$

Уравнение Лиувилля является интегрируемым, и его общее решение приведено в разделе 29.4.

28.3.3 Конформная калибровка для лоренцевых поверхностей

На псевдоримановой поверхности $\mathbb{M}$ интервал в конформной калибровке имеет вид

$$ds^2 = \Phi(dt^2 - dx^2) = \Phi du dv \quad (28.50)$$

где $\Phi = \Phi(u, v)$ – конформный множитель и u и v – координаты светового конуса (1.93). Соответствующая метрика и ее обратная в конусных координатах имеют вид

$$g_{\alpha\beta} = \frac{\Phi}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad g^{\alpha\beta} = \frac{2}{\Phi} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (28.51)$$

Теорема 28.3.2. *Если метрика дважды дифференцируема $g_{\alpha\beta} \in C^2(\mathbb{M})$, то каждая точка $\mathbb{M}$ имеет координатную окрестность, в которой метрика является вейлевски лоренцевой (28.50).*

Доказательство. Докажем, что преобразованием координат произвольную обратную метрику можно привести к виду (28.51). Рассмотрим преобразование координат $x^\alpha \mapsto y^\alpha(x)$ в произвольной координатной окрестности. Тогда компоненты обратной метрики преобразуются по правилу

$$g^{\alpha\beta} \mapsto \tilde{g}^{\alpha\beta} = g^{\gamma\delta} \frac{\partial y^\alpha}{\partial x^\gamma} \frac{\partial y^\beta}{\partial x^\delta}.$$

Для упрощения формул введем обозначение

$$g^{\alpha\beta} = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

где $a, b, c, d \in C^2(\mathbb{M})$. Не ограничивая общности, можно считать, что исходная система координат выбрана таким образом, что $a > 0$. Потребуем, чтобы преобразованные компоненты метрики имели вид (28.51). Тогда функции перехода должны удовлетворять двум уравнениям:

$$\begin{aligned} a \left(\frac{\partial y^0}{\partial x^0} \right)^2 + 2b \frac{\partial y^0}{\partial x^0} \frac{\partial y^0}{\partial x^1} + c \left(\frac{\partial y^0}{\partial x^1} \right)^2 &= 0, \\ a \left(\frac{\partial y^1}{\partial x^0} \right)^2 + 2b \frac{\partial y^1}{\partial x^0} \frac{\partial y^1}{\partial x^1} + c \left(\frac{\partial y^1}{\partial x^1} \right)^2 &= 0. \end{aligned} \quad (28.52)$$

Эти квадратные уравнения относительно $\partial y^0/\partial x^0$ и $\partial y^1/\partial x^0$ легко решаются. В результате получаем два эквивалентных линейных уравнения:

$$\frac{\partial y^0}{\partial x^0} + \lambda_0 \frac{\partial y^0}{\partial x^1} = 0, \quad \frac{\partial y^1}{\partial x^0} + \lambda_1 \frac{\partial y^1}{\partial x^1} = 0, \quad (28.53)$$

где

$$\lambda_0 := \frac{b - \sqrt{b^2 - ac}}{a}, \quad \lambda_1 := \frac{b + \sqrt{b^2 - ac}}{a}.$$

Поскольку метрика лоренцева, то $b^2 - ac > 0$.

Предположим, что существуют решения уравнений (28.53) такие, что $\{\partial_\alpha y^0\} \neq 0$ и $\{\partial_\alpha y^1\} \neq 0$. Тогда кривые

$$y^0(x) = C^0, \quad y^1(x) = C^1, \quad C^{0,1} = \text{const}, \quad (28.54)$$

определяют два семейства характеристик (16.58) для волнового уравнения на лоренцевой поверхности. Действительно, уравнения (28.52) это и есть уравнения характеристик.

Для дальнейшего нам понадобится

Лемма 28.3.1. Пусть функция $W(x) \in \mathcal{C}^2$ такова, что $\partial W / \partial x^1 \neq 0$. Для того, чтобы семейство кривых $W = C = \text{const}$ давало семейство характеристик волнового уравнения (16.56), необходимо и достаточно, чтобы выражение $W = C$ было общим интегралом одного из обыкновенных дифференциальных уравнений

$$\frac{dx^1}{dx^0} = \lambda_0(x^0, x^1), \quad \frac{dx^1}{dx^0} = \lambda_1(x^0, x^1). \quad (28.55)$$

Доказательство. Из теоремы о неявной функции следует, что уравнение $W = C$ при $\partial_1 W \neq 0$ определяет неявную функцию $x^1 = x^1(x^0, C)$, при этом

$$\frac{dx^1}{dx^0} = -\frac{\partial_0 W}{\partial_1 W}.$$

Если функция $W = C$ является характеристикой волнового уравнения, то она удовлетворяет квадратному уравнению

$$a\partial_0 W^2 + 2b\partial_0 W \partial_1 W + c\partial_1 W^2 = 0,$$

которое приводит к уравнению для неявной функции

$$a \left(\frac{dx^1}{dx^0} \right)^2 - 2b \frac{dx^1}{dx^0} + c = 0,$$

которое сводится к одному из двух уравнений (28.55).

Легко проверить также обратное утверждение. □

Согласно общей теории дифференциальных уравнений решения обыкновенных дифференциальных уравнений (28.55) существуют в, возможно, меньшей окрестности. Согласно доказанной лемме, это означает существование двух семейств характеристик $y^0(x) = C_0$ и $y^1(x) = C_1$ класса $\mathcal{C}^1$. Поскольку $\lambda_{0,1} \in \mathcal{C}^2$, то и $y^{0,1} \in \mathcal{C}^2$.

Далее, из уравнений (28.53) следует неравенство

$$\det \frac{\partial y}{\partial x} := \partial_0 y^0 \partial_1 y^1 - \partial_1 y^0 \partial_0 y^1 = (\lambda_1 - \lambda_0) \partial_1 y^0 \partial_1 y^1 = \frac{2\sqrt{b^2 - ac}}{a} \partial_1 y^0 \partial_1 y^1 \neq 0.$$

Это означает, что преобразование координат $x \mapsto y$ невырождено. □

Замечание. Приведенное доказательство существования системы координат, в которой метрика имеет вейлевски лоренцев вид (28.50), можно обобщить на евклидов случай (см., например, [210], §1.3, раздел 4). Однако при этом нужно потребовать вещественную аналитичность компонент метрики. □

Система координат, в которой метрика является вейлевски лоренцевой, определена неоднозначно. Преобразования светоподобных координат

$$u, v \mapsto u'(u), v'(v), \quad (28.56)$$

которые приводят только к изменению конформного множителя, будем, как и в евклидовом случае, называть конформными.

Выражения для символов Кристоффеля и тензора кривизны аналогичны соответствующим формулам в римановом случае. В конусных координатах отличны от нуля только два символа Кристоффеля:

$$\tilde{\Gamma}_{uu}^u = \frac{\partial_u \Phi}{\Phi}, \quad \tilde{\Gamma}_{vv}^v = \frac{\partial_v \Phi}{\Phi}. \quad (28.57)$$

Линейно независимые отличные от нуля компоненты тензора кривизны имеют вид

$$\tilde{R}_{uv}^v = \tilde{R}_{vu}^u = \frac{\Phi \partial_{uv}^2 \Phi - \partial_u \Phi \partial_v \Phi}{\Phi^2}. \quad (28.58)$$

У тензора Риччи отличны от нуля только недиагональные компоненты:

$$\tilde{R}_{uv} = \tilde{R}_{vu} = \frac{\Phi \partial_{uv}^2 \Phi - \partial_u \Phi \partial_v \Phi}{\Phi^2}. \quad (28.59)$$

Это приводит к следующему выражению для скалярной кривизны

$$\tilde{R} = 4 \frac{\Phi \partial_{uv}^2 \Phi - \partial_u \Phi \partial_v \Phi}{\Phi^3} = \frac{4}{\Phi} \partial_{uv}^2 \ln |\Phi|. \quad (28.60)$$

Если конформный множитель Φ положителен, то его можно параметризовать с помощью экспоненты

$$\Phi := e^{2\phi}, \quad (28.61)$$

где $\phi = \phi(u, v)$ – некоторая функция координат. Тогда скалярная кривизна (28.60) принимает вид

$$\tilde{R} = 2e^{-2\phi} \square \phi, \quad (28.62)$$

где $\square$ – оператор Даламбера на плоскости Минковского (1.123).

Для поверхностей постоянной кривизны (см. раздел 29) скалярная кривизна постоянна,

$$\tilde{R} = -2K = \text{const}, \quad (28.63)$$

где K – гауссова кривизна. Это уравнение в конформной калибровке сводится к уравнению Лиувилля

$$\square \phi + K e^{2\phi} = 0. \quad (28.64)$$

Уравнение Лиувилля является интегрируемым, и его общее решение приведено в разделе 29.4.

28.4 Координаты светового конуса

В трех предыдущих разделах мы использовали преобразования координат для уменьшения числа неизвестных компонент метрики. В настоящем разделе мы рассмотрим удобный способ введения базиса в касательном пространстве для лоренцевой поверхности, оставляя полную свободу в выборе координат на многообразии.

Если на поверхности задана метрика лоренцевой сигнатуры, то при проведении вычислений бывает удобно использовать координаты светового конуса для векторов касательного пространства, которые совпадают с координатами светового конуса (1.93) на плоскости Минковского. Преимущество этих координат связано с тем, что локальное лоренцево вращение не перемешивает компоненты векторного поля,

т.к. каждая из них преобразуется по неприводимому одномерному представлению связной компоненты единицы группы $\mathbb{SO}_0(1, 1)$.

Преобразование компонент вектора $X = X^a e_a$ относительно ортонормального базиса в касательном пространстве при переходе от декартовых координат к *координатам светового конуса* определим следующим образом:

$$X^\pm := \frac{1}{\sqrt{2}}(X^0 \pm X^1), \quad (28.65)$$

Якобиан преобразования $X^0, X^1 \mapsto X^+, X^-$ равен единице. Обратное преобразование имеет вид

$$X^0 = \frac{1}{\sqrt{2}}(X^+ + X^-), \quad X^1 = \frac{1}{\sqrt{2}}(X^+ - X^-). \quad (28.66)$$

В частности, конусные координаты репера (диады) равны

$$e_\alpha^\pm = \frac{1}{\sqrt{2}}(e_\alpha^0 \pm e_\alpha^1). \quad (28.67)$$

Нетрудно проверить, что определитель репера можно записать в виде

$$\sqrt{|g|} = \det e_\alpha^a = e_0^- e_1^+ - e_0^+ e_1^- = \hat{\varepsilon}^{\alpha\beta} e_\alpha^+ e_\beta^-, \quad (28.68)$$

что эквивалентно тождеству

$$\varepsilon^{\alpha\beta} e_\alpha^+ e_\beta^- = 1. \quad (28.69)$$

При проведении расчетов часто приходится обращаться диаду. Выпишем соответствующие формулы:

$$\begin{aligned} e^0_+ &= -\frac{1}{\sqrt{|g|}} e^1_-, & e^1_+ &= \frac{1}{\sqrt{|g|}} e^0_- \\ e^0_- &= \frac{1}{\sqrt{|g|}} e^1_+, & e^1_- &= -\frac{1}{\sqrt{|g|}} e^0_+. \end{aligned} \quad (28.70)$$

Скалярное произведение векторов в координатах светового конуса имеет вид

$$X^a Y^b \eta_{ab} = X^- Y^+ + X^+ Y^-. \quad (28.71)$$

Замечание. Важным является то обстоятельство, что лоренцев квадрат вектора,

$$X^2 = 2X^+ X^-, \quad (28.72)$$

линеен по компонентам, а не квадратичен, как в декартовой системе координат. Это помогает в ряде случаев записать явно решение некоторых уравнений, не извлекая квадратного корня. $\square$

Скалярное произведение можно записывать в виде $X^\pm Y^\pm \eta_{\pm\pm}$, где индекс $\pm = \{+, -\}$ пробегает два значения, а лоренцева метрика и ее обратная в координатах светового конуса имеют одинаковый вид,

$$\eta_{\pm\pm} = \eta^{\pm\pm} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (28.73)$$

Ко- и контравариантные компоненты векторов связаны простыми соотношениями:

$$X^+ = X_-, \quad X^- = X_+, \quad (28.74)$$

при этом соотношения между ковариантными компонентами такие же, как и для контравариантных векторов (28.65):

$$X_{\pm} = \frac{1}{\sqrt{2}}(X_0 \pm X_1).$$

Псевдоскаляр, построенный из двух векторов, в координатах светового конуса принимает вид

$$X^a Y^b \varepsilon_{ab} = X^{\pm} Y^{\pm} \varepsilon_{\pm\pm} = X^{-} Y^{+} - X^{+} Y^{-}, \quad (28.75)$$

где

$$\varepsilon_{\pm\pm} = -\varepsilon^{\pm\pm} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad (28.76)$$

$$\varepsilon^{\pm}_{\pm} = -\varepsilon_{\pm}^{\pm} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (28.77)$$

Используя выражение для псевдоследа тензора кручения (28.17), получаем его компоненты в координатах светового конуса

$$T^{*\pm} = 2\varepsilon^{\alpha\beta}(\partial_{\alpha} \mp \omega_{\alpha})e_{\beta}^{\pm}. \quad (28.78)$$

Отсюда следует, что каждая компонента репера $e_{\alpha}^{\pm}$ определяет одну компоненту псевдоследа. Тожество (28.24) в координатах светового конуса принимает вид

$$R = \tilde{R} + 2\nabla_a T^a - 4T^{+}T^{-}, \quad (28.79)$$

где

$$\nabla_{\alpha} T^{*\pm} = \partial_{\alpha} T^{*\pm} \mp \omega_{\alpha} T^{*\pm}.$$

Координаты светового конуса играют большую роль в общем решении уравнений движения двумерной гравитации.

Глава 29

Поверхности постоянной кривизны

Поверхности постоянной кривизны являются относительно простыми объектами, которые хорошо изучены и в некоторых случаях классифицированы. Универсальными накрывающими для любой полной римановой поверхности постоянной кривизны являются либо сфера $\mathbb{S}^2$, либо евклидова плоскость $\mathbb{R}^2$, либо одна компонента связности двуполостного гиперboloида $\mathbb{H}^2$, соответственно, для положительной, нулевой и отрицательной гауссовой кривизны. Все полные римановы поверхности постоянной кривизны получаются из этих трех как фактор пространства по действию свободной и собственно разрывной группы преобразований (теорема 11.2.4).

В случае псевдоримановых поверхностей, метрика которых имеет лоренцеву сигнатуру, можно не различать поверхности постоянной положительной и отрицательной кривизны, т.к. они связаны между собой простым изменением знака метрики: $g_{\alpha\beta} \mapsto -g_{\alpha\beta}$, не меняющим сигнатуру. Для нулевой кривизны, универсальной накрывающей является плоскость Минковского $\mathbb{R}^{1,1}$. Если кривизна отлична от нуля, то универсальной накрывающей поверхности постоянной кривизны является универсальная накрывающая однополостного гиперboloида $\mathbb{L}^2$.

В настоящей главе мы подробно опишем все универсальные накрывающие поверхности постоянной кривизны, за исключением евклидовой плоскости $\mathbb{R}^2$ и двумерного пространства Минковского $\mathbb{R}^{1,1}$, которые достаточно полно были рассмотрены во Введении.

29.1 Сфера $\mathbb{S}^2$

Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовыми координатами x, y, z . Двумерная сфера $\mathbb{S}_a^2$ радиуса $a > 0$ с центром в начале системы координат задается уравнением (см. рис. 29.1, a)

$$x^2 + y^2 + z^2 = a^2. \quad (29.1)$$

Сфера, очевидно, является связным и односвязным многообразием.

Вложение сферы в трехмерное евклидово пространство $\mathbb{S}_a^2 \hookrightarrow \mathbb{R}^3$, заданное уравнением (29.1), индуцирует на ней положительно определенную метрику с помощью возврата отображения (см. раздел 2.9). Явный вид индуцированной метрики проще всего выглядит в сферической системе координат r, θ, φ (см. раздел 7.1). Чтобы получить его, в выражение для евклидовой метрики в $\mathbb{R}^3$, записанной в сферической системе координат,

$$ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)$$

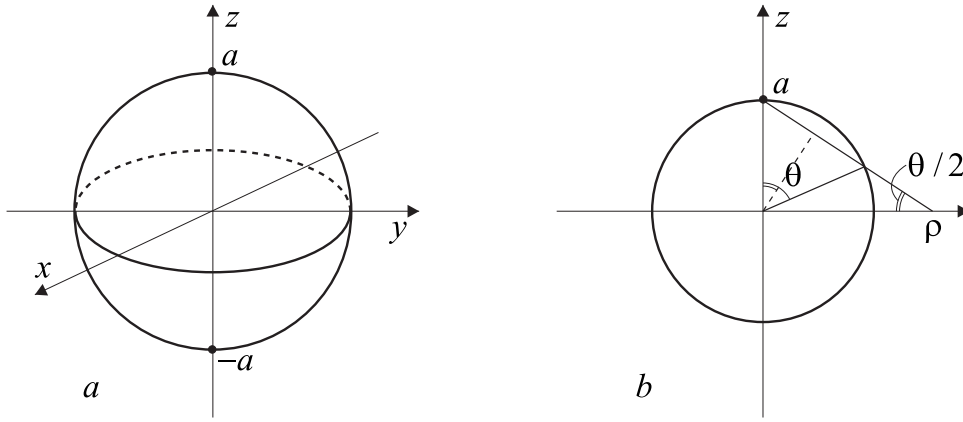


Рис. 29.1: Сфера радиуса a с центром в начале координат (a) . Стереографическая проекция точки сферы на плоскость x, y (b) .

достаточно подставить уравнение сферы $r = a$. В результате метрика, индуцированная на сфере, в сферической системе координат примет вид

$$d\Omega := a^2(d\theta^2 + \sin^2\theta d\varphi^2), \quad (29.2)$$

где координаты на сфере θ и φ называются, соответственно, *азимутальным* и *полярным* углами. Координаты θ, φ покрывают всю сферу, но не определяют глобальную систему координат на сфере, т.к. необходимо провести отождествление, например, $\varphi \sim \varphi + 2\pi$. Соответствующую метрику и ее обратную можно записать в матричных обозначениях:

$$g_{\alpha\beta} = a^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}, \quad g^{\alpha\beta} = \frac{1}{a^2} \begin{pmatrix} 1 & 0 \\ 0 & \sin^{-2}\theta \end{pmatrix}.$$

Из символов Кристоффеля (6.24) отличны от нуля только три:

$$\Gamma_{\theta\varphi}^\varphi = \Gamma_{\varphi\theta}^\varphi = \frac{\cos\theta}{\sin\theta}, \quad \Gamma_{\varphi\varphi}^\theta = -\sin\theta \cos\theta.$$

Тензор кривизны имеет только две нетривиальные линейно независимые компоненты:

$$R_{\theta\varphi\theta}^\varphi = -1, \quad R_{\theta\varphi\varphi}^\theta = \sin^2\theta. \quad (29.3)$$

Тензор кривизны со всеми опущенными индексами имеет только одну независимую нетривиальную компоненту

$$R_{\theta\varphi\theta\varphi} = -a^2 \sin^2\theta.$$

Тензор Риччи диагонален:

$$R_{\theta\theta} = -1, \quad R_{\varphi\varphi} = -\sin^2\theta.$$

Соответствующая скалярная кривизна (6.90) равна константе

$$R = -2K = -\frac{2}{a^2}.$$

Отсюда следует, что согласно нашему определению тензора кривизны, тензора Риччи и скалярной кривизны, скалярная кривизна сферы единичного радиуса равна -2 .

Поэтому мы также используем *гауссову кривизну* K , которая отличается от скалярной кривизны множителем -2 . Для сферы единичного радиуса гауссова кривизна равна единице, $K = 1$.

Экстремали (линии наименьшей длины) на сфере определяются следующей теоремой.

Теорема 29.1.1. *Сечения сферы плоскостями, проходящими через ее центр и только они являются экстремальями. Все экстремали полны.*

Доказательство. См., например, [211], раздел 7. □

Таким образом, все экстремали на сфере являются большими окружностями и замкнуты. Ясно, что они полны, т.к. их можно проходить бесконечное число раз. Любые две точки, которые не являются диаметрально противоположными, можно соединить двумя экстремальями. Одна из них будет кратчайшей, а другая – длиннейшей. Диаметрально противоположные точки соединяются бесконечно большим числом экстремалей, которые имеют одинаковую длину.

Метрика евклидова пространства $\mathbb{R}^3$ инвариантна относительно действия группы трехмерных вращений $\mathbb{SO}(3)$. Кроме того уравнение сферы (29.1) также инвариантно относительно вращений. Отсюда следует, что метрика, индуцированная на сфере, допускает по крайней мере три вектора Киллинга по числу параметров группы Ли преобразований. С другой стороны, поверхность постоянной кривизны допускает существование не более трех линейно независимых векторов Киллинга (теорема 15.3.2). Поэтому сфера представляет собой максимально симметричную поверхность постоянной положительной кривизны. Любой вектор Киллинга на сфере представляет собой линейную комбинацию векторов Киллинга, соответствующих вращениям. В декартовой системе координат $\{x^i\} = (x, y, z)$, $i = 1, 2, 3$, векторы Киллинга евклидова пространства $\mathbb{R}^3$ имеют следующие компоненты:

$$K_i = \varepsilon_{ij}{}^k x^j \partial_k. \quad (29.4)$$

где ε_{ijk} – полностью симметричный тензор третьего ранга. Они удовлетворяют коммутационным соотношениям (1.130) алгебры $\mathfrak{so}(3)$:

$$[K_i, K_j] = -\varepsilon_{ij}{}^k K_k. \quad (29.5)$$

Используя формулы перехода от декартовых координат к сферическим (7.1), получаем явный вид векторных полей Киллинга на сфере $\mathbb{S}_a^2$ в сферических координатах:

$$\begin{aligned} K_1 &= -\sin \varphi \partial_\theta - \frac{\cos \theta \cos \varphi}{\sin \theta} \partial_\varphi, \\ K_2 &= \cos \varphi \partial_\theta - \frac{\cos \theta \sin \varphi}{\sin \theta} \partial_\varphi, \\ K_3 &= \partial_\varphi. \end{aligned} \quad (29.6)$$

Можно проверить, что они удовлетворяют той же алгебре (29.5), что и векторы Киллинга евклидова пространства.

В приложениях часто используются *стереографические координаты* ρ, φ на сфере $\mathbb{S}_a^2$, в которых индуцированная метрика (29.2) является вейлевски евклидовой. Они строятся следующим образом. Спроектируем точку (θ, φ) , не находящуюся на северном полюсе, на плоскость x, y из северного полюса, как показано на рис. 29.1, *b*. При

этом точки нижней и верхней полусферы отображаются, соответственно, на внутренность и дополнение к кругу радиуса a на плоскости x, y . Тогда полярный радиус ρ на плоскости x, y связан с азимутальным углом θ следующим соотношением

$$\frac{a}{\rho} = \operatorname{tg} \frac{\theta}{2} = \frac{1 - \cos \theta}{\sin \theta}.$$

Отсюда находим радиус

$$\rho = a \frac{\sin \theta}{1 - \cos \theta}. \quad (29.7)$$

Подставляя соотношение между дифференциалами

$$d\rho = -a \frac{1}{1 - \cos \theta} d\theta$$

в выражение для индуцированной метрики (29.2), получаем выражение для метрики на сфере в стереографических координатах

$$ds^2 = \frac{4}{(1 + \rho^2/a^2)^2} (d\rho^2 + \rho^2 d\varphi^2), \quad 0 \leq \rho < \infty, \quad 0 \leq \varphi < 2\pi. \quad (29.8)$$

Отсюда следует, что в стереографических координатах метрика на сфере является вейлевски евклидовой.

Векторы Киллинга в стереографических координатах имеют вид:

$$\begin{aligned} K_1 &= \frac{\rho^2 + a^2}{2a} \sin \varphi \partial_\rho - \frac{\rho^2 - a^2}{2a\rho} \cos \varphi \partial_\varphi, \\ K_2 &= -\frac{\rho^2 + a^2}{2a} \cos \varphi \partial_\rho - \frac{\rho^2 - a^2}{2a\rho} \sin \varphi \partial_\varphi, \\ K_3 &= \partial_\varphi. \end{aligned} \quad (29.9)$$

Стереографические координаты покрывают все точки сферы, за исключением северного полюса, который отображается в бесконечность на плоскости x, y . Аналогично можно построить проекцию точек сферы из южного полюса. Тогда сфера будет покрыта двумя картами.

Рассмотрим плоскость x, y как комплексную плоскость $z = x + iy \in \mathbb{C}$. В комплексных координатах метрика (29.8) примет вид

$$ds^2 = \frac{4}{(1 + z\bar{z}/a^2)^2} dz d\bar{z}. \quad (29.10)$$

При стереографической проекции северный полюс сферы отображается в бесконечно удаленную точку на комплексной плоскости. Комплексная плоскость, объединенная с бесконечно удаленной точкой, называется *расширенной комплексной плоскостью* и обозначается $\overline{\mathbb{C}}$. Комплексная плоскость $\mathbb{C} \approx \mathbb{R}^2$ и расширенная комплексная плоскость $\overline{\mathbb{C}} \approx \mathbb{S}^2$, как было отмечено в примере 1.2.17, имеют разные топологии. Сфера $\mathbb{S}^2$, на которую взаимно однозначно отображается расширенная комплексная плоскость с помощью стереографической проекции называется *сферой Римана*.

Отображение расширенной комплексной плоскости $\overline{\mathbb{C}}$ на сферу $\mathbb{S}^2$ можно записать в комплексном виде. Обозначим декартовы координаты точки, расположенной на сфере, через ξ, η, t . По определению, они удовлетворяют уравнению

$$\xi^2 + \eta^2 + t^2 = a^2,$$

т.е. в качестве координат на сфере можно выбрать координаты ξ и η или комплексную переменную $\zeta := \xi + i\eta$. Они связаны с комплексными координатами на плоскости $z = x + iy$ соотношением

$$\frac{\zeta}{z} = \frac{a - t}{a}.$$

Исключив из этих соотношений t , получим равенство

$$\zeta = \frac{2z}{1 + z\bar{z}/a^2}.$$

Выпишем также формулы, для координаты t и обратного преобразования $\xi, \eta, t \rightarrow z$

$$t = \frac{z\bar{z} - a^2}{z\bar{z} + a^2}a, \quad z = \frac{a}{a - t}\zeta.$$

Эти формулы часто используются в приложениях.

29.2 Двуполостный гиперboloид $\mathbb{H}^2$

Рассмотрим трехмерное пространство Минковского $\mathbb{R}^{1,2}$ с декартовыми координатами t, x, y . Метрика Лоренца на этом многообразии определяется квадратичной формой

$$ds^2 = dt^2 - dx^2 - dy^2. \quad (29.11)$$

Двуполостный гиперboloид $\mathbb{H}^2$ “радиуса” $a > 0$, вложенный в пространство Минковского $\mathbb{R}^{1,2}$, задается уравнением

$$t^2 - x^2 - y^2 = a^2, \quad a > 0, \quad (29.12)$$

и показан на рис. 29.2, *a*. Это уравнение имеет решения только при $|t| \geq a$. Верхняя и нижняя полы гиперboloида соответствуют значениям $t \geq a$ и $t \leq -a$. Как многообразие двуполостный гиперboloид состоит из двух компонент связности: верхней и нижней полы. Каждая пола является связным и односвязным многообразием. Для определенности выберем верхнюю полу гиперboloида и обозначим ее $\mathbb{H}^2$. В дальнейшем, говоря про двуполостный гиперboloид, будем подразумевать именно его верхнюю полу $\mathbb{H}^2$. Верхняя (или нижняя) пола двуполостного гиперboloида называется также *плоскостью Лобачевского* или *гиперболической плоскостью*.

Вложение $\mathbb{H}^2 \hookrightarrow \mathbb{R}^{1,2}$ двуполостного гиперboloида в пространство Минковского (29.12) индуцирует на нем метрику евклидовой сигнатуры (риманову метрику). Это очевидно, поскольку все касательные векторы к поверхности (29.12) пространственноподобны.

Вычисление геометрических характеристик двуполостного гиперboloида наиболее просто проводится в гиперболических координатах:

$$\begin{aligned} x &= r \operatorname{sh} \theta \cos \varphi, \\ y &= r \operatorname{sh} \theta \sin \varphi, \\ t &= \pm r \operatorname{ch} \theta, \end{aligned} \quad (29.13)$$

которые мы обозначим теми же буквами r, θ, φ , что и сферические координаты. Преобразования координат (29.13) определены при всех значениях координат r, θ, φ . В

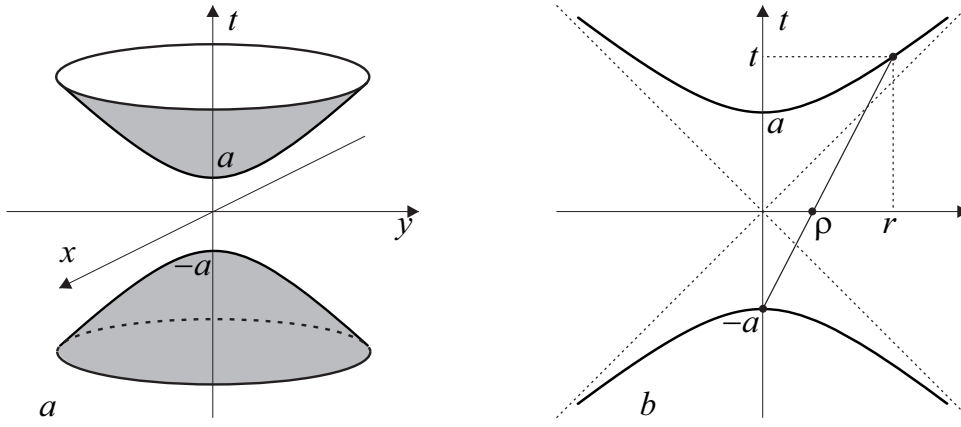


Рис. 29.2: Двуполостный гиперboloид $\mathbb{H}^2$, вложенный в трехмерное пространство Минковского $\mathbb{R}^{1,2}$ (a). Стереографические координаты ρ, φ на двуполостном гиперboloиде (b).

приведенных формулах знак $\pm$ в формуле для t относится к гиперболическим координатам, соответственно, в верхнем $t > 0$ и нижнем $t < 0$ полупространствах.

Якобиан преобразования легко вычислить

$$J = \frac{\partial(t, x, y)}{\partial(r, \theta, \varphi)} = \pm r^2 \operatorname{sh} \theta.$$

Таким образом, преобразование координат вырождено при $r = 0$ и $\theta = 0$, т.е. на оси времени t .

Обратные преобразования запишем в следующем виде

$$\begin{aligned} r &= \sqrt{t^2 - x^2 - y^2}, \\ \theta &= \pm \operatorname{arctg} \frac{\sqrt{x^2 + y^2}}{t}, \\ \varphi &= \operatorname{arctg} \frac{y}{x}. \end{aligned} \quad (29.14)$$

Они определены при $t^2 \geq x^2 + y^2$, и знак $\pm$ относится, соответственно, к верхней и нижней поле гиперboloида. Мы видим, что внутренность светового конуса будущего, $t > 0$, с вершиной в начале координат и удаленной полуплоскостью $y = 0$, $x \geq 0$ отображается на область

$$0 < r < \infty, \quad 0 < \theta < \infty, \quad 0 < \varphi < 2\pi. \quad (29.15)$$

При этом внутренность светового конуса прошлого, $t < 0$, с удаленной полуплоскостью также отображается на область (29.15). Заметим, что полуплоскость необходимо удалить из-за полярного угла $0 < \varphi < 2\pi$.

В гиперболических координатах двуполостный гиперboloид (29.12) задается особенно просто: $r = a = \text{const} > 0$. Поэтому углы θ и φ можно рассматривать в качестве координат на двуполостном гиперboloиде $\mathbb{H}^2$.

Лоренцева метрика (29.11) в гиперболических координатах (29.13) принимает вид

$$ds^2 = dr^2 - r^2(d\theta^2 + \operatorname{sh}^2 \theta d\varphi^2).$$

Поэтому индуцированная метрика на двуполостном гиперboloиде $\mathbb{H}^2$, $r = a = \text{const}$, задается квадратичной формой

$$dl^2 = a^2(d\theta^2 + \text{sh}^2\theta d\varphi^2). \quad (29.16)$$

При этом мы изменили общий знак двумерной индуцированной метрики для того, чтобы она стала положительно определенной. Соответствующая метрика и ее обратная в матричных обозначениях имеют вид

$$g_{\alpha\beta} = a^2 \begin{pmatrix} 1 & 0 \\ 0 & \text{sh}^2\theta \end{pmatrix} \quad g^{\alpha\beta} = a^{-2} \begin{pmatrix} 1 & 0 \\ 0 & \text{sh}^{-2}\theta \end{pmatrix}. \quad (29.17)$$

Замечание. Если бы мы рассматривали вложение (29.12) двуполостного гиперboloида в трехмерное евклидово пространство $\mathbb{R}^3$, то индуцированная метрика имела бы другой вид. Поэтому важно, что рассматривается вложение именно в пространство Минковского. $\square$

Прямые вычисления показывают, что только три символа Кристоффеля (6.24) отличны от нуля:

$$\Gamma_{\theta\varphi}^\varphi = \Gamma_{\varphi\theta}^\varphi = \frac{\text{ch}\theta}{\text{sh}\theta}, \quad \Gamma_{\varphi\varphi}^\theta = -\text{sh}\theta \text{ch}\theta.$$

Полный тензор кривизны имеет две отличные от нуля независимые компоненты:

$$R_{\theta\varphi\theta}^\varphi = 1, \quad R_{\theta\varphi\varphi}^\theta = -\text{sh}^2\theta \quad (29.18)$$

При этом тензор кривизны со всеми опущенными индексами имеет только одну отличную от нуля независимую компоненту

$$R_{\theta\varphi\theta\varphi} = a^2 \text{sh}^2\theta.$$

Соответствующий тензор Риччи диагонален:

$$R_{\theta\theta} = 1, \quad R_{\varphi\varphi} = \text{sh}^2\theta,$$

и скалярная кривизна постоянна

$$\tilde{R} = -2K = \frac{2}{a^2}.$$

Для гиперboloида единичного радиуса, $a = 1$, гауссова кривизна $K = -1$. Таким образом, двуполостный гиперboloид является поверхностью постоянной отрицательной гауссовой кривизны.

Очевидно, что группа Лоренца $\mathbb{O}(1, 2)$, действующая глобально в пространстве Минковского $\mathbb{R}^{1,2}$, является группой симметрии и лоренцевой метрики (29.11), и двуполостного гиперboloида (29.12). Поэтому группой симметрии индуцированной метрики на двуполостном гиперboloиде (29.16) является группа Ли псевдоортогональных вращений (группа Лоренца) $\mathbb{SO}(1, 2)$. Соответствующая алгебра $\mathfrak{so}(1, 2)$ в пространстве Минковского $\mathbb{R}^{1,2}$ задается тремя векторами Киллинга:

$$\begin{aligned} K_0 &= -x\partial_y + y\partial_x = -\partial_\varphi, \\ K_1 &= y\partial_t + t\partial_y = \sin\varphi\partial_\theta + \frac{\text{ch}\theta \cos\varphi}{\text{sh}\theta}\partial_\varphi, \\ K_2 &= -t\partial_x - x\partial_t = -\cos\varphi\partial_\theta + \frac{\text{ch}\theta \sin\varphi}{\text{sh}\theta}\partial_\varphi, \end{aligned} \quad (29.19)$$

которые мы записали в декартовых и гиперболических координатах. Векторы Киллинга в гиперболических координатах не содержат компоненты вдоль ∂_r . Это означает, что они касательны к двуполостному гиперboloиду. Коммутационные соотношения для векторных полей Киллинга имеют вид

$$[K_0, K_1] = K_2, \quad [K_2, K_0] = K_1, \quad [K_1, K_2] = -K_0. \quad (29.20)$$

Отметим, что инверсия времени $t \rightarrow -t$ преобразует $K_{1,2} \rightarrow -K_{1,2}$ и не меняет K_0 . При этом алгебра Ли полей Киллинга остается неизменной.

Введем обозначение $\{K_i\} = \{K_0, K_1, K_2\}$, $i = 0, 1, 2$. Тогда алгебру $\mathfrak{so}(1, 2)$ можно записать в компактном виде:

$$[K_i, K_j] = -\varepsilon_{ij}^{\quad k} K_k,$$

где ε_{ijk} — полностью антисимметричный тензор третьего ранга, $\varepsilon_{012} = 1$, а подъем индексов, в отличие от алгебры вращений $\mathfrak{so}(3)$, осуществляется с помощью обратной метрики Минковского $\eta^{ij} = \text{diag}(+ - -)$.

Поскольку в общем случае поверхность допускает не более трех векторов Киллинга, то двуполостный гиперboloид представляет собой максимально симметричную поверхность.

Замечание. Внимательный читатель, конечно, заметил, что формулы для двуполостного гиперboloида и сферы, рассмотренной в разделе 29.1, очень похожи. Соответствующие выражения отличаются частью знаков и заменой тригонометрических функций от θ на гиперболические. $\square$

Группа $\mathbb{SO}_0(1, 2)$ некомпактна и ее алгебра Ли $\mathfrak{so}(1, 2)$ имеет нетривиальную неабелеву подалгебру, натянутую на векторные поля K_1 и $K_0 + K_2$. Действительно, из коммутационных соотношений (29.20) следует равенство

$$[K_1, K_0 + K_2] = -(K_0 + K_2). \quad (29.21)$$

Соответствующая двумерная неабелева группа Ли подробно рассмотрена в разделе 8.7, где $L_x = -\alpha K_1$, $L_y = K_0 + K_2$. Отметим, что алгебра вращений $\mathfrak{so}(3)$ не имеет неабелевых подгрупп.

Легко найти интегральные кривые полей Киллинга в пространстве Минковского $\mathbb{R}^{1,2}$. Например, интегральными кривыми поля K_2 в плоскости t, x являются гиперболы

$$t^2 - x^2 = \pm c^2, \quad c \in \mathbb{R}.$$

Нетрудно проверить, что двуполостные гиперboloиды, параметризуемые “радиусом” a , являются интегральными подмногообразиями полей Киллинга (29.19).

Нетрудно проверить, что векторы Киллинга линейно зависимы:

$$tK_0 + xK_1 + yK_2 = 0. \quad (29.22)$$

С точки зрения теоремы Фробениуса 2.11.1 векторные поля K_1 и $K_0 + K_2$ линейно независимы и определяют распределение в пространстве Минковского $\mathbb{R}^{1,2}$. Они находятся в инволюции и, следовательно, определяют семейство интегральных подмногообразий, которыми являются двуполостные гиперboloиды.

Двуполостный гиперboloид является римановым многообразием постоянной отрицательной гауссовой кривизны. Выше было показано, что эта поверхность естественным образом вкладывается в трехмерное пространство Минковского $\mathbb{R}^{1,2}$, что

связано с $\mathbb{SO}(1, 2)$ симметрией метрики (29.16). Можно доказать, что не существует $\mathcal{C}^2$ вложения полной поверхности постоянной отрицательной гауссовой кривизны в трехмерное евклидово пространство [212]. В то же время в классе $\mathcal{C}^1$ вложение полной поверхности постоянной отрицательной гауссовой кривизны уже возможно [213].

Теперь обсудим экстремали.

Теорема 29.2.1. *Сечения двуполостного гиперboloида $\mathbb{H}^2$ плоскостями, проходящими через начало координат, и только они являются экстремальными. Все экстремали на двуполостном гиперboloиде полны.*

Доказательство. Уравнения для экстремалей интегрируются в явном виде. $\square$

Из рис. 29.2,а ясно, что все экстремали незамкнуты. Через две произвольные точки на гиперboloиде $\mathbb{H}^2$ и начало координат проходит одна и только одна плоскость. Это означает, что две произвольные точки на гиперboloиде можно соединить единственной экстремалью, и это будет линия наименьшей длины.

При преобразовании Лоренца плоскость переходит в плоскость. Поэтому экстремаль переходит в экстремаль. Это следовало ожидать, поскольку метрика двуполостного гиперboloида (29.16) инвариантна относительно преобразований Лоренца.

Для того, чтобы записать метрику двуполого гиперboloида (29.16) в вейлевски евклидовом виде, введем стереографические координаты ρ, φ . Это делается аналогично введению стереографических координат на сфере. Сначала введем обычные полярные координаты r, φ на плоскости x, y . Выберем точку $(t, r, \varphi) \in \mathbb{H}^2$ на верхней полё гиперboloида и спроектируем ее из вершины нижней полы $x = y = 0, t = -a$ на плоскость x, y . На плоскости x, y ей будет соответствовать точка с координатами ρ, φ (см. рис. 29.2,б). Из рисунка ясно, что справедливо соотношение

$$\frac{r}{\rho} = \frac{t + a}{a}.$$

Подставив сюда время $t = \sqrt{a^2 + r^2}$ из уравнения гиперboloида (29.12), получим связь радиусов:

$$r = \frac{2a^2\rho}{a^2 - \rho^2}.$$

Это приводит к соотношению между дифференциалами

$$dr = 2a^2 \frac{a^2 + \rho^2}{(a^2 - \rho^2)^2} d\rho.$$

Чтобы записать метрику двуполостного гиперboloида в стереографических координатах, запишем ее сначала в полярных координатах (после изменения общего знака):

$$ds^2 = \frac{a^2}{a^2 + r^2} dr^2 + r^2 d\varphi^2. \quad (29.23)$$

Для получения этой формулы надо в евклидовой метрике (29.11) произвести замену координат $x, y \mapsto r, \varphi$ и подставить выражение для времени $t = \sqrt{a^2 + r^2}$. Замена $r \mapsto \rho$ приводит к метрике в стереографических координатах

$$ds^2 = \frac{4}{(1 - \rho^2/a^2)^2} (d\rho^2 + \rho^2 d\varphi^2). \quad (29.24)$$

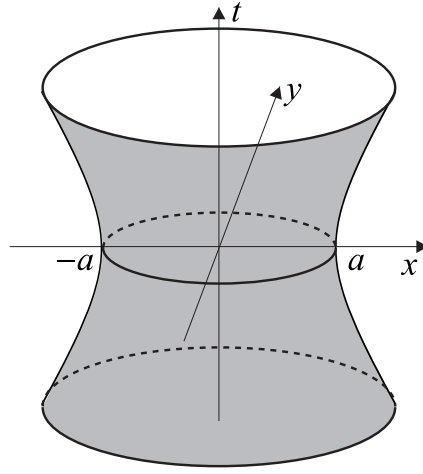


Рис. 29.3: Однополостный гиперboloид $\mathbb{L}^2$, вложенный в трехмерное пространство Минковского $\mathbb{R}^{1,2}$.

Эта метрика является вейлевски евклидовой и определена при

$$0 \leq \rho < a, \quad 0 \leq \varphi < 2\pi.$$

Она отличается от метрики сферы (29.8) только знаком перед ρ^2 в знаменателе, и этим определяется различие в области определения стереографических координат для сферы $\mathbb{S}^2$ и двуполостного гиперboloида $\mathbb{H}^2$.

29.3 Однополостный гиперboloид $\mathbb{L}^2$

Как и раньше, будем считать, что в трехмерном пространстве Минковского $\mathbb{R}^{1,2}$ заданы декартовы координаты t, x, y и метрика Лоренца (29.11). Рассмотрим *однополостный гиперboloид* $\mathbb{L}^2$, вложенный в пространство-время Минковского $\mathbb{R}^{1,2}$. Он задается уравнением (см. рис. 29.3).

$$t^2 - x^2 - y^2 = -a^2, \quad a > 0, \quad (29.25)$$

которое отличается от уравнения для двуполостного гиперboloида (29.12) знаком правой части. Это уравнение имеет решения только при $x^2 + y^2 \geq a^2$.

С топологической точки зрения однополостный гиперboloид является связным, но не односвязным многообразием. Его фундаментальная группа совпадает с фундаментальной группой окружности, $\pi_1(\mathbb{L}^2) = \mathbb{Z}$, а универсальная накрывающая, очевидно, диффеоморфна евклидовой плоскости $\mathbb{R}^2$.

Из рисунка ясно, что в любой точке гиперboloида касательное пространство содержит как времениподобные, так и пространственноподобные векторы по отношению к метрике Лоренца (29.11). Это значит, что вложение $\mathbb{L}^2 \hookrightarrow \mathbb{R}^{1,2}$ индуцирует на поверхности $\mathbb{L}^2$ метрику лоренцевой сигнатуры.

Для проведения вычислений удобно воспользоваться гиперболическими координатами r, θ, φ :

$$\begin{aligned} x &= r \operatorname{ch} \theta \cos \varphi, \\ y &= r \operatorname{ch} \theta \sin \varphi, \\ t &= r \operatorname{sh} \theta. \end{aligned} \quad (29.26)$$

Эти формулы определены при всех значениях r, θ, φ . Якобиан преобразования координат (29.26) равен

$$J = \frac{\partial(t, x, y)}{\partial(r, \theta, \varphi)} = r^2 \operatorname{ch} \theta.$$

Таким образом, преобразование координат вырождено только в начале координат $r = 0$. Поскольку якобиан преобразования координат неотрицателен, то переход к новым координатам сохраняет ориентацию.

Обратные преобразования запишем в виде

$$\begin{aligned} r &= \sqrt{x^2 + y^2 - t^2}, \\ \theta &= \operatorname{arccth} \frac{\sqrt{x^2 + y^2}}{t}, \\ \varphi &= \operatorname{arctg} \frac{y}{x}. \end{aligned} \quad (29.27)$$

Они определены при $x^2 + y^2 \geq t^2$. Мы видим, что внешность светового конуса с вершиной в начале координат и удаленной полуплоскостью $y = 0, x \geq 0$ отображается на область

$$0 < r < \infty, \quad -\infty < \theta < \infty, \quad 0 < \varphi < 2\pi.$$

В гиперболических координатах однополостный гиперboloид определяется уравнением $r = a = \text{const}$. Поэтому углы θ и φ можно выбрать в качестве координат на однополостном гиперboloиде $\mathbb{L}^2$.

Метрика Лоренца (29.11) в гиперболических координатах (29.27) принимает вид

$$ds^2 = -dr^2 + r^2(d\theta^2 - \operatorname{ch}^2 \theta d\varphi^2).$$

Следовательно, индуцированная метрика на однополостном гиперboloиде $r = a = \text{const}$ имеет вид

$$ds^2 = a^2(d\theta^2 - \operatorname{ch}^2 \theta d\varphi^2). \quad (29.28)$$

Она имеет, очевидно, лоренцеву сигнатуру.

В матричных обозначениях метрика и ее обратная имеют вид

$$g_{\alpha\beta} = a^2 \begin{pmatrix} 1 & 0 \\ 0 & -\operatorname{ch}^2 \theta \end{pmatrix} \quad g^{\alpha\beta} = a^{-2} \begin{pmatrix} 1 & 0 \\ 0 & -\operatorname{ch}^{-2} \theta \end{pmatrix}. \quad (29.29)$$

Из символов Кристоффеля только три отличны от нуля:

$$\Gamma_{\theta\varphi}^\varphi = \Gamma_{\varphi\theta}^\varphi = \frac{\operatorname{sh} \theta}{\operatorname{ch} \theta}, \quad \Gamma_{\varphi\varphi}^\theta = \operatorname{sh} \theta \operatorname{ch} \theta.$$

У полного тензора кривизны только две линейно независимые компоненты отличны от нуля:

$$R_{\theta\varphi\theta}^\varphi = 1, \quad R_{\theta\varphi\varphi}^\theta = \operatorname{ch}^2 \theta.$$

Тензор кривизны со всеми опущенными индексами имеет только одну линейно независимую компоненту

$$R_{\theta\varphi\theta\varphi} = -a^2 \operatorname{ch}^2 \theta.$$

Тензор Риччи диагонален:

$$R_{\theta\theta} = 1, \quad R_{\varphi\varphi} = -\operatorname{ch}^2 \theta.$$

Наконец, скалярная кривизна равна

$$R = -2K = \frac{2}{a^2}.$$

Таким образом, однополостный гиперboloид $\mathbb{L}^2$ является поверхностью постоянной отрицательной гауссовой кривизны, $K = -1/a^2$. Чтобы получить поверхность постоянной положительной гауссовой кривизны с метрикой лоренцевой сигнатуры, можно рассмотреть тот же однополостный гиперboloид, но вложенный в трехмерное пространство с метрикой

$$ds^2 = -dt^2 + dx^2 + dy^2,$$

что приведет к изменению знака метрики на поверхности, а, значит, и знака скалярной кривизны.

Метрика трехмерного пространства Минковского (29.11) и уравнение гиперboloида (29.25) инвариантны относительно трехмерной группы лоренцевых вращений $\mathbb{SO}(1, 2)$. Это означает, что преобразования из этой группы являются изометриями. Группа Лоренца $\mathbb{SO}(1, 2)$ трехмерна, и ее действие в пространстве Минковского $\mathbb{R}^{1,2}$ задается тремя векторами Киллинга, которые в декартовых координатах имеют тот же вид, что и в случае двуполостного гиперboloида (29.19):

$$\begin{aligned} K_0 &= -x\partial_y + y\partial_x, \\ K_1 &= y\partial_t + t\partial_y, \\ K_2 &= -t\partial_x - x\partial_t. \end{aligned} \tag{29.30}$$

Конечно, они удовлетворяют той же алгебре Ли $\mathfrak{so}(1, 2)$ (29.20), поскольку пространство Минковского $\mathbb{R}^{1,2}$ не изменилось. Изменится только их вид в гиперболических координатах:

$$\begin{aligned} K_0 &= -\partial_\varphi, \\ K_1 &= \sin \varphi \partial_\theta + \frac{\operatorname{sh} \theta \cos \varphi}{\operatorname{ch} \theta} \partial_\varphi, \\ K_2 &= -\cos \varphi \partial_\theta + \frac{\operatorname{sh} \theta \sin \varphi}{\operatorname{ch} \theta} \partial_\varphi. \end{aligned} \tag{29.31}$$

Поскольку максимальное число полей Киллинга для поверхности равно трем, то для однополостного гиперboloида векторы Киллинга (29.31) определяют максимальный набор изометрий.

Как и в предыдущем разделе, векторы Киллинга (29.31) на однополостном гиперboloиде линейно зависимы (29.22). Векторы K_1 и $K_0 + K_2$ образуют подалгебру в $\mathfrak{so}(1, 2)$ и определяют двумерное инволютивное распределение в пространстве Минковского $\mathbb{R}^{1,2}$. Согласно теореме Фробениуса 2.11.1 это распределение определяет семейство интегральных подмногообразий, которыми являются однополостные гиперboloиды.

Теперь исследуем экстремали (16.21) на однополостном гиперboloиде $\mathbb{L}^2$.

Теорема 29.3.1. *Сечения однополостного гиперboloида $\mathbb{L}^2$ плоскостями, проходящими через начало координат, и только они являются экстремалими. Все экстремали на однополостном гиперboloиде полны.*

Доказательство. См., например, [214]. □

Обозначим угол между секущей плоскостью, проходящей через начало координат, и плоскостью x, y через ψ . В зависимости от величины этого угла все экстремали делятся на три класса:

$$\begin{array}{ll} \psi \in [0, \pi/4) & - \text{пространственноподобные,} \\ \psi = \pi/4 & - \text{светоподобные,} \\ \psi \in (\pi/4, \pi/2] & - \text{времениподобные.} \end{array}$$

Пространственноподобные экстремали замкнуты, а остальные – нет.

Через любые две различные точки на гиперboloиде $\mathbb{L}^2$ и начало координат можно провести одну и только одну плоскость. Это значит, что две произвольные точки однополостного гиперboloида можно соединить по крайней мере одной экстремалью. Если экстремали, соединяющие две точки пространственноподобны, то таких экстремалей две, причем по крайней мере одна из них имеет минимальную длину. В остальных случаях экстремаль, соединяющая две точки, единственна.

При преобразовании Лоренца плоскость переходит в плоскость. Следовательно, экстремаль отображается в экстремаль.

Вычислим длину пространственноподобной замкнутой экстремали. Рассмотрим плоскость

$$t = kx, \quad 0 < k < 1,$$

которая параллельна оси y и определяет пространственноподобную экстремаль. В гиперболических координатах это уравнение принимает вид

$$\operatorname{th} \theta = k \cos \varphi.$$

Отсюда следует равенство для дифференциалов

$$d\theta = -\frac{k \sin \varphi}{1 - k^2 \cos^2 \varphi} d\varphi.$$

Подставив это выражение в метрику (29.28), получим квадрат элемента длины пространственноподобной экстремали

$$ds^2 = -a^2 \frac{1 - k^2}{(1 - k^2 \cos^2 \theta)^2} d\varphi^2.$$

Чтобы вычислить длину экстремали, надо извлечь квадрат и проинтегрировать:

$$L := a\sqrt{1 - k^2} \int_0^{2\pi} \frac{d\varphi}{1 - k^2 \cos^2 \varphi} = 2\pi a.$$

Поскольку задача симметрична относительно вращений вокруг оси t , то отсюда следует, что длина произвольной пространственноподобной экстремали не зависит от угла, под которым плоскость пересекает двуполостный гиперboloид, и равна длине окружности, по которой гиперboloид пересекает плоскость x, y .

Теперь обсудим *стереографические координаты*, для однополостного гиперboloида, в которых она является конформно плоской. Их введение в данном случае не так наглядно, как для сферы, и сложнее. Определим стереографические координаты τ, σ для гиперboloида, как координаты точки на плоскости t, x , которая является центральной проекцией точки гиперboloида из точки $t = x = 0, y = -a$.

Чтобы явно построить функции перехода требуется некоторое геометрическое построение. Пусть точка на гиперboloиде имеет координаты t, x, y . По определению,

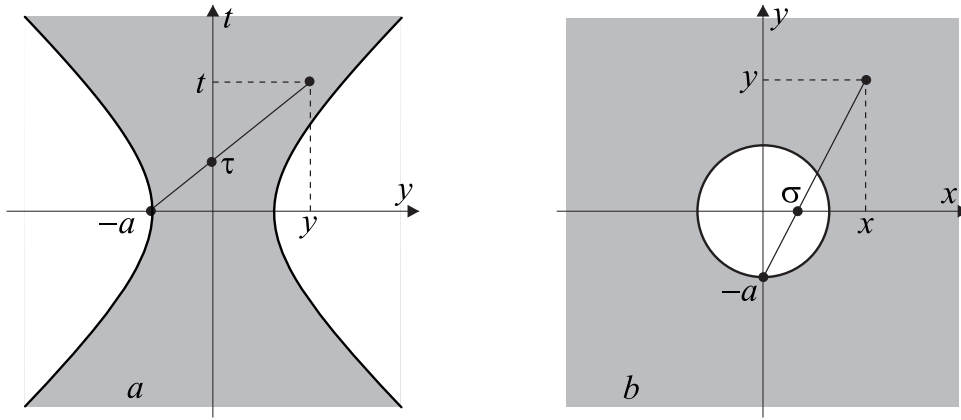


Рис. 29.4: Стереографические координаты τ, σ на однополостном гиперboloиде. Проекция на плоскости t, y (a) и x, y (b).

они удовлетворяют равенству (29.25), т.е. в качестве координат на гиперboloиде можно выбрать, например, декартовы координаты t, x , которые покрывают половину гиперboloида. Соединим эту точку с точкой $t = x = 0, y = -a$ прямой. Соответствующие проекции на плоскости t, y и x, y показаны на рис.29.4 a и b. Из рисунков следуют равенства:

$$\frac{t}{\tau} = \frac{y + a}{a}, \quad \frac{x}{\sigma} = \frac{y + a}{a}. \quad (29.32)$$

Из этих уравнений выразим t и x и подставим в уравнение гиперboloида (29.25). В результате возникнет квадратное уравнение на y , которое можно решить:

$$y = a \frac{a^2 + \tau^2 - \sigma^2}{a^2 - \tau^2 + \sigma^2},$$

где мы выбрали больший корень. Тогда

$$\frac{y + a}{a} = \frac{2a^2}{a^2 - \tau^2 + \sigma^2}.$$

Теперь уравнения (29.32) дают функции перехода:

$$t = \tau \frac{2a^2}{a^2 - \tau^2 + \sigma^2}, \quad x = \sigma \frac{2a^2}{a^2 - \tau^2 + \sigma^2}. \quad (29.33)$$

Якобиан преобразования координат $\tau, \sigma \mapsto t, x$ легко вычислить

$$J = \frac{\partial(t, x)}{\partial(\tau, \sigma)} = 4a^4 \frac{a^2 + \tau^2 - \sigma^2}{(a^2 - \tau^2 + \sigma^2)^3}.$$

Он равен нулю и обращается в бесконечность, соответственно, на гиперболах $\tau^2 - \sigma^2 = -a^2$ и $\tau^2 - \sigma^2 = a^2$ и имеет разный знак по разные стороны гипербол. Таким образом, стереографические координаты пробегает всю плоскость $-\infty < \tau, \sigma < \infty$ за исключением четырех ветвей гипербол.

На гиперболах $\tau^2 - \sigma^2 = -a^2$ из формул преобразования координат (29.33) следуют равенства $t = \tau$ и $x = \sigma$, и, следовательно, $t^2 - x^2 = -a^2$. То есть гипербола $\tau^2 - \sigma^2 = -a^2$ соответствует сечениям гиперboloида плоскостью $y = 0$.

Отображение однополостного гиперboloида на плоскость τ, σ показано на рис.29.5. Для наглядности изображен гиперboloид конечной высоты. Мы предполагаем, что

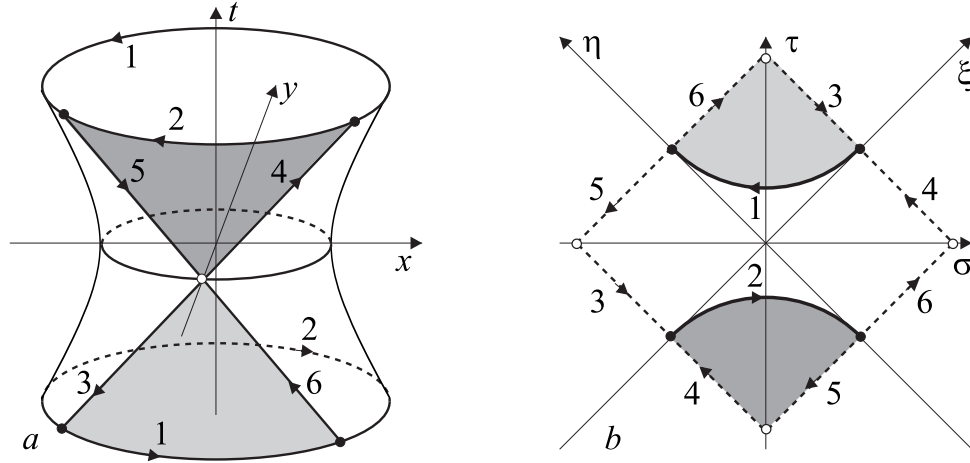


Рис. 29.5: Стереографические координаты на однополостном гиперboloиде. Центральная проекция гиперboloида с координатами t, x (a). Стереографические координаты τ, σ (b).

его верхняя и нижняя граничные окружности бесконечно удалены. Также для наглядности мы сжали плоскость τ, σ вдоль светоподобных направлений

$$\xi = \tau + \sigma, \quad \eta = \tau - \sigma, \quad (29.34)$$

и изобразили ее в виде квадрата. Этого можно добиться с помощью невырожденной замены координат, например,

$$\xi \mapsto \arctg \xi, \quad \eta \mapsto \arctg \eta.$$

Мы также приняли следующие обозначения. Закрашенные черным кружки обозначают точки, находящиеся на бесконечном расстоянии, а полые – на конечном. Точке $t = x = 0, y = -a$, из которой ведется проектирование, соответствуют четыре незакрашенные вершины квадрата. Кривые между выделенными точками на гиперboloиде и их образы обозначены одинаковыми цифрами. Стрелки показывают выбранную ориентацию кривых.

При отображении на плоскость τ, σ весь гиперboloид разбивается на три области, ограниченные кривыми (1,5,3,2,6,4), (2,5,4) и (1,6,3). То есть гиперboloид разрезается вдоль двух прямых $t = \pm x, y = -a$. Центральной точке $t = x = 0, y = -a$ при стереографической проекции соответствуют четыре точки – вершины квадрата – и их необходимо отождествить. Проще всего проследить отображение областей, рассмотрев отображение их границ. Область, ограниченная кривыми (1,5,3,2,6,4) отображается на среднюю часть плоскости τ, σ между гиперболами $\tau^2 - \sigma^2 = a^2$. Области, ограниченные кривыми (2,5,4) и (1,6,3), отображаются, соответственно, в области на плоскости τ, σ , расположенные ниже гиперболы $\tau = -\sqrt{a^2 + \sigma^2}$ и выше гиперболы $\tau = \sqrt{a^2 + \sigma^2}$.

Каждая точка гиперboloида из областей (1,5,3,2,6,4), (2,5,4) и (1,6,3) имеет единственные стереографические координаты. При этом две светоподобные прямые $t^2 = x^2, y = -a$ (линии 3,4 и 5,6 на рис. 29.5) отображаются в бесконечность на плоскости τ, σ . Точки прямых $t^2 = x^2, y = -a$ лежат на конечном расстоянии, и поэтому стороны квадрата на плоскости τ, σ нарисованы пунктиром. Напротив, бесконечно удаленная граница гиперboloида отображается на гиперболы $\tau^2 - \sigma^2 = a^2$, и поэтому нарисованы жирной сплошной линией.

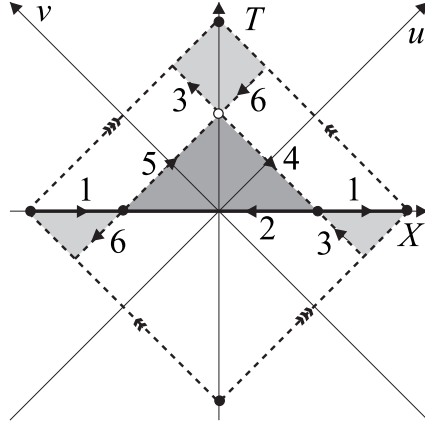


Рис. 29.6: Координаты на однополостном гиперboloиде, в которых метрика зависит только от времени.

Противоположные стороны квадрата на τ, σ плоскости лежат на конечном расстоянии и отображаются в одну прямую на гиперboloиде. Это означает, что их необходимо отождествить (за исключением бесконечно удаленных точек $\xi = 0$, $\eta = \pm\infty$ и $\xi = \pm\infty$, $\eta = 0$).

Чтобы получить метрику однополостного гиперboloида в стереографических координатах, сначала запишем ее в координатах t, x . Это легко сделать путем исключения координаты y из метрики Лоренца (29.11) с помощью уравнения (29.25). В результате метрика примет вид

$$ds^2 = \frac{(a^2 - x^2)dt^2 + 2txdt dx - (a^2 + t^2)dx^2}{a^2 + t^2 - x^2}. \quad (29.35)$$

Подставив сюда выражения для дифференциалов из формул преобразования координат (29.33), мы получим, что в стереографических координатах метрика однополостного гиперboloида является вейлевски лоренцевой:

$$ds^2 = \frac{4}{\left(1 - \frac{\tau^2 - \sigma^2}{a^2}\right)^2} (d\tau^2 - d\sigma^2). \quad (29.36)$$

В светоподобных координатах (29.34) метрика примет вид

$$ds^2 = \frac{4}{\left(1 - \xi\eta/a^2\right)^2} d\xi d\eta. \quad (29.37)$$

Произведем еще одну замену координат так, чтобы интервал зависел только от времениподобной координаты. В этой системе координат двумерное пространство-время является однородным и анализ экстремалей значительно упрощается. Сначала с помощью дробно линейного преобразования преобразуем светоподобные координаты $\xi, \eta \mapsto u, v$:

$$\xi := a \frac{u + a}{u - a}, \quad \eta := a \frac{v + a}{v - a}. \quad (29.38)$$

На рис. 29.6 показан однополостный гиперboloид в новых координатах. Цифрами обозначены образы соответствующих кривых на рис. 29.3. Ось абсцисс нарисована жирной линией, что означает ее полноту (бесконечную удаленность). Верхняя и нижняя вершины квадрата также лежат на бесконечном расстоянии и показаны закрашенными кружками. Противоположные стороны квадрата (за исключением вершин)

лежат на конечном расстоянии и показаны пунктиром. Их необходимо отождествить, т.к. они соответствуют одним и тем же точкам гиперboloида. Все четыре вершины квадрата также отождествляются. Центральная точка стереографической проекции $t = x = 0$, $y = -a$ (четыре вершины квадрата на рис. 29.3, *b*) отображается в точку $u = v = a$. Следуя стрелкам, легко проследить отображение областей $(1,5,3,2,6,4)$, $(2,5,4)$ и $(1,6,3)$.

Из формул преобразования координат (29.38) вытекают формулы преобразования дифференциалов:

$$\begin{aligned} d\xi &= -2a^2 \frac{du}{(u-a)^2}, \\ d\eta &= -2a^2 \frac{dv}{(v-a)^2}. \end{aligned}$$

Якобиан преобразования координат сингулярен на линиях $u = a$ и $v = a$:

$$\frac{\partial(\xi, \eta)}{\partial(u, v)} = \frac{4a^2}{(u-a)^2(v-a)^2}.$$

В новых светоподобных координатах интервал принимает вид

$$ds^2 = \frac{4a^2}{(u+v)^2} du dv. \quad (29.39)$$

Теперь снова введем временную T и пространственную X координаты с помощью соотношений:

$$u := T + X, \quad v := T - X.$$

Тогда метрика однополостного гиперboloида примет вид

$$ds^2 = \frac{a^2}{T^2} (dT^2 - dX^2). \quad (29.40)$$

Она является вейлевски лоренцевой, а соответствующее пространство – однородным.

При решении уравнений движения в различных моделях гравитации мы получаем для метрики некоторое выражение, например, формулу (29.36) или (29.40). То есть мы имеем плоскость τ, σ или T, X с заданной на ней метрикой. Восстановить по этой метрике глобальную структуру пространства-времени (однополостный гиперboloид) довольно сложно, но можно. В разделе 30 будет описан общий подход к этой проблеме, основанный на анализе экстремалей. Если окажется, что экстремали для найденного решения уравнений движения неполны, то решение необходимо продолжить и восстановить глобальную структуру. Это можно сделать с помощью метода конформных блоков. Процедура восстановления поверхности по виду метрики, допускающей один вектор Киллинга, конструктивна, и в результате возникает диаграмма Картера–Пенроуза, представляющая глобальную структуру пространства-времени.

29.4 Уравнение Лиувилля

Рассмотрим евклидову плоскость $\mathbb{R}^2$ с декартовыми координатами x, y или, что эквивалентно, комплексную плоскость $\mathbb{C}$ с координатами $z := x + iy$ и $\bar{z} := x - iy$.

Пусть задана вещественнозначная достаточно гладкая функция $\phi(z, \bar{z})$. Тогда уравнением Лиувилля для функции ϕ называется нелинейное уравнение второго порядка в частных производных

$$\Delta\phi + K e^{2\phi} = 0, \quad K = \text{const}, \quad (29.41)$$

где $\Delta := 4\partial_{z\bar{z}}^2$ – оператор Лапласа.

Уравнение Лиувилля часто встречается в моделях математической физики и является одним из немногих нелинейных явно интегрируемых уравнений.

Теорема 29.4.1. *Общее решение уравнения Лиувилля в односвязной области имеет вид*

$$e^{2\phi} = \frac{4F'\bar{F}'}{(aF\bar{F} + bF + \bar{b}\bar{F} + d)^2}, \quad (29.42)$$

где $F = F(z)$ – произвольная голоморфная функция, введены обозначения:

$$F' := \partial_z F, \quad \bar{F} := \bar{F}(\bar{z}), \quad \bar{F}' := \partial_{\bar{z}} \bar{F},$$

и произвольные постоянные $a, d \in \mathbb{R}$ и $b \in \mathbb{C}$ удовлетворяют одному условию

$$ad - b\bar{b} = K. \quad (29.43)$$

При этом, если $K = 0$, то по крайней мере одна из постоянных a, b , или d должна быть отлична от нуля. Мы также предполагаем, что $F' \neq 0$.

Доказательство. См. [215]. □

Таким образом, произвольное решение уравнения Лиувилля параметризуется одной голоморфной функцией $F(z)$, двумя вещественными (a, d) и одной комплексной (b) постоянными, на которые наложено одно условие (29.42).

Если задана поверхность $\mathbb{M}$ с римановой метрикой $g \in \mathcal{C}^3(\mathbb{M})$, то согласно теореме 28.3.1 локально существует изотермическая система координат, в которой метрика является вейлевски евклидовой:

$$ds^2 = e^{2\phi} dz d\bar{z}.$$

Если поверхность имеет постоянную кривизну, то ее гауссова кривизна также постоянна, $K = \text{const}$. Это условие в изотермических координатах сводится к уравнению Лиувилля (29.41) (см. раздел 28.3.3). Отсюда вытекает следующая

Теорема 29.4.2. *Любая поверхность с римановой метрикой $g \in \mathcal{C}^3(\mathbb{M})$ постоянной кривизны локально изометрична сфере $\mathbb{S}^2$ ($K > 0$), евклидовой плоскости $\mathbb{R}^2$ ($K = 0$) или двуполостному гиперболоиду $\mathbb{H}^2$ ($K < 0$).*

Доказательство. При умножении метрики на положительную постоянную k , $g_{\alpha\beta} \mapsto k g_{\alpha\beta}$, гауссова кривизна делится на ту же постоянную, $K \mapsto K/k$. Поэтому достаточно рассмотреть три случая $K = 1, 0, -1$. Выберем произвольную точку на поверхности. Тогда в некоторой окрестности этой точки существуют изотермические координаты. Следовательно, для конформного множителя возникнет уравнение Лиувилля. Рассмотрим три случая.

$K = 1$. Положим в общем решении уравнения Лиувилля

$$d = 1, \quad a = 1, \quad b = 0, \quad F = z.$$

Тогда решение (29.42) примет вид

$$ds^2 = \frac{4}{(1 + z\bar{z})^2} dz d\bar{z}.$$

Это – метрика сферы единичного радиуса (29.10).

$K = 0$. Положим

$$d = 1, \quad a = 0, \quad b = 0, \quad F = z/2.$$

Тогда получим евклидову метрику $ds^2 = dz d\bar{z}$.

$K = -1$. Положим

$$d = 1, \quad a = -1, \quad b = 0, \quad F = z.$$

Тогда метрика примет вид

$$ds^2 = \frac{4}{(1 - z\bar{z})^2} dz d\bar{z}.$$

Это – метрика (29.24) двуполостного гиперболоида единичного “радиуса” в комплексных координатах. $\square$

Следствие. Риманова метрика поверхности постоянной кривизны бесконечно дифференцируема. $\square$

Доказательство. Поскольку функция $F(z)$ голоморфна, то функция (29.42) бесконечно дифференцируема. $\square$

Доказанная теорема локальна. Согласно теореме 11.3.3 каждая поверхность имеет единственную с точностью до диффеоморфизма универсальную накрывающую поверхность. Поскольку поверхности $\mathbb{S}^2$, $\mathbb{R}^2$ и $\mathbb{H}^2$ связны и односвязны, то они являются универсальными накрывающими поверхностями для всех поверхностей, соответственно, положительной, нулевой и отрицательной кривизны. Кроме того, они полны, т.е. любую экстремаль можно продолжить до бесконечного значения канонического параметра в обе стороны. Чтобы найти все полные поверхности постоянной кривизны, для поверхностей $\mathbb{S}^2$, $\mathbb{R}^2$ и $\mathbb{H}^2$ необходимо найти все группы преобразований $\mathbb{G}$, действующие собственно разрывно и свободно. Тогда все полные поверхности постоянной кривизны будут являться фактор пространствами $\tilde{\mathbb{M}}/\mathbb{G}$, где $\tilde{\mathbb{M}} = \mathbb{S}^2, \mathbb{R}^2, \mathbb{H}^2$ (см. главу 11).

Теперь рассмотрим уравнение Лиувилля на плоскости Минковского $\mathbb{R}^{1,1}$. Обозначим декартовы координаты через t, x . Введем также конусные координаты

$$u := t + x, \quad v := t - x.$$

Тогда аналогом уравнения Лиувилля (29.41) будет являться нелинейное уравнение в частных производных

$$\square\phi + Ke^{2\phi} = 0, \quad K = \text{const.} \quad (29.44)$$

где $\square := 4\partial_{uv}^2$ – оператор Даламбера на плоскости Минковского (1.123). Это уравнение также будем называть *уравнением Лиувилля*.

Уравнение Лиувилля на плоскости Минковского также явно интегрируется. Его общее решение имеет вид

$$e^{2\phi} = \frac{4F'G'}{(aFG + bF + cG + d)^2}, \quad (29.45)$$

где $F = F(u)$ и $G = G(v)$ – произвольные дифференцируемые монотонные функции от светоподобных координат u и v , соответственно, $F'G' \neq 0$. Четыре вещественные постоянные a, b, c и d , должны удовлетворять одному условию

$$ad - bc = K. \quad (29.46)$$

Если $K = 0$, то хотя бы одна из постоянных a, b, c или d должна быть отлична от нуля.

Рассмотрим псевдориманову поверхность. Пусть метрика дважды непрерывно дифференцируема. Тогда согласно теореме 28.3.2 в некоторой окрестности произвольной точки поверхности существует система координат, в которой метрика является вейлевски лоренцевой (28.50). Если поверхность имеет постоянную кривизну, то для конформного множителя возникает уравнение Лиувилля (28.64).

Теорема 29.4.3. *Любая псевдориманова поверхность с метрикой $g \in \mathcal{C}^2(\mathbb{M})$ постоянной кривизны локально изометрична однополостному гиперboloиду $\mathbb{L}^2$ ($K \neq 0$) или плоскости Минковского $\mathbb{R}^{1,1}$ ($K = 0$).*

Доказательство. Для лоренцевых поверхностей изменение знака метрики меняет знак гауссовой кривизны. С другой стороны, это эквивалентно перестановке местами времени и пространства. Поэтому можно не различать псевдоримановы поверхности положительной и отрицательной кривизны. Так же, как и для римановых поверхностей метрику можно умножить на положительную постоянную. Следовательно, достаточно рассмотреть два случая $K = 1$ и $K = 0$.

$K = 1$. Положим в решении уравнения Лиувилля

$$d = 1, \quad a = 1, \quad b = c = 0, \quad F = u, \quad G = v.$$

Тогда метрика примет вид

$$ds^2 = \frac{4}{(1 + uv)^2} dudv.$$

Это – метрика однополостного гиперboloида (29.37) единичного “радиуса” (с точностью до замены общего знака метрики).

$K = 0$. Положим

$$d = 1, \quad a = 0, \quad b = c = 0, \quad F = u/2, \quad G = v/2.$$

Тогда получим метрику Лоренца $ds^2 = dudv$. □

Доказанная теорема локальна. Как и в случае римановых поверхностей любая полная псевдориманова поверхность является фактор пространством $\tilde{\mathbb{M}}/\mathbb{G}$, где $\tilde{\mathbb{M}} = \mathbb{L}^2, \mathbb{R}^{1,1}$ для поверхностей, соответственно, ненулевой и нулевой гауссовой кривизны, а $\mathbb{G}$ – группа преобразований, действующая собственнo разрывно и свободно.

Произвольные функции $F(u)$ и $G(v)$, входящие в решение (29.45) уравнения Лиувилля, соответствуют конформным преобразованиям светоподобных координат (28.56), в которых метрика является вейлевски лоренцевой. Поскольку $F' \neq 0$ и $G' \neq 0$, то всегда можно совершить конформное преобразование координат $u, v \mapsto F, G$, в которых метрика примет вид

$$ds^2 = e^{2\phi} dudv = \frac{4dF dG}{(aFG + bF + cG + d)^2}. \quad (29.47)$$

Область определения решения уравнения Лиувилля (29.45) зависит от выбора функций F и G , а также от значения постоянных a , b , c и d .

При $K > 0$, положим в решении (29.45)

$$F = u, \quad G = v, \quad a = d = 0, \quad b = -c = \sqrt{K}.$$

Тогда метрика примет вид

$$ds^2 = \frac{4dudv}{K(u-v)^2} = \frac{1}{K} \frac{dt^2 - dx^2}{x^2}. \quad (29.48)$$

Эта метрика определена либо в правой, $x > 0$, либо в левой, $x < 0$, полуплоскости.

Аналогично, при $K < 0$ положим

$$F = u, \quad G = v, \quad a = d = 0, \quad b = c = \sqrt{|K|}.$$

Тогда метрика примет вид

$$ds^2 = \frac{4dudv}{|K|(u+v)^2} = \frac{1}{|K|} \frac{dt^2 - dx^2}{t^2}. \quad (29.49)$$

В таком виде метрика определена в верхней, $t > 0$, или нижней, $t < 0$, полуплоскости. Она уже была получена ранее (29.40).

Два последних выражения для метрики связаны между собой перестановкой координат $t \leftrightarrow x$ и изменением знака метрики $g_{\alpha\beta} \rightarrow -g_{\alpha\beta}$. Это означает, что поверхности постоянной положительной и отрицательной кривизны, как уже отмечалось, в случае псевдоримановой метрики эквивалентны.

Замечание. Модели математической физики, как правило, помимо метрики включают также другие поля. Поэтому может случиться так, что суммарное действие для всех полей не будет инвариантно относительно изменения знака метрики. В этом случае знак гауссовой кривизны является существенным. $\square$

При любом значении гауссовой кривизны K мы можем положить

$$F = \frac{u}{2}, \quad G = \frac{v}{2}, \quad a = K, \quad d = 1, \quad b = c = 0.$$

Тогда метрика примет вид

$$ds^2 = \frac{dudv}{(1 + \frac{K}{4}uv)^2}.$$

Эта метрика определена на всей плоскости за исключением двух ветвей гиперболы $uv = -4/K$.

Положим $F = u$ и $G = v$ в общем решении уравнения Лиувилля (29.45). Тогда метрика примет вид

$$ds^2 = \frac{4dudv}{(auv + bu + cv + d)^2}. \quad (29.50)$$

Предложение 29.4.1. При дробно линейном преобразовании светоподобных координат

$$u \mapsto u' := \frac{\alpha_1 u + \beta_1}{\gamma_1 u + \delta_1}, \quad v \mapsto v' := \frac{\alpha_2 v + \beta_2}{\gamma_2 v + \delta_2}, \quad (29.51)$$

где $\alpha_1, \dots, \delta_1, \alpha_2, \dots, \delta_2$ – некоторые постоянные, такие что $\det_1 := \alpha_1 \delta_1 - \beta_1 \gamma_1 \neq 0$ и $\det_2 := \alpha_2 \delta_2 - \beta_2 \gamma_2 \neq 0$, метрика (29.50) принимает вид

$$ds^2 = \frac{4 \det_1 \det_2 du' dv'}{(a'u'v' + b'u' + c'v' + d')^2}, \quad (29.52)$$

где

$$\begin{aligned} a' &= a\alpha_1\alpha_2 + b\alpha_1\gamma_2 + c\gamma_1\alpha_2 + d\gamma_1\gamma_2, \\ b' &= a\alpha_1\beta_2 + b\alpha_1\delta_2 + c\gamma_1\beta_2 + d\gamma_1\delta_2, \\ c' &= a\beta_1\alpha_2 + b\beta_1\gamma_2 + c\delta_1\alpha_2 + d\delta_1\gamma_2, \\ d' &= a\beta_1\beta_2 + b\beta_1\delta_2 + c\delta_1\beta_2 + d\delta_1\delta_2. \end{aligned}$$

Доказательство. Прямые вычисления. □

Дробно линейные преобразования (29.51) образуют подгруппу группы конформных преобразований координат (28.56). Если $\det_1 = \det_2 = 1$, то соответствующие дробно линейные преобразования образуют подгруппу (модулярная группа), которая сохраняет вид метрики (29.50). Эти преобразования можно использовать для фиксирования некоторых постоянных из a, b, c и d .

Глава 30

Лоренцевы поверхности с одним вектором Киллинга

Перед тем, как приступить к построению глобальных решений в моделях гравитации, дадим определение, что значит “продолжение решения”.

Определение. Рассмотрим два (псевдо)римановых многообразия (M_1, g_1) и (M, g) одинаковой размерности с метриками g_1 и g , соответственно. Пусть задано вложение $M_1 \xrightarrow{f_1} M$. Возврат отображения индуцирует на многообразии M_1 метрику f_1^*g . Если индуцированная метрика совпадает с исходной, $f_1^*g = g_1$, то мы говорим, что (псевдо)риманово многообразие (M, g) является *продолжением* (псевдо)риманова многообразия (M_1, g_1) .

Пусть задано еще одно (псевдо)риманово многообразие (M_2, g_2) той же размерности и вложение $M_2 \xrightarrow{f_2} M$. Пусть также $f_2^*g = g_2$. Если

$$M = \overline{f_1(M_1)} \cup \overline{f_2(M_2)} \quad \text{и} \quad f_1(M_1) \cap f_2(M_2) = \emptyset,$$

где черта обозначает замыкание в топологии M , то мы говорим, что (псевдо)римановы многообразия (M_1, g_1) и (M_2, g_2) *склеены вдоль* $\overline{f_1(M_1)} \cap \overline{f_2(M_2)}$. $\square$

Данное определение продолжения (псевдо)римановых многообразий аналогично определению сужения и продолжения функции, данному в разделе 1.4. Разница заключается лишь в том, что при отображении метрика преобразуется нетривиально.

Ясно, что продолжений (псевдо)римановых многообразий существует бесконечно много. В общей теории относительности метрика ищется как решение уравнений Эйнштейна. Поэтому мы требуем дополнительно, чтобы при продолжении и метрика g_1 на M_1 , и метрика g на M удовлетворяли одним и тем же уравнениям Эйнштейна. В этом случае мы говорим о продолжении решений уравнений Эйнштейна. Допустим, что решения уравнений Эйнштейна единственны в том смысле, что если две метрики g_1 и g_2 , определенные в некоторой области $U \subset M$, удовлетворяют уравнениям Эйнштейна и совпадают в некоторой подобласти $V \subset U$, то они совпадают на всем U . В этом случае продолжение решений уравнений Эйнштейна является единственным. Такое продолжение аналогично аналитическому продолжению функций на комплексной плоскости. Поэтому часто говорят об аналитическом продолжении решений уравнений Эйнштейна. Однако даже аналитическое продолжение функций на неодносвязном двумерном многообразии является неоднозначной операцией.

Поскольку среди уравнений Эйнштейна есть гиперболические, то ожидать однозначности решений в указанном выше смысле не приходится. Насколько известно

автору, в настоящее время единственность *статических* решений уравнений Эйнштейна не доказана даже для односвязных многообразий. В то же время не известно ни одного примера неоднозначности статических решений. Поэтому термин “аналитическое продолжение” следует употреблять с большой осторожностью.

Теперь возникает вопрос о том, до каких пор следует продолжать найденное решение уравнений Эйнштейна? Ответ на данный вопрос однозначен: решение необходимо продолжить настолько, на сколько это возможно. Критерием здесь является полнота геодезических (экстремалей).

Согласно общим теоремам о единственности решений систем обыкновенных дифференциальных уравнений через каждую точку (псевдо)риманова многообразия (M, g) в данном направлении проходит одна и только одна геодезическая (экстремаль). Поэтому при склейке двух многообразий задается канонический способ продолжения геодезических многообразия (M_1, g_1) в геодезические многообразия (M_2, g_2) . Поэтому можно говорить, что решение уравнений Эйнштейна продолжается вдоль геодезических. Конечно, и в этом случае единственность продолжения определяется единственностью решений уравнений Эйнштейна.

Определение. Пара (M, g) , где M – многообразие и g – заданная на нем метрика, называется *глобальным решением* в данной модели гравитации, если метрика удовлетворяет соответствующим уравнениям движения на M , а само многообразие M является максимально продолженным. $\square$

Напомним, что максимальное продолжение означает, что любую экстремаль (геодезическую) можно либо продолжить в обе стороны до бесконечного значения канонического параметра, либо при конечном значении канонического параметра она попадает в сингулярную точку, где один из геометрических инвариантов становится сингулярным. Это определение инвариантно и не зависит от выбора системы отсчета, т.к. канонический параметр инвариантен относительно общих преобразований координат и определен с точностью до линейного преобразования.

Пример 30.0.1. Нетривиальным глобальным решением в общей теории относительности является расширение Крускала–Секереша [216, 217] решения Шварцшильда [218] (независимо найденного также Дросте [219, 220, 221]) (см. главу 34). В этом случае четырехмерное пространство-время представляет собой топологическое произведение сферы S^2 на псевдориманову поверхность с одним вектором Киллинга, которая обычно изображается в виде диаграммы Картера–Пенроуза и будет описана в настоящем разделе. $\square$

В настоящей главе развит конструктивный метод конформных блоков построения глобальных решений (M, g) для двумерных метрик лоренцевой сигнатуры специального вида, обладающих одним вектором Киллинга.

Необходимость построения глобальных решений, т.е. одновременного построения и самих многообразий и заданных на них метрик, связана с инвариантностью моделей гравитации относительно общих преобразований координат. Действительно, для решения уравнений движения, как правило, фиксируется некоторая система координат, что уменьшает число искомых функций. Но, с другой стороны, фиксирование системы отсчета означает, что, за исключением тривиальных случаев, найденное решение является только локальным и представляет собой, как правило, только часть некоторого большего пространства-времени. Отметим также, что только глобальная структура многообразия позволяет дать физическую интерпретацию решений уравнений движения для метрики.

Построение глобальных решений в гравитации является трудной задачей, поскольку, помимо решения уравнений движения, что само по себе сложно, предполагает решение уравнений для экстремалей, анализ их полноты и продолжения многообразия, если оно оказалось неполным. В общей теории относительности из-за сложности уравнений движения известно лишь небольшое число глобальных решений. Некоторые из них будут описаны в главе 33). В двумерных моделях гравитации ситуация проще и удалось построить все глобальные решения [214] в двумерной гравитации с кручением [222], а также в широком классе моделей дилатонной гравитации [223, 224, 225]. Конструктивный метод конформных блоков построения глобальных решений для псевдоримановых поверхностей был предложен в [214] для двумерной гравитации с кручением в конформной калибровке. Для поверхностей с римановой метрикой метод конформных блоков был развит в статье [226].

Замечание. В настоящей главе мы рассматриваем продолжение решений только вдоль экстремалей. В аффинной геометрии с кручением и неметричностью геодезические и экстремали в общем случае различны. Поэтому имеет смысл рассматривать максимальное продолжение многообразий как вдоль экстремалей, так и вдоль геодезических. Если геодезические и экстремали совпадают как, например, в общей теории относительности, то продолжение решения вдоль экстремалей автоматически влечет за собой продолжение вдоль геодезических. $\square$

30.1 Локальный вид лоренцевой метрики

Рассмотрим плоскость $\mathbb{R}^2$ с декартовыми координатами $\{\zeta^\alpha\} = (\tau, \sigma)$, $\alpha = 0, 1$. Пусть в некоторой области на плоскости задана метрика лоренцевой сигнатуры, которая имеет вейлевски плоский вид

$$ds^2 = |\Phi(q)|(d\tau^2 - d\sigma^2). \quad (30.1)$$

Будем считать, что конформный множитель $\Phi(q) \in C^l$, $l \geq 2$, является l раз непрерывно дифференцируемой функцией одного аргумента $q \in \mathbb{R}$ за исключением конечного числа степенных особенностей. Этого достаточно для того, чтобы компоненты метрики (30.1) могли удовлетворять некоторой системе уравнений второго порядка вне особых точек. Пусть аргумент q зависит только от одной из координат, τ или σ , и эта связь задается обыкновенным дифференциальным уравнением

$$\left| \frac{dq}{d\zeta} \right| = \pm \Phi(q), \quad (30.2)$$

где выполняется следующее правило знаков:

$$\begin{aligned} \Phi > 0 : \quad \zeta &= \sigma, & \text{знак } + & \text{(статическое локальное решение),} \\ \Phi < 0 : \quad \zeta &= \tau, & \text{знак } - & \text{(однородное локальное решение).} \end{aligned} \quad (30.3)$$

На первый взгляд двумерная метрика (30.1) имеет очень специфический вид. Однако, это – метрика достаточно общего вида. В дальнейшем будет показано, что некоторые модели гравитации при решении полной системы уравнений движения приводят к двумерной метрике вида (30.1). Например, при отыскании сферически симметричных вакуумных решений уравнений Эйнштейна в четырехмерном пространстве-времени (см. следующую главу), метрика (30.1) возникает на двумерных поверхностях, точки которых параметризуются временем и радиусом. Метрика вида (30.1)

возникает также при решении уравнений движения двумерной гравитации. По этой причине в настоящей главе мы будем, для краткости, называть метрику вида (30.1) просто локальным решением.

Мы допускаем, что конформный множитель метрики (30.1) может обращаться в нуль и иметь особенности в конечном числе точек, которые мы пронумеруем в порядке возрастания: q_i , $i = 1, \dots, k$. В эту последовательность включены также бесконечно удаленные точки $q_1 = -\infty$ и $q_k = \infty$. Будем считать, что вблизи каждой точки q_i конформный множитель ведет себя степенным образом:

$$|q_i| < \infty : \quad \Phi(q) \sim (q - q_i)^m, \quad (30.4)$$

$$|q_i| = \infty : \quad \Phi(q) \sim q^m. \quad (30.5)$$

При конечных значениях q_i показатель степени не равен нулю, $m \neq 0$, т.к. конформный множитель в этой точке, по предположению, либо равен нулю, либо сингулярен.

Строго говоря, показатель m различен для различных точек, и мы должны были бы писать m_i , а не просто m . Однако эта упрощенная запись не приводит к недоразумениям, поскольку при анализе полноты экстремалей будет рассматриваться одна конкретная точка q_i . Показатель степени m может быть любым вещественным числом, для которого правые части (30.4), (30.5) определены.

В промежутках между нулями и особенностями, где функция Φ или положительна, или отрицательна, локальное решение, соответственно, статично с вектором Киллинга $K = \partial_\tau$ или однородно с вектором Киллинга $K = \partial_\sigma$. Квадрат длины вектора Киллинга в обоих случаях равен конформному множителю, $K^2 = \Phi$. На горизонтах, которые, как будет показано ниже, соответствуют нулям конформного множителя, $\Phi(q) = 0$, вектор Киллинга является светоподобным. При этом он отличен от нуля, что будет продемонстрировано в координатах Эддингтона–Финкельштейна (см. раздел 30.6).

При конформных преобразованиях вид метрики (30.1), (30.2) меняется, т.к. переменная q становится зависящей от обеих координат на плоскости. При растяжке координат $\zeta \mapsto \tilde{\zeta} = k\zeta$, $k = \text{const} \neq 0$, которые образуют подгруппу группы конформных преобразований, метрика сохраняет свой вид с $\tilde{q} = k^{-1}q$, $\tilde{\Phi} = k^{-2}\Phi$.

По сути дела, формулы (30.1), (30.2) задают четыре различные метрики: из-за наличия знака модуля у производной в (30.2) есть две области со статической метрикой и две области с однородной метрикой, отличающиеся знаком производной $dq/d\zeta$. Будем обозначать эти области римскими цифрами:

$$\begin{aligned} \text{I} : \quad & \Phi > 0, \quad dq/d\sigma > 0, \\ \text{II} : \quad & \Phi < 0, \quad dq/d\tau < 0, \\ \text{III} : \quad & \Phi > 0, \quad dq/d\sigma < 0, \\ \text{IV} : \quad & \Phi < 0, \quad dq/d\tau > 0. \end{aligned} \quad (30.6)$$

Порядок следования областей будет ясен из дальнейшего построения.

Статическое локальное решение в области III получается из локального решения в области I пространственным отражением $\sigma \rightarrow -\sigma$, а однородное локальное решение в области IV связано с локальным решением в области II обращением времени $\tau \rightarrow -\tau$. Поскольку изменение знака конформного множителя у метрики можно компенсировать перестановкой пространственной и временной координаты $\tau \leftrightarrow \sigma$, то однородные локальные решения в областях II и IV можно получить из стационарных локальных решений для конформного множителя $-\Phi$ в областях III и I с последующим поворотом плоскости τ, σ на угол $\pi/2$ по часовой стрелке.

Пример 30.1.1. Рассмотрим t, r компоненты метрики Шварцшильда вне горизонта при $r > 2M > 0$

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{2M}{r}}.$$

Эту двумерную метрику можно записать в вейлевски плоском виде

$$ds^2 = \left(1 - \frac{2M}{r}\right) (dt^2 - d\sigma^2),$$

преобразовав радиальную координату $r = r(\sigma)$, где функция $r(\sigma)$ удовлетворяет уравнению

$$\left|\frac{dr}{d\sigma}\right| = 1 - \frac{2M}{r}.$$

Это – статические области I, III. Здесь роль параметра q играет радиус r .

Если координата r лежит в интервале $0 < r < 2M$, то ее следует назвать временем, т.е. произвести замену $r \leftrightarrow t$. Следовательно, при $0 < t < 2M$ (под горизонтом) метрика Шварцшильда имеет вид

$$ds^2 = -\frac{dt^2}{1 - \frac{2M}{t}} + \left(1 - \frac{2M}{t}\right) dr^2.$$

Ее также можно записать в вейлевски плоском виде

$$ds^2 = \left|1 - \frac{2M}{t}\right| (d\tau^2 - dr^2),$$

где

$$\left|\frac{dt}{d\tau}\right| = -\left(1 - \frac{2M}{t}\right).$$

Таким образом мы получили однородную метрику в областях II, IV. В этом случае роль параметра q играет параметр t .

Как видим, двумерная t, r часть метрики Шварцшильда в точности имеет вид (30.1) с конформным множителем

$$\Phi(q) = 1 - \frac{2M}{q}, \quad (30.7)$$

где $q = r, t$. Конформный множитель $\Phi(q)$ для метрики Шварцшильда имеет простой полюс при $q = 0$ и простой нуль (горизонт) в точке $q = 2M$. $\square$

Для метрики (30.1) в областях I и III можно перейти к координатам τ, q , в которых метрика примет вид

$$ds^2 = \Phi(q) d\tau^2 - \frac{dq^2}{\Phi(q)}.$$

В областях II и IV можно перейти к координатам q, σ :

$$ds^2 = -\frac{dq^2}{\Phi(q)} + \Phi(q) d\sigma^2.$$

Координаты τ, q в областях I, III и q, σ в областях II, IV называются *координатами Шварцшильда*. Метрика в этих координатах имеет простой вид, причем все компоненты метрики заданы явно. Недостатком координат Шварцшильда является то, что они не различают между собой области I и III, а также II и IV. Это существенно, потому что для построения глобальных решений необходимо использовать все области.

Перейдем к вычислению геометрических объектов для метрики (30.1). Символы Кристоффеля имеют различный вид в различных областях. Ниже приведены явные выражения только для ненулевых компонент:

$$\text{I : } \Gamma_{\tau\tau}^{\sigma} = \Gamma_{\tau\sigma}^{\tau} = \Gamma_{\sigma\tau}^{\tau} = \Gamma_{\sigma\sigma}^{\sigma} = \frac{1}{2}\Phi', \quad (30.8)$$

$$\text{II : } \Gamma_{\tau\tau}^{\tau} = \Gamma_{\tau\sigma}^{\sigma} = \Gamma_{\sigma\tau}^{\sigma} = \Gamma_{\sigma\sigma}^{\tau} = \frac{1}{2}\Phi', \quad (30.9)$$

$$\text{III : } \Gamma_{\tau\tau}^{\sigma} = \Gamma_{\tau\sigma}^{\tau} = \Gamma_{\sigma\tau}^{\tau} = \Gamma_{\sigma\sigma}^{\sigma} = -\frac{1}{2}\Phi', \quad (30.10)$$

$$\text{IV : } \Gamma_{\tau\tau}^{\tau} = \Gamma_{\tau\sigma}^{\sigma} = \Gamma_{\sigma\tau}^{\sigma} = \Gamma_{\sigma\sigma}^{\tau} = -\frac{1}{2}\Phi', \quad (30.11)$$

где штрих обозначает производную по q . В статичных областях отличны от нуля символы Кристоффеля с нечетным числом пространственных индексов, а в однородных областях – с нечетным числом временных индексов. В статичных и однородных областях символы Кристоффеля отличаются знаком. Это является следствием того, что однотипные области связаны преобразованием $\zeta \rightarrow -\zeta$, а символы Кристоффеля линейны по производным. У тензора кривизны в каждой области отличны от нуля только четыре компоненты:

$$\text{I, III : } R_{\tau\sigma\tau}^{\sigma} = -R_{\sigma\tau\tau}^{\sigma} = R_{\tau\sigma\sigma}^{\tau} = -R_{\sigma\tau\sigma}^{\tau} = -\frac{1}{2}\Phi''\Phi, \quad (30.12)$$

$$\text{II, IV : } R_{\tau\sigma\tau}^{\sigma} = -R_{\sigma\tau\tau}^{\sigma} = R_{\tau\sigma\sigma}^{\tau} = -R_{\sigma\tau\sigma}^{\tau} = \frac{1}{2}\Phi''\Phi. \quad (30.13)$$

Тензор Риччи диагонален:

$$\text{I, III : } R_{\tau\tau} = -R_{\sigma\sigma} = -\frac{1}{2}\Phi''\Phi, \quad (30.14)$$

$$\text{II, IV : } R_{\tau\tau} = -R_{\sigma\sigma} = \frac{1}{2}\Phi''\Phi. \quad (30.15)$$

При этом во всех четырех областях скалярная кривизна одинакова,

$$\text{I, II, III, IV : } R = -\Phi''. \quad (30.16)$$

Замечание. При заданном конформном множителе Φ это алгебраическое уравнение связывает q с R . Отсюда вытекает, что функция $q = q(\tau, \sigma)$, так же, как и скалярная кривизна, может рассматриваться, как функция на плоскости. В той области, где уравнение (30.16) разрешимо относительно q в качестве одной из координат можно выбрать скалярную кривизну R . $\square$

Максимальное продолжение многообразия означает, что, если у поверхности (пространства-времени) с метрикой (30.1) есть край, лежащий на конечном расстоянии, т.е. соответствующий конечному значению канонического параметра для экстремалей, то он может быть только сингулярным. В противном случае многообразие можно и необходимо продолжить.

Поскольку в двумерном случае все компоненты тензора кривизны определяются скалярной кривизной, то проанализируем ее поведение подробнее.

Сначала рассмотрим

Пример 30.1.2. Из уравнения (30.16) вытекает, что для поверхностей постоянной кривизны, которые рассмотрены в разделе 30.5.5, конформный множитель представляет собой квадратичный полином:

$$\Phi(q) = -(aq^2 + bq + c), \quad a, b, c = \text{const.} \quad (30.17)$$

При этом $R = 2a$. □

Из уравнения (30.16) следует, что скалярная кривизна сингулярна вблизи точки q_i при следующих показателях степени асимптотического поведения (30.4):

$$|q_i| < \infty : \quad m < 0, \quad 0 < m < 1, \quad 1 < m < 2, \quad (30.18)$$

$$|q_i| = \infty : m > 2. \quad (30.19)$$

Отметим, что при конечных значениях q_i и при $m = 1$ скалярная кривизна может иметь отличные от нуля значения за счет поправок второго порядка. При бесконечных значениях параметра, $q \rightarrow \pm\infty$, скалярная кривизна стремится к отличной от нуля постоянной при $m = 2$ и к нулю при $m < 2$. Отметим, что ненулевое значение кривизны в конечной точке $|q_i| < \infty$ может возникнуть и при $m = 1$ за счет поправок следующего порядка в разложении (30.4).

30.2 Конформные блоки

При построении глобальных решений будет использовано понятие конформного блока, который ставится в соответствие каждому интервалу (q_i, q_{i+1}) . Внутри каждого интервала конформный множитель Φ либо строго положителен, либо строго отрицателен.

Сначала найдем область определения соответствующей метрики (30.1) на плоскости τ, σ . Для определенности, рассмотрим статическое локальное решение I на интервале (q_i, q_{i+1}) . Тогда временная координата пробегает всю числовую ось, $\tau \in \mathbb{R}$. Уравнение (30.2) для пространственной координаты σ принимает вид

$$\frac{dq}{d\sigma} = \Phi(q). \quad (30.20)$$

Постоянная интегрирования этого уравнения соответствует сдвигу пространственной координаты, $\sigma \mapsto \sigma + \text{const}$, т.е. выбору начала отсчета σ . Из уравнения (30.20) следует, что область изменения координаты σ зависит от сходимости интеграла

$$\sigma_{i,i+1} \sim \int^{q_i, q_{i+1}} \frac{dq}{\Phi(q)} \quad (30.21)$$

в граничных точках. Интеграл (30.21) сходится или расходится в зависимости от показателя степени m :

$$\begin{aligned} |q_i| < \infty : & \begin{cases} m < 1 & - \text{сходится,} & \text{прямая,} \\ m \geq 1 & - \text{расходится,} & \text{угол,} \end{cases} \\ |q_i| = \infty : & \begin{cases} m \leq 1 & - \text{расходится,} & \text{угол,} \\ m > 1 & - \text{сходится,} & \text{прямая.} \end{cases} \end{aligned} \quad (30.22)$$

В этой таблице справа указана форма границы соответствующих конформных блоков, которые будут введены ниже. Если на обоих концах интервала (q_i, q_{i+1}) интеграл расходится, то $\sigma \in (-\infty, \infty)$, и метрика определена на всей плоскости τ, σ . Если на одном из концов q_{i+1} или q_i интеграл сходится, то метрика определена на полуплоскости $\sigma \in (-\infty, \sigma_{i+1})$ или $\sigma \in (\sigma_i, \infty)$, соответственно. При этом выбор граничных точек σ_{i+1} и σ_i произволен, и, не ограничивая общности, можно положить $\sigma_{i,i+1} = 0$. Если на обоих концах интервала (q_i, q_{i+1}) интеграл сходится, то локальное решение определено в полосе $\sigma \in (\sigma_i, \sigma_{i+1})$, при этом совместить с началом координат можно только один из концов интервала.

Для наглядного изображения максимально продолженных решений введем понятие конформного блока. С этой целью отобразим плоскость τ, σ на квадрат вдоль светоподобных направлений:

$$\xi = \tau + \sigma, \quad \eta = \tau - \sigma. \quad (30.23)$$

Для этого произведем конформную замену переменных:

$$u = u(\xi), \quad v = v(\eta), \quad u, v \in \mathcal{C}^{l+1}, \quad (30.24)$$

где функции u и v ограничены на всей оси, а их первые производные положительны. Например, $u = \arctg \xi$, $v = \arctg \eta$. Предположение о классе гладкости функций перехода сохраняет класс гладкости метрики после преобразования координат. Тогда статическому локальному решению, определенному на всей плоскости τ, σ , будет поставлен в соответствие квадратный *конформный блок*, изображенный на рис. 30.1, а.

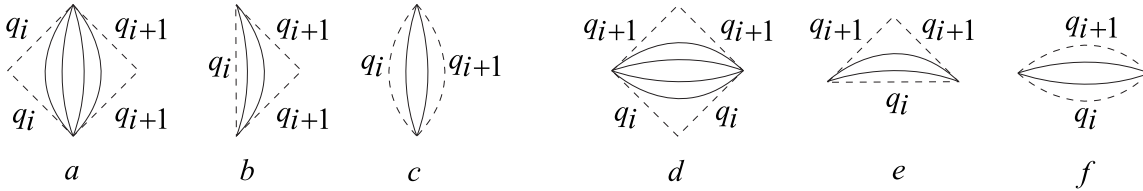


Рис. 30.1: Конформные блоки для статических (a, b, c) и однородных (d, e, f) локальных решений. Тонкие сплошные линии обозначают траектории Киллинга.

В общем случае под конформным блоком понимается конечная часть плоскости u, v и заданная на ней метрика (30.1) в координатах (30.23), (30.24). Если потребовать, чтобы экстремали подходили к границе конформного блока под всеми возможными углами, то это однозначно фиксирует асимптотическое поведение функций перехода (30.24) [214]. В остальном функции перехода остаются произвольными. Мы не будем останавливаться на этом вопросе подробно, т.к. гладкость метрики на горизонтах будет доказана путем перехода к координатам Эддингтона–Финкельштейна в разделе 30.6.

Значение переменной q и, значит, скалярной кривизны постоянно вдоль времени-подобных траекторий Киллинга, которые внутри блока показаны тонкими сплошными линиями. Переменная q непрерывно меняется слева направо: монотонно возрастает в области I и монотонно убывает в области III. При этом на двух левых сторонах квадрата она принимает значение q_i , а на двух правых – значение q_{i+1} . Нижняя и верхняя вершины квадрата являются существенно особыми точками, т.к. в этих точках предел q зависит от пути, по которому произвольно выбранная последовательность подходит к вершине. Будем говорить, что у статического конформного блока

есть две границы: левая и правая, на которых переменная q принимает значения q_i и q_{i+1} , соответственно.

Если решение уравнения (30.20) определено на полуинтервале ($\sigma_i = 0, \sigma_{i+1} = \infty$), то статическому локальному решению ставится в соответствие треугольный конформный блок, пример которого показан на рис. 30.1, *b*. При этом конечной граничной точкой $\sigma_i = 0$ ставится в соответствие вертикальная прямая. Этого всегда можно добиться путем выбора функций $u(\xi)$ и $v(\eta)$.

В случае, когда решение уравнения (30.20) определено на конечном интервале (σ_i, σ_{i+1}), конформный блок изображается в виде линзы, изображенной на рис. 30.1, *c*. За счет выбора функций $u(\xi)$ и $v(\eta)$ левую или правую границу этого конформного блока можно сделать вертикальной, но не обе границы одновременно.

Напомним, что каждому статическому интервалу (q_i, q_{i+1}) ставится в соответствие два конформных блока, т.к. благодаря знаку модуля, уравнение (30.2) инвариантно относительно инверсии пространственной координаты $\sigma \leftrightarrow -\sigma$.

Аналогично каждому однородному локальному решению ставится в соответствие один из конформных блоков, изображенных на рис. 30.1, *d, e, f*, или перевернутый блок, полученный обращением времени $\tau \leftrightarrow -\tau$. Для этих конформных блоков переменная q постоянна вдоль пространственноподобных траекторий Киллинга и монотонно возрастает или убывает от значения q_i на нижней границе, до значения q_{i+1} на верхней границе. Левая и правая вершины однородных конформных блоков являются существенно особыми точками.

Конформные блоки для статических и однородных локальных решений будем называть, соответственно, статическими и однородными.

Итак, мы показали, что лоренцева поверхность, на которой задана конформно плоская метрика (30.1) на каждом интервале (q_i, q_{i+1}) диффеоморфна одному из конформных блоков. Для построения максимально продолженных решений необходимо найти и проанализировать экстремали на этих поверхностях, при необходимости продолжить многообразие и метрику, а также доказать дифференцируемость метрики при склеивании конформных блоков. Использование конформных блоков для построения глобальных решений удобно и наглядно, т.к. все светоподобные экстремали изображаются в виде двух семейств прямых линий, проходящих под углом $\pm\pi/4$ к оси времени. Поэтому при продолжении поверхности конформные блоки можно склеивать вдоль ребер, сохраняя гладкость светоподобных экстремалей.

30.3 Экстремали

30.3.1 Форма экстремалей

Для того, чтобы понять, как устроено глобальное решение для метрики (30.1), необходимо проанализировать поведение экстремалей $\{\zeta^\alpha(t)\} = \{\tau(t), \sigma(t)\}$, $t \in \mathbb{R}$, которые подчиняются системе обыкновенных дифференциальных уравнений второго порядка

$$\ddot{\zeta}^\alpha = -\Gamma_{\beta\gamma}{}^\alpha \dot{\zeta}^\beta \dot{\zeta}^\gamma, \quad (30.25)$$

где точка обозначает дифференцирование по каноническому параметру t и $\Gamma_{\beta\gamma}{}^\alpha$ — символы Кристоффеля.

Исследуем подробно поведение экстремалей для статического локального решения в области

$$\text{I:} \quad ds^2 = \Phi(q)(d\tau^2 - d\sigma^2), \quad \frac{dq}{d\sigma} = \Phi(q) > 0. \quad (30.26)$$

Используя выражение для символов Кристоффеля (30.8), получаем уравнения для экстремалей

$$\ddot{\tau} = -\Phi' \dot{\tau} \dot{\sigma}, \quad (30.27)$$

$$\ddot{\sigma} = -\frac{1}{2}\Phi'(\dot{\tau}^2 + \dot{\sigma}^2), \quad (30.28)$$

где штрих обозначает производную по q . Эта система уравнений имеет два первых интеграла (см. раздел 16.3):

$$\Phi(\dot{\tau}^2 - \dot{\sigma}^2) = C_0 = \text{const}, \quad (30.29)$$

$$\Phi \dot{\tau} = C_1 = \text{const}. \quad (30.30)$$

Интеграл движения (30.29) существует для любой метрики, поскольку для параметризации экстремалей используется канонический параметр. Значение постоянной C_0 определяет тип экстремали:

$C_0 > 0$ – времениподобные экстремали,

$C_0 = 0$ – светоподобные экстремали,

$C_0 < 0$ – пространственноподобные экстремали.

Отметим, что тип экстремали не может меняться от точки к точке. Второй из интегралов (30.30) связан с наличием вектора Киллинга, т.е. с симметрией задачи, и для произвольной метрики он отсутствует.

В системе уравнений для экстремалей (30.27), (30.28) конформный множитель Φ рассматривается, как сложная функция от σ : $\Phi = \Phi(q(\sigma))$, где $q(\sigma)$ – решение уравнения (30.26). Если форма экстремали $\sigma(\tau)$ найдена, то конформный множитель можно рассматривать так же, как сложную функцию от τ вдоль каждой экстремали: $\Phi = \Phi(\sigma(\tau))$. В дальнейшем анализе это будет подразумеваться.

Теорема 30.3.1. *Любая экстремаль в статическом пространстве-времени типа I, III принадлежит одному из следующих четырех классов.*

1) *Светоподобные экстремали*

$$\tau = \pm\sigma + \text{const}, \quad (30.31)$$

для которых канонический параметр определен уравнением

$$\dot{\tau} = \frac{1}{\Phi}. \quad (30.32)$$

2) *Экстремали общего вида, форма которых определена уравнением*

$$\frac{d\tau}{d\sigma} = \pm \frac{1}{\sqrt{1 - C_2 \Phi}}, \quad (30.33)$$

где C_2 – произвольная отличная от нуля постоянная, причем значения $C_2 < 0$ и $C_2 > 0$ описывают, соответственно, пространственно- и времениподобные экстремали. Канонический параметр определяется любым из двух уравнений:

$$\dot{\tau} = \frac{1}{\Phi}, \quad (30.34)$$

$$\dot{\sigma} = \pm \frac{\sqrt{1 - C_2 \Phi}}{\Phi}. \quad (30.35)$$

При этом в уравнениях (30.33) и (30.35) знаки плюс или минус выбираются одновременно.

3) Прямые пространственноподобные экстремали, параллельные оси σ и проходящие через каждую точку $\tau = \text{const}$. Канонический параметр определен уравнением

$$\dot{\sigma} = \frac{1}{\sqrt{\Phi}}. \quad (30.36)$$

4) Прямые вырожденные времениподобные экстремали, параллельные оси τ и проходящие через критические точки $\sigma_0 = \text{const}$, в которых

$$\Phi'(\sigma_0) = 0. \quad (30.37)$$

Канонический параметр для них совпадает с временной координатой,

$$t = \tau. \quad (30.38)$$

Доказательство. Доказательство теоремы сводится к интегрированию системы уравнений (30.27), (30.28). При получении интеграла движения (30.30) необходимо делить на $\dot{\tau}$ и $\dot{\sigma}$, поэтому эти вырожденные случаи рассматриваются отдельно.

Начнем с вырожденных случаев. Для $\tau = \text{const}$ уравнение (30.27) выполняется автоматически, а уравнение (30.28) после интегрирования приводит к условию

$$-\Phi\dot{\sigma}^2 = C_0 < 0.$$

Это уравнение сводится к (30.36) растяжкой канонического параметра.

Второй вырожденный случай соответствует $\dot{\sigma} = 0$. При этом уравнение (30.28) выполняется только в тех точках $\sigma = \sigma_0$, в которых $\Phi'(\sigma_0) = 0$, а уравнение (30.27) сводится к уравнению $\ddot{\tau} = 0$. Отсюда вытекает, что координату τ можно выбрать в качестве канонического параметра. Здесь можно лишь заметить справедливость равенства

$$\frac{d\Phi}{d\sigma} = \Phi \frac{d\Phi}{dq} = 0,$$

и, поскольку $\Phi > 0$, то в уравнении (30.37) дифференцирование по q можно заменить дифференцированием по σ . Таким образом исчерпываются оба специальных случая.

Остальные экстремали можно получить из анализа интегралов (30.29), (30.30), причем константа C_0 может быть произвольной, а $C_1 \neq 0$, т.к. случай $C_1 = 0$ относится к прямым экстремалим третьего класса. Уравнение (30.29) с учетом (30.30) приводит к равенству

$$\dot{\sigma} = \pm \frac{C_1}{\Phi} \sqrt{1 - C_2 \Phi}, \quad (30.39)$$

где

$$C_2 := \frac{C_0}{C_1^2}. \quad (30.40)$$

Уравнение для формы экстремалей общего вида (30.33) является следствием (30.30) и (30.39), причем значение постоянной C_2 определяет тип экстремали. Уравнения для канонического параметра получаются из (30.30) и (30.39) после растяжки. В частном случае $C_2 = 0$ получаем светоподобные экстремали (30.31). $\square$

Замечание. Из уравнения (30.33) следует, что постоянная C_2 параметризует угол, под которым экстремаль общего вида проходит через заданную точку. Можно проверить, что через произвольную точку в каждом направлении проходит одна и только одна экстремаль. Это подтверждает тот факт, что найдены все экстремали. $\square$

Доказанная теорема позволяет качественно понять поведение экстремалей для произвольного конформного множителя.

На рис. 30.2 в первом ряду показано типичное поведение конформного множителя $\Phi(q)$ с одним и тремя локальными экстремумами между двумя нулями и с двумя локальными экстремумами между нулем и особенностью для статических конформных блоков. Во втором ряду показана зависимость этих конформных множителей $\Phi(\sigma)$ от пространственной координаты. В третьем ряду показано качественное поведение времени- и пространственноподобных экстремалей. Чтобы не загромождать рисунок, мы опустили светоподобные экстремали, проходящие через каждую точку пространства-времени, а также экстремали, параллельные оси σ . Все экстремали можно сдвигать вдоль времениподобной координаты τ .

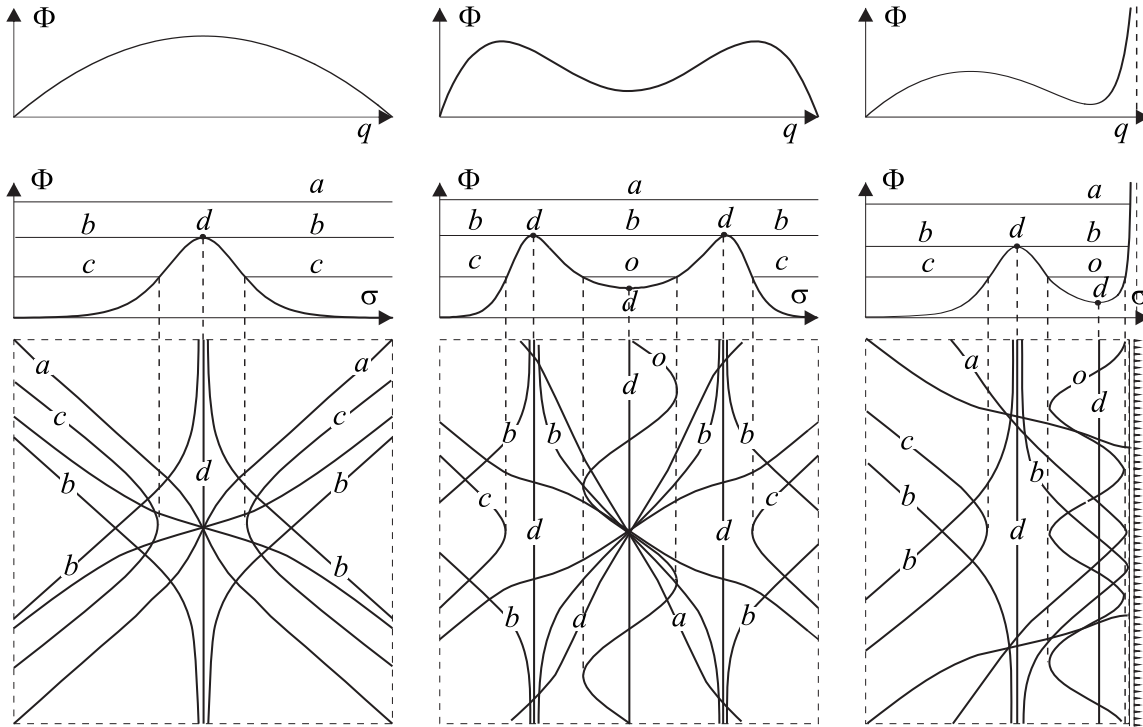


Рис. 30.2: Типичное поведение конформного множителя $\Phi(q)$ между двумя нулями и нулем и сингулярностью для статического локального решения (верхний ряд). Зависимость конформного множителя $\Phi(\sigma)$ от пространственной координаты (средний ряд). Типичные времени- и пространственно подобные экстремали (нижний ряд). Буквами отмечены времениподобные экстремали, имеющие качественно различное поведение при разных значениях постоянной C_2 . Через каждый локальный экстремум проходит вырожденная экстремаль d . Во втором и третьем случае имеются осциллирующие экстремали o .

Если обе граничные точки интервала q_i и q_{i+1} являются нулями, то из непрерывности конформного множителя Φ следует, что у него есть по крайней мере один максимум, через который проходит вырожденная экстремаль d . В общем случае через каждый локальный экстремум функции $\Phi(\sigma)$ проходит вертикальная экстремаль. На рисунках они помечены буквой d .

Пространственноподобные экстремали общего вида соответствуют отрицательным значениям постоянной C_2 в уравнении (30.33), и, поскольку $\Phi > 0$, подынтегральное выражение всюду определено. Они начинаются на левой границе кон-

формного блока и заканчиваются на правой. Пространственноподобные экстремали общего вида показаны на рисунке без каких-либо пометок.

Времениподобные экстремали общего вида имеют качественно различное поведение при различных значениях постоянной $C_2 > 0$. Они помечены буквами a, b, c и o .

Если постоянная C_2 достаточно мала,

$$C_2 < \frac{1}{\max \Phi(\sigma)},$$

то времениподобная экстремаль общего вида определена при всех значениях $\sigma \in (\sigma_i, \sigma_{i+1})$. Такие экстремали соединяют левую и правую границу конформного блока, пересекая все вырожденные экстремали, и помечены на рисунке буквой a .

Во втором ряду рис.30.2 горизонтальными линиями показаны области определения координаты σ ,

$$C_2 \Phi(\sigma) \leq 1, \quad (30.41)$$

при различных значениях постоянной C_2 .

Если у функции Φ имеется локальный минимум, которому также соответствует вырожденная экстремаль, то среди экстремалей общего вида существуют времениподобные экстремали, осциллирующие вблизи локального минимума, которые помечены буквой o . Это следует из уравнения (30.33), т.к. при достаточно больших положительных значениях C_2 область изменения координаты σ определяется неравенством (30.41). Критические точки σ_* , в которых $\Phi(\sigma_*) = 1/C_2$, являются точками поворота при конечном значении $\tau = \tau_*$ тогда и только тогда, когда интеграл

$$\int^{\tau_*} d\tau = \pm \int^{\sigma_*} \frac{d\sigma}{\sqrt{1 - C_2 \Phi}} \quad (30.42)$$

сходится. Этот интеграл сходится, если точка σ_* не совпадает с локальным максимумом. То есть поворот экстремали происходит при конечном значении τ_* и канонического параметра, что вытекает из уравнения (30.35).

При том же значении постоянной C_2 существуют времениподобные неосциллирующие экстремали общего вида, которые приходят и уходят на одну и ту же (левую или правую) границу. У них есть точка поворота при конечных значениях τ . Такие экстремали на рисунке помечены буквой c .

Если критическая точка σ_* совпадает с локальным максимумом, то

$$\Phi \simeq \frac{1}{C_2} - C_3(\sigma - \sigma_*)^2 + \dots, \quad C_3 = \text{const} > 0,$$

и интеграл (30.42) расходится. Это означает, что критическая точка достигается при бесконечном значении τ_* и канонического параметра. То есть соответствующая экстремаль в этом направлении полна. На рисунке такие экстремали помечены буквой b .

Приведенная теорема описывает все экстремали для статических локальных решений. Поведение экстремалей для однородных локальных решений аналогично. Они получаются заменой $\Phi \mapsto -\Phi$, перестановкой временной и пространственной координат $\tau \leftrightarrow \sigma$ и поворотом на угол $\pi/2$ всей плоскости.

30.3.2 Асимптотика экстремалей

Для решения вопроса о том является ли метрика, определенная на τ, σ плоскости, полной по экстремалиям или локальное решение следует продолжить, необходимо исследовать асимптотику и полноту экстремалей вблизи границы конформного блока. Забегая вперед, заметим, что все экстремали, за исключением осциллирующих, имеют асимптотику при приближении к границе конформного блока.

Рассмотрим квадратный конформный блок. В верхнюю и нижнюю вершины конформного блока попадают все экстремали, для которых

$$\lim_{\tau \rightarrow \pm\infty} \left| \frac{d\tau}{d\sigma} \right| > 1,$$

если предел существует. В левую и правую вершины конформного блока попадают экстремали, для которых

$$\lim_{\sigma \rightarrow \pm\infty} \left| \frac{d\tau}{d\sigma} \right| < 1. \quad (30.43)$$

На стороны квадратного конформного блока могут попасть только экстремали, имеющие светоподобную асимптотику:

$$\lim_{\tau, \sigma \rightarrow \pm\infty} \left| \frac{d\tau}{d\sigma} \right| = 1. \quad (30.44)$$

В принципе, в вершину конформного блока могут попасть также экстремали, имеющие светоподобную асимптотику. Однако, как будет показано ниже, этого не происходит.

Если одна из границ q_i , например, статического конформного блока прямая, то на нее попадают экстремали со светоподобной (30.44) и пространственноподобной (30.43) асимптотикой.

Осциллирующие экстремали всегда зажаты прямыми вырожденными экстремалиями. Поэтому они попадают в ту же вершину конформного блока, что и прямые вырожденные экстремали 4) из теоремы 30.3.1.

Рассмотрим асимптотику экстремалей вблизи границы статических конформных блоков. Любая светоподобная экстремаль на квадратном конформном блоке начинается и заканчивается на противоположных сторонах. В случае треугольного конформного блока или блока типа линзы светоподобные экстремали начинаются или заканчиваются на времениподобных границах.

Для статических локальных решений типа I, III прямые экстремали, параллельные пространственной оси σ начинаются в левой и заканчиваются в правой вершине квадратного конформного блока. В случае треугольного блока и блока типа линзы эти экстремали начинаются или (и) заканчиваются на времениподобной границе. Вырожденные экстремали начинаются в нижней и заканчиваются в верхней вершине статического конформного блока. Осциллирующие экстремали зажаты вырожденными экстремалиями и поэтому также начинаются и заканчиваются в нижней и верхней вершине.

Таким образом, мы поняли в какие граничные точки конформного блока попадают все прямые экстремали, а также осциллирующие экстремали общего вида.

Осталось рассмотреть поведение неосциллирующих экстремалей общего вида вблизи границы. Сначала определим точки границы конформного блока, в которые они попадают. Допустим, что граница q_i конформного блока является нулем, $\Phi(q_i) = 0$,

$|q_i| < \infty$. Этому соответствует положительный показатель $m > 0$ (30.4). Соответствующая граница конформного блока является углом при $m \geq 1$ и прямой при $0 < m < 1$ (30.22). В этом случае из уравнения для экстремалей (30.33) следует, что векторы, касательные к ним, имеют светоподобную асимптотику

$$\lim_{q \rightarrow q_i} \left| \frac{d\tau}{d\sigma} \right| = 1.$$

Если граница конформного блока является углом, то экстремали общего вида попадают на стороны угла, а не в вершину. Действительно, из уравнений (30.34), (30.35) для экстремалей, идущих в правую верхнюю сторону квадратного конформного блока, следует равенство

$$\frac{d\eta}{d\xi} = \frac{1 - \sqrt{1 - C_2\Phi}}{1 + \sqrt{1 + C_2\Phi}}.$$

Для малых $\Phi \ll 1$ разложим правую часть этого равенства в ряд. В первом порядке по Φ получим равенство

$$\frac{d\eta}{d\xi} \simeq \frac{1}{2} \left[1 - \left(1 - \frac{C_2\Phi}{2} \right) \right] = \frac{C_2}{4}\Phi.$$

Откуда вытекает, что интеграл

$$\int d\eta \sim \int d\xi \Phi \sim \int d\sigma \Phi \sim \int^{q_i} dq = q_i$$

сходится. Это означает, что точка $\xi \rightarrow \infty$ достигается при конечных значениях η , и это соответствует некоторой точке на стороне угла, а не вершина. Если граница q_i конформного блока является времениподобной, то экстремали общего вида попадают на нее.

Пусть теперь на границе $\Phi(q_i) = \infty$, $|q_i| < \infty$. В этом случае $m < 0$, граница времениподобна и достигается при конечных σ_i . Из равенства (30.33) следует, что касательный вектор к пространственноподобной экстремали при $\Phi \rightarrow \infty$ имеет асимптотику

$$\frac{d\tau}{d\sigma} \sim \frac{1}{\sqrt{\Phi}} \rightarrow 0.$$

То есть, они попадают на границу $|\sigma_i| < \infty$ под прямым углом.

Времениподобные экстремали общего вида не могут попасть на времениподобную границу, потому что правая часть в (30.33) при $C_2 > 0$ вблизи особенности $\Phi(q_i) = \infty$ становится отрицательной. Эти экстремали вблизи границы $|\sigma_i| < \infty$ имеют точку поворота при конечных значениях τ , по той же причине, что и осциллирующие экстремали.

30.3.3 Полнота экстремалей

Теорема 30.3.1 позволяет проанализировать полноту экстремалей при подходе к границе конформного блока. Для определенности рассмотрим стационарное локальное решение. Прежде всего заметим, что вырожденные экстремали всегда полны, т.к. канонический параметр для них совпадает с временем (30.38). Осциллирующие экстремали также полны поскольку делают бесконечное число осцилляций, каждая из

которых соответствует конечному изменению канонического параметра. Если же экстремаль общего вида при $\tau \rightarrow \pm\infty$ приближается к вырожденной экстремали, то она полна, поскольку из уравнения для канонического параметра (30.34) следует, что

$$\lim_{\sigma \rightarrow \sigma_0} t \rightarrow \int^{\infty} d\tau \Phi(\sigma_0) \rightarrow \infty.$$

Для стационарного локального решения в нижнюю и верхнюю вершины попадают только вырожденные и осциллирующие экстремали, и поэтому они всегда полны. Возможна ситуация, когда эти экстремали отсутствуют. В этом случае будем считать эти вершины полными, потому что любая времениподобная кривая имеет бесконечную длину при $\tau \rightarrow \pm\infty$. Это означает, что нижняя и верхняя вершины стационарного блока всегда полны, т.е. представляют собой временную бесконечность прошлого и будущего, соответственно.

Полнота светоподобных экстремалей определяется уравнением (30.32), которое позволяет вычислить предел

$$\lim_{\sigma \rightarrow \pm\infty} t \rightarrow \int^{\sigma_i} d\sigma \Phi = \int^{q_i} dq = q_i, \quad (30.45)$$

поскольку в данном случае $\dot{\tau} = \dot{\sigma}$. Это значит, что при приближении к границе конформного блока они неполны при конечных q_i и полны при $|q_i| = \infty$.

Полнота неосциллирующих экстремалей общего вида, попадающих на границу конформного блока, соответствующую значению q_i , вытекает из уравнения (30.35):

$$\lim_{q \rightarrow q_i} t \rightarrow \int^{\sigma_i} d\sigma \frac{\Phi}{\sqrt{1 - C_2 \Phi}} \sim \int^{q_i} \frac{dq}{\sqrt{1 - C_2 \Phi}}. \quad (30.46)$$

Вблизи нулей $\Phi \rightarrow 0$ поведение экстремалей общего вида такое же, как и у светоподобных экстремалей (30.45). Особенность $\Phi \rightarrow \infty$ достигается только пространственноподобными экстремальями, $C_2 < 0$, полнота которых определяется интегралом

$$\lim_{q \rightarrow q_i} t \rightarrow \int^{q_i} \frac{dq}{\sqrt{\Phi}}. \quad (30.47)$$

т.к. единицу и постоянную C_2 в формуле (30.46) можно отбросить. Эти экстремали неполны в конечных точках $|q_i| < \infty$ при $m < 0$. В бесконечно удаленных точках $|q_i| = \infty$ пространственноподобные экстремали полны при $0 < m \leq 2$ и неполны при $m > 2$.

Полнота прямых экстремалей, параллельных оси σ , определяется уравнением (30.36) и при приближении к границе q_i дается интегралом (30.47). Это значит, что их полнота вблизи особенностей такая же, как и у экстремалей общего вида. Вблизи нулей они всегда полны за исключением случая простого нуля в конечной точке, где они неполны.

Таким образом проанализирована полнота всех элементов границы. Итог анализа приведен на рис. 30.3. Здесь в зависимости от показателя степени m (30.4) для конечных и бесконечных значений q_i показаны соответствующие элементы границы конформных блоков. Для определенности, показана правая граница статических конформных блоков типа I. На рисунке времениподобная граница показана вертикальной линией, а светоподобная – углом. Если на границе скалярная кривизна сингулярна (30.18), (30.19), то на ней присутствуют выступы. Неполные и полные границы показаны, соответственно, пунктиром и жирной сплошной линией. Исключение

пространства-времени, и только они определяют неполную по экстремалиям границу конформного блока, на которой кривизна конечна. Во всех остальных случаях граница либо полна, либо на ней кривизна сингулярна. Поэтому решения вида (30.1) необходимо продолжить только через нули функции $\Phi(q)$. Напомним, что нулям конформного множителя (горизонтам) соответствуют границы в виде угла (см. рис. 30.3).

Сформулируем правила, по которым осуществляется максимальное продолжение поверхности с метрикой (30.1) при заданной функции $\Phi(q)$. Дифференцируемость максимально продолженной по экстремалиям псевдоримановой поверхности, построенной по этим правилам дается приведенной ниже теоремой 30.4.1.

1. Каждое глобальное решение для метрики (30.1) соответствует некоторому интервалу значений переменной $q \in (q_-, q_+)$, где $q_{\pm}$ либо бесконечно удаленные точки $q_{\pm} = \pm\infty$, либо сингулярности кривизны, определяемые условием (30.18). Внутри интервала сингулярности должны отсутствовать.
2. Если внутри интервала (q_-, q_+) конформный блок не имеет нулей, то соответствующий конформный блок представляет собой глобальное решение.
3. Если внутри интервала (q_-, q_+) нули существуют, то пронумеруем их в порядке возрастания, $\Phi(q_j) = 0$, $j = 1, \dots, n$, и поставим в соответствие каждому интервалу $(q_-, q_1), \dots, (q_n, q_+)$ пару статических или однородных конформных блоков для $\Phi > 0$ и $\Phi < 0$, соответственно.
4. Склеим конформные блоки вдоль границ, соответствующих нулям q_j , причем склеиваются только блоки, соответствующие соседним интервалам (q_{j-1}, q_j) и (q_j, q_{j+1}) .
5. Диаграмма Картера–Пенроуза получается склеиванием всех соседних конформных блоков. Она представляет собой связную фундаментальную область, если внутри интервала (q_-, q_+) конформный множитель меняет знак. Если $\Phi \geq 0$ или $\Phi \leq 0$ всюду внутри интервала (q_-, q_+) , то получим две фундаментальные области, которые получаются друг из друга отражением пространства или времени.
6. При наличии в интервале (q_-, q_+) только одного нуля нечетного порядка граница фундаментальной области состоит из границ конформных блоков, соответствующих точкам q_- и q_+ , и диаграмма Картера–Пенроуза представляет глобальное решение.
7. При наличии одного нуля четного порядка или двух и более нулей произвольного порядка граница фундаментальной области включает нули конформного множителя, и ее необходимо либо продолжить периодически в пространстве и (или) во времени, либо отождествить противоположные стороны вдоль горизонтов.
8. Если у диаграммы Картера–Пенроуза фундаментальная группа тривиальна, то она представляет собой универсальную накрывающую глобального решения.
9. Если у диаграммы Картера–Пенроуза фундаментальная группа нетривиальна, то следует построить соответствующую универсальную накрывающую.

Определение. *Диаграммой Картера–Пенроуза* называется образ на плоскости максимально продолженной поверхности, соответствующей интервалу (q_-, q_+) и полученный в результате склеивания всех конформных блоков для интервала (q_-, q_+) по правилам 1)–9). $\square$

Диаграммы Картера–Пенроуза являются наглядным изображением максимально продолженной поверхности, т.к. оба класса непересекающихся между собой светоподобных экстремалей изображаются двумя классами перпендикулярных прямых, как на плоскости Минковского.

Поясним приведенные правила. Утверждение 1) является следствием того, что продолжение решения через точки $q_{\pm}$ невозможно, т.к. эти точки либо полны, либо сингулярны. Решения продолжаются только через горизонты $|q_j| < \infty$ с $m = 1$ или $m \geq 2$. В этих случаях граница конформного блока является углом. Если точка q_j – нуль нечетного порядка, то все четыре конформных блока, соответствующих интервалам (q_{j-1}, q_j) и (q_j, q_{j+1}) склеиваются согласно правилам 3) и 4) вокруг вершины q_j , причем эта точка является седловой для конформного множителя $\Phi(q)$, рассматриваемого как функция на диаграмме Картера–Пенроуза. Если q_j – нуль четного порядка, то вдоль соответствующего горизонта склеиваются между собой только области одного типа, например, I–I или III–III. Если внутри интервала (q_-, q_+) функция Φ не меняет знака, то возникают две несвязные между собой фундаментальные области, каждую из которых можно периодически продолжить. При изменении знака Φ возникает седловая точка, и области различных типов образуют связную фундаментальную область, которую можно либо периодически продолжить, либо нет, в зависимости от значений q на границе. Это составляет содержание правил 5) и 6).

Правила 5), 6) и 7) являются утверждениями, которые будут доказаны дальнейшим построением.

Сформулируем основную теорему, оправдывающую приведенные выше правила построения глобальных решений.

Теорема 30.4.1. *Универсальное накрывающее пространство, построенное по правилам 1)–9), представляет собой максимально продолженную вдоль экстремалей дифференцируемую псевдориманову поверхность класса $\mathcal{C}^{l+1}$ с метрикой класса $\mathcal{C}^l$, $l \geq 2$, такую, что внутренность каждого конформного блока изометрична поверхности с метрикой (30.1).*

Доказательство. По построению, внутренность каждого конформного блока покрывается одной картой и является многообразием класса $\mathcal{C}^{l+1}$, что следует из (30.24). Поэтому внутри конформного блока метрика принадлежит тому же классу гладкости, что и конформный множитель. При этом функции перехода (30.24) изометрично отображают плоскость τ, σ или ее часть, на которой определена метрика, на внутренность конформного блока. Поэтому необходимо доказать дифференцируемость многообразия и метрики только на горизонтах и в седловых точках. Это сделано путем перехода к новым системам координат в разделах 30.6 и 30.7, соответственно. $\square$

Перед завершением доказательства теоремы рассмотрим несколько примеров, чтобы пояснить, как пользоваться правилами построения глобальных решений в конкретных случаях, и какие глобальные решения могут возникнуть.

30.5 Примеры

Если в результате решения уравнений движения какой-либо модели гравитации возникает двумерная метрика вида (30.1), то, следуя правилам 1)–9) из предыдущего раздела можно построить глобальное решение, не заботясь о переходе к новым координатным системам, покрывающим большие области. Достоинством рассматриваемого метода является его конструктивность, т.к. для построения глобальных решений достаточно элементарного анализа конформного множителя $\Phi(q)$. Начнем с трех известных примеров из общей теории относительности.

30.5.1 Решение Шварцшильда

Конформный множитель для решения Шварцшильда имеет вид (30.7). Он имеет простой полюс в точке $q = 0$ ($m = -1$), который соответствует сингулярности кривизны (30.18). Точка $q_1 = 2M$ ($m = 1$) является простым нулем и соответствует горизонту. При $M > 0$ нуль лежит на положительной полуоси. Значения $q = \pm\infty$ ($m = 0$) соответствуют асимптотически плоской пространственной бесконечности ($R = 0$). Поведение конформного множителя показано на рис. 30.4 слева.

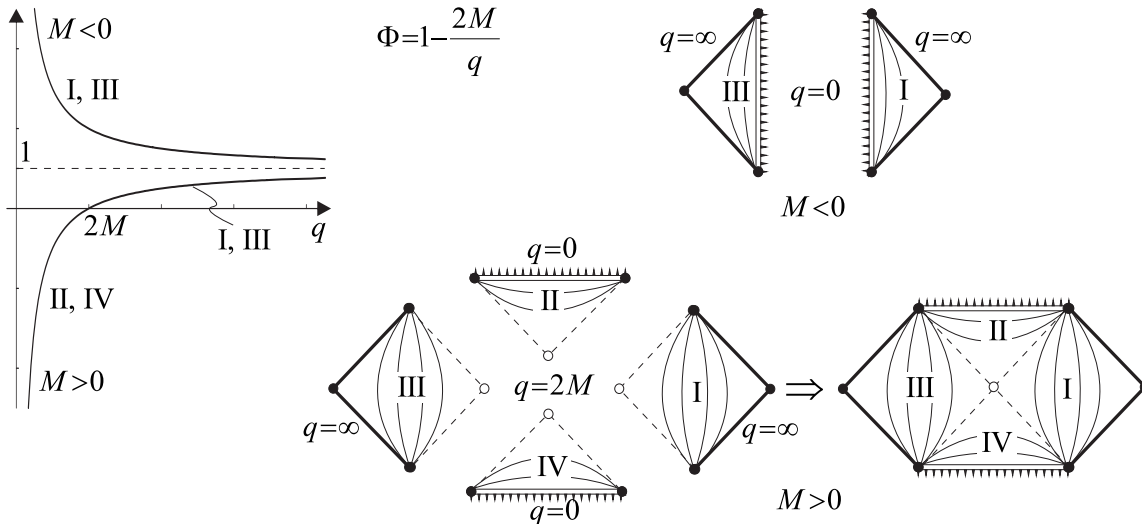


Рис. 30.4: Поведение конформного множителя для решения Шварцшильда для положительной и отрицательной массы. При $M < 0$ существует два глобальных решения, представленных треугольными диаграммами Картера–Пенроуза. При $M > 0$ четыре конформных блока, соответствующих двум интервалам $(0, 2M)$ и $(2M, \infty)$, склеиваются в одно глобальное решение.

Бесконечному интервалу $q \in (-\infty, \infty)$ соответствует два глобальных решения для положительных и отрицательных q . Рассмотрим случай $M > 0$. Для положительных значений $q \in (0, \infty)$ имеем $q_- = 0$, $q_1 = 2M$ и $q_+ = \infty$. Каждому интервалу $(q_-, q_1) = (0, 2M)$ и $(q_1, q_+) = (2M, \infty)$ соответствует по два однородных (II, IV) и статических (I, III) конформных блока, показанных на рис. 30.4 внизу в центре. При этом элементы границы определяются по рис. 30.3. Из этих четырех конформных блоков по правилу 4) единственным образом склеивается глобальное решение, показанное на рис. 30.4. Эта диаграмма Картера–Пенроуза представляет собой расширение Крускала–Секереша решения Шварцшильда [216, 217]. Напомним, что расширение Крускала–Секереша представляет собой запись метрики Шварцшильда в

такой системе координат, которая покрывает сразу все области I–IV. Картер впервые изобразил это расширение в виде конечной области на плоскости [227].

Поскольку фундаментальная группа для полученной диаграммы тривиальна, то построенное глобальное решение является универсальной накрывающей для других глобальных решений. Оба статических конформных блока I или III диффеоморфны внешнему решению Шварцшильда при $r > 2M$. Теорема 30.4.1 обеспечивает, в данном случае, гладкость соответствующего глобального решения, и нет необходимости в явном построении глобальной системы координат.

Аналогичные глобальные решения будут возникать для широкого класса метрик, у которых конформный множитель ведет себя качественно так же, как и нижняя ветвь на рис. 30.4. То есть определен на полуинтервале (q_-, ∞) , имеет особенность в q_- , один нуль и стремится к постоянной на бесконечности.

Отметим, что поверхность, представленная диаграммой Картера–Пенроуза имеет непостоянную скалярную кривизну,

$$R = \frac{4M}{q^3}. \quad (30.48)$$

Это – двумерная скалярная кривизна на лоренцевой поверхности. Скалярная кривизна, соответствующая четырехмерной метрике Шварцшильда с учетом угловой зависимости тождественно равна нулю в силу уравнений Эйнштейна. Заметим, что двумерная скалярная кривизна (30.48) совпадает с инвариантным собственным значением четырехмерного тензора Вейля [169]. Центр диаграммы Картера–Пенроуза является седловой точкой для переменной q и, следовательно, для скалярной кривизны (30.48). Эта точка неполна и отмечена незакрашенной окружностью.

Физическая интерпретация диаграммы Картера–Пенроуза для решения Шварцшильда состоит в следующем. В общей теории относительности пространство-время четырехмерно. Поэтому максимально продолженное сферически симметричное решение вакуумных уравнений Эйнштейна представляет собой топологическое произведение сферы S^2 на диаграмму Картера–Пенроуза, изображенную на рис. 30.4 и представляющую собой ту часть пространства-времени, которая описывается координатами t и r . Нижняя и верхняя сингулярные границы диаграммы Картера–Пенроуза, умноженные на сферу, называются, соответственно, *белой* и *черной дырой*. Сфера $r = 2M$ называется *горизонтом* черной дыры. Конформный блок II, лежащий под горизонтом черной дыры, соответствует внутренности черной дыры. Если выбрать произвольную точку в области II, то любая времениподобная или светоподобная кривая неизбежно попадет на сингулярную пространственноподобную границу (черную дыру) за конечное собственное время. При этом нет никакой возможности пересечь горизонт. Внешние области I и III соответствуют двум разным вселенным для одной и той же черной дыры. Эти вселенные причинно не связаны между собой, поскольку их точки можно соединить только пространственноподобными кривыми. Если наблюдатель находится, например, во вселенной I, то у него есть две возможности: либо за конечное время упасть на черную дыру, либо жить бесконечно долго, поскольку бесконечность будущего (верхняя вершина конформного блока I) полна.

Для отрицательных значений q горизонты отсутствуют. Поэтому треугольные конформные блоки, показанные на рис. 30.4 вверху справа, представляют собой максимально продолженные решения. В этом пространстве-времени сингулярность кривизны располагается вдоль времениподобной границы, находящейся на конечном расстоянии и не окруженной горизонтом. Такие сингулярности в общей теории относительности называются *голыми*. Заметим, однако, что на голые сингулярности

попадают только свето- и пространственноподобные экстремали при конечном значении канонического параметра. Времениподобные экстремали отражаются вблизи сингулярности (см. рис. 30.2).

Глобальное решение для отрицательных q можно рассматривать как решение для положительных q (что соответствует интерпретации координаты q , как радиуса), но с отрицательной массой, $M < 0$. Эти решения рассматриваются в общей теории относительности, как нефизические.

30.5.2 Решение Рейснера–Нордстрема

Конформный множитель для решения Рейснера–Нордстрема [228, 229], описывающего заряженную черную дыру, имеет вид

$$\Phi = 1 - \frac{2M}{q} + \frac{Q^2}{q^2}, \quad 0 < Q < M, \quad (30.49)$$

где M и Q – масса и заряд черной дыры. Функцию Φ при заданных соотношениях между константами можно записать в виде

$$\Phi = \frac{(q - q_1)(q - q_2)}{q^2}, \quad (30.50)$$

где

$$q_{1,2} = M \pm \sqrt{M^2 - Q^2}.$$

Отсюда следует, что конформный множитель имеет полюс второго порядка при $q = 0$ и два положительных простых нуля (горизонта) $q_{1,2}$. Его поведение для положительных q показано на рис. 30.5 слева.

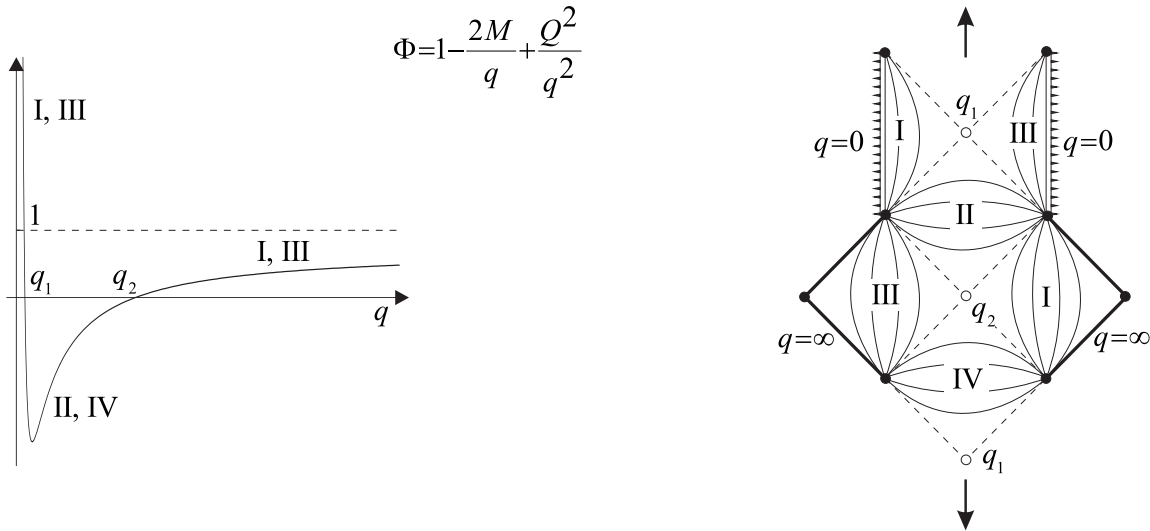


Рис. 30.5: Конформный множитель и фундаментальная область для решения Рейснера–Нордстрема. Стрелки показывают возможное продолжение или отождествление решения вдоль горизонтов.

Глобальная структура решения для положительных q , по сравнению с решением Шварцшильда, меняется качественно из-за наличия двух горизонтов. Каждому из интервалов $(0, q_1)$, (q_1, q_2) и (q_2, ∞) соответствует по два конформных блока. Поскольку внутри интервала $(0, \infty)$ конформный множитель меняет знак, то по правилу 6)

диаграмма Картера–Пенроуза, склеенная из шести конформных блоков, рис. 30.5, представляет собой фундаментальную область для решения Рейснера–Нордстрема. Ее граница состоит не только из сингулярных и бесконечно удаленных точек, но также включает отрезки, соответствующие горизонтам. В соответствии с правилом 7) ее можно периодически продолжить, склеивая одинаковые фундаментальные области вдоль горизонтов в направлениях, показанных стрелками. В этом случае мы получим универсальную накрывающую для решения Рейснера–Нордстрема. Другая возможность заключается в том, что после склеивания произвольного конечного числа фундаментальных областей, граничные точки на горизонтах снизу и сверху можно отождествить. Тогда глобальное решение с топологической точки зрения будет представлять собой цилиндр.

Сингулярности в решении Рейснера–Нордстрема являются голыми.

Скалярная кривизна (30.16) для решения Рейснера–Нордстрема непостоянна:

$$R = \frac{4M}{q^3} - \frac{6Q^2}{q^4}. \quad (30.51)$$

Это – двумерная скалярная кривизна поверхности, а не четырехмерного пространства-времени.

Левая ветвь конформного множителя $-\infty < q < 0$ для решения Рейснера–Нордстрема соответствует голой сингулярности и рассматриваться не будет, т.к. изменение порядка полюса не влияет на структуру глобального решения.

30.5.3 Экстремальная черная дыра

Экстремальная черная дыра возникает из решения Рейснера–Нордстрема (30.49) в том случае, когда заряд равен массе, $Q = M$. Соответствующий конформный множитель,

$$\Phi = \frac{(q - M)^2}{q^2}, \quad (30.52)$$

имеет полюс второго порядка при $q = 0$ и положительный нуль $q = M$ также второго порядка. Его поведение показано на рис. 30.6 слева. Левая ветвь $-\infty < q < 0$, как и

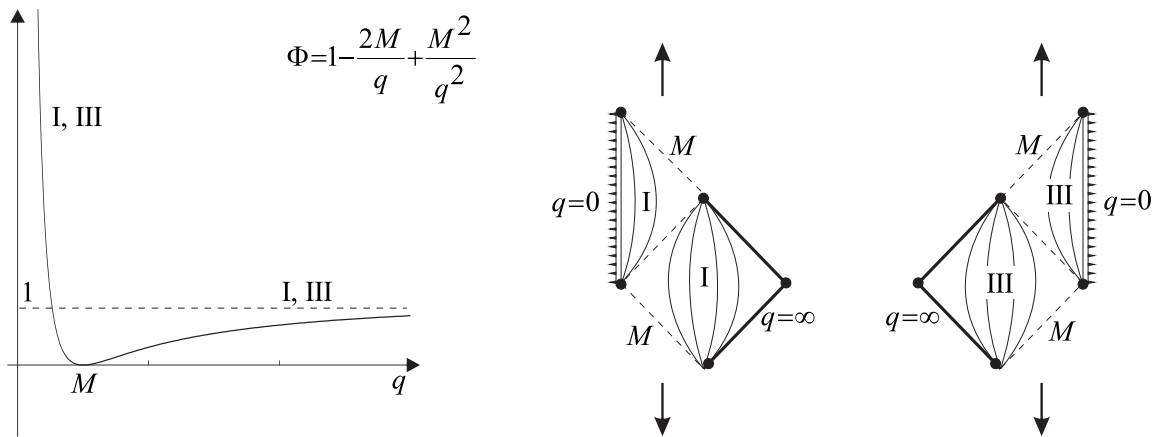


Рис. 30.6: Конформный множитель и две фундаментальные области для экстремальной черной дыры. Стрелками показаны направления, вдоль которых решения периодически продолжают.

в предыдущих случаях, описывает голую сингулярность.

Положительные значения q разбиваются на два интервала $(q_-, q_1) = (0, M)$ и $(q_1, q_+) = (M, \infty)$. Этим интервалам ставятся в соответствие два конформных блока и их пространственное отражение. Поскольку конформный множитель не меняет знак, то в соответствии с правилом 6) имеются две несвязные фундаментальные области, показанные на рис. 30.6. Вторая фундаментальная область, построенная из областей типа III, получается пространственным отражением из области типа I. Как и в случае решения Рейснера–Нордстрема границы фундаментальных областей включают горизонты и их можно либо склеивать до бесконечности, либо отождествить, что приводит к универсальной накрывающей и цилиндрам, соответственно.

30.5.4 Плоскость Минковского

Предыдущие примеры демонстрируют правила построения глобальных лоренцевых поверхностей непостоянной кривизны с одним вектором Киллинга. Следующие два примера показывают, как эти правила работают в случае поверхностей постоянной кривизны, которые имеют по три вектора Киллинга. Пример плоскости Минковского важен также потому, что будет использован в дальнейшем для доказательства дифференцируемости глобальных решений в седловых точках.

Для плоскости Минковского скалярная кривизна равна нулю, а конформный множитель является линейной функцией

$$\Phi = bq + c, \quad b, c = \text{const.} \quad (30.53)$$

При этом возможны два качественно отличных случая: $b = 0$ и $b \neq 0$.

При $b = 0$ метрика (30.1) является метрикой Минковского

$$ds^2 = c(d\tau^2 - d\sigma^2), \quad (30.54)$$

где, для определенности, будем считать, что $c > 0$. Конформный множитель $\Phi = c$ не имеет ни особенностей, ни нулей. При этом уравнение (30.2) имеет общее решение

$$q = \pm c\sigma,$$

где мы отбросили несущественную постоянную интегрирования, соответствующую сдвигу σ . Знак $\pm$ соответствует различной ориентации переменной q относительно пространственной координаты. Поскольку эта переменная не входит в метрику, то ее можно исключить из рассмотрения. Поэтому интервалу $q \in (-\infty, \infty)$ ставится в соответствие один квадратный конформный блок, показанный на рис. 30.7 слева.

Теперь рассмотрим случай $b \neq 0$. Не ограничивая общности, положим $b > 0$. Тогда конформный множитель (30.53) имеет один простой нуль в точке $q_1 = -c/b$. Интервалам $(-\infty, q_1)$ и (q_1, ∞) соответствует по два квадратных конформных блока, которые склеиваются вдоль горизонтов, как показано на рис. 30.7 справа.

Покажем, как связано это глобальное решение с предыдущим представлением плоскости Минковского в виде одного конформного блока. Не ограничивая общности, положим $c = 0$, т.е.

$$ds^2 = b|q|(d\tau^2 - d\sigma^2), \quad b > 0. \quad (30.55)$$

Этого всегда можно добиться путем сдвига $q \rightarrow q + \text{const}$, который не влияет на глобальную структуру решения. Рассмотрим по очереди все четыре области.

Область I. Уравнение (30.2) для $\Phi = bq$ имеем следующее решение

$$q = e^{b\sigma} \quad (30.56)$$

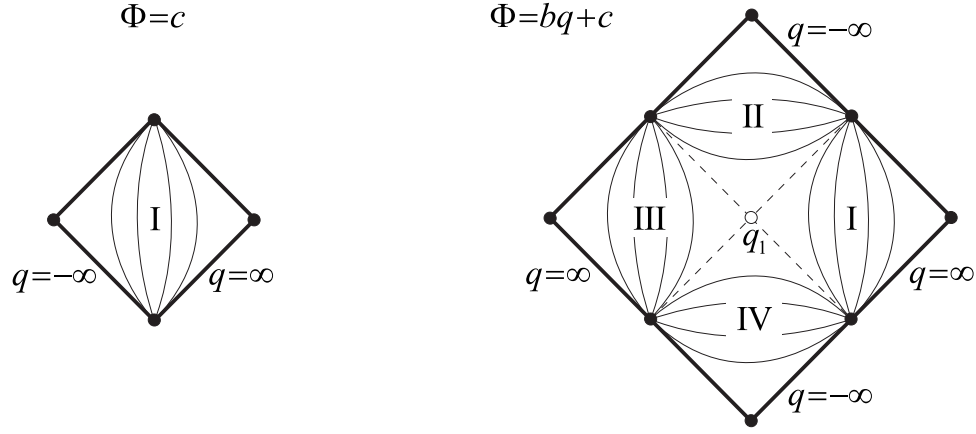


Рис. 30.7: Два представления плоскости Минковского при отсутствии и наличии одного горизонта. В последнем случае плоскость Минковского склеивается из четырех конформных блоков.

с точностью до сдвига σ . Соответствующая метрика статична,

$$ds^2 = b e^{b\sigma} (d\tau^2 - d\sigma^2), \quad \tau, \sigma \in \mathbb{R}^2, \quad (30.57)$$

и определена на всей плоскости. Переход к конусным координатам (30.23) и конформное преобразование

$$u = \frac{2}{b} e^{\frac{b\xi}{2}} > 0, \quad v = -\frac{2}{b} e^{-\frac{b\eta}{2}} < 0 \quad (30.58)$$

приводит к метрике Минковского

$$ds^2 = b du dv, \quad (30.59)$$

определенной в первом квадранте.

Область III. Для переменной q и метрики справедливы равенства:

$$q = e^{-b\sigma}, \quad (30.60)$$

$$ds^2 = b e^{-b\sigma} (d\tau^2 - d\sigma^2), \quad \tau, \sigma \in \mathbb{R}^2. \quad (30.61)$$

Конформное преобразование, приводящее к метрике Минковского (30.59), имеет вид

$$u = -\frac{2}{b} e^{-\frac{b\xi}{2}} < 0, \quad v = \frac{2}{b} e^{\frac{b\eta}{2}} > 0, \quad (30.62)$$

что соответствует третьему квадранту плоскости Минковского.

Область II. Эта область однородна:

$$q = -e^{b\tau}, \quad (30.63)$$

$$ds^2 = b e^{b\tau} (d\tau^2 - d\sigma^2), \quad \tau, \sigma \in \mathbb{R}^2. \quad (30.64)$$

Конформное преобразование, приводящее к метрике Минковского (30.59), имеет вид

$$u = \frac{2}{b} e^{\frac{b\xi}{2}} > 0, \quad v = \frac{2}{b} e^{\frac{b\eta}{2}} > 0, \quad (30.65)$$

что соответствует второму квадранту.

Область IV. Аналогично, для переменной q и метрики справедливы равенства

$$q = -e^{-b\tau}, \quad (30.66)$$

$$ds^2 = b e^{-b\tau} (d\tau^2 - d\sigma^2), \quad \tau, \sigma \in \mathbb{R}^2. \quad (30.67)$$

Конформное преобразование, приводящее к метрике Минковского имеет вид

$$u = -\frac{2}{b} e^{-\frac{b\xi}{2}} < 0, \quad v = -\frac{2}{b} e^{-\frac{b\eta}{2}} < 0, \quad (30.68)$$

что соответствует четвертому квадранту.

Для всех четырех областей горизонт $q_1 = 0$ соответствует координатным прямым $u = 0$ и $v = 0$. Простые вычисления показывают, что во всех областях

$$q = -\frac{b^2}{4}(t^2 - x^2), \quad (30.69)$$

где

$$u = t + x, \quad v = t - x.$$

Эти координаты дают простейший пример координат Крускала–Секереша для метрики (30.55). То есть с точностью до знака переменная q представляет собой квадрат радиуса гиперболической полярной системы координат (1.112) на плоскости Минковского. Для гиперболического полярного угла в статичных областях

$$\varphi = \operatorname{arcth} \left(\frac{t}{x} \right),$$

справедливы равенства:

$$\text{I: } \varphi = \frac{b}{2}\tau, \quad \text{III: } \varphi = -\frac{b}{2}\tau.$$

В однородных областях

$$\varphi = \operatorname{arcth} \left(-\frac{x}{t} \right),$$

и

$$\text{II: } \varphi = -\frac{b}{2}\sigma, \quad \text{IV: } \varphi = \frac{b}{2}\sigma.$$

Поскольку конформные преобразования (30.58), (30.62), (30.65) и (30.68) во всех четырех областях дают одну и ту же метрику Минковского (30.59), но в разных квадрантах, то это доказывает гладкость метрики на горизонтах и в седловой точке для переменной q (30.69).

Таким образом, линейный конформный множитель (30.53) при $b \neq 0$ описывает плоскость Минковского, рис. 30.7 справа, в гиперболической полярной системе координат.

30.5.5 Поверхности постоянной кривизны

Рассмотрим поверхности постоянной ненулевой кривизны, для которых конформный множитель является квадратичным полиномом (30.17). Для определенности, рассмотрим поверхности положительной скалярной (отрицательной гауссовой) кривизны, $a > 0$. Поверхности отрицательной скалярной кривизны получаются из поверхностей положительной кривизны преобразованием $\tau \leftrightarrow \sigma$, т.е. поворотом всех диаграмм на угол $\pi/2$.

В зависимости от значений постоянных a, b, c уравнение $\Phi(q) = 0$ может не иметь решений или иметь один или два корня. Рассмотрим последовательно все три случая, которые приводят к разным метрикам и различным диаграммам Картера–Пенроуза для поверхностей постоянной кривизны. Не ограничивая общности, положим $b = 0$, чего всегда можно добиться сдвигом переменной q .

Отсутствие горизонта. Рассмотрим конформный множитель вида

$$\Phi = -(aq^2 + c), \quad a > 0, \quad c > 0. \quad (30.70)$$

Поскольку $\Phi < 0$, а сингулярности и нули отсутствуют, то глобальное решение соответствует бесконечному интервалу $q \in (-\infty, \infty)$. Оно представляется однородным конформным блоком типа линзы II или IV, показанными на рис. 30.8 слева. В обла-

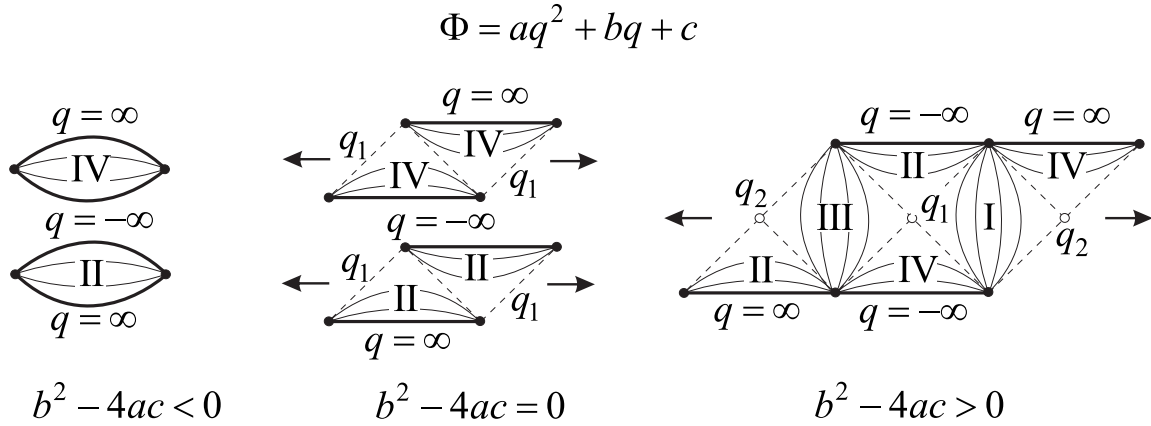


Рис. 30.8: Диаграммы Картера–Пенроуза для поверхностей постоянной положительной скалярной кривизны. Возможно три случая: отсутствие, один или два горизонта.

сти IV уравнение (30.2) для конформного множителя (30.70) приводит к равенству

$$q = \sqrt{\frac{c}{a}} \operatorname{tg}(\sqrt{ac} \tau) \quad (30.71)$$

с точностью до сдвига τ . Временная координата меняется в конечном интервале

$$-\frac{\pi}{2} < \sqrt{ac} \tau < \frac{\pi}{2},$$

что соответствует диаграмме типа линзы. Метрика в плоскости τ, σ имеет вид

$$ds^2 = \frac{c}{\cos^2(\sqrt{ac} \tau)} (d\tau^2 - d\sigma^2). \quad (30.72)$$

Эта метрика является глобальной и определена на всей универсальной накрывающей постоянной кривизны.

В области II метрика имеет такой же вид, меняется только ориентация переменной q по отношению к τ .

Один горизонт. Пусть конформный множитель имеет вид

$$\Phi = -aq^2, \quad a > 0. \quad (30.73)$$

В точке $q_1 = 0$ он имеет нуль второго порядка, соответствующий горизонту. В области IV для обоих интервалов $q \in (-\infty, 0)$ и $q \in (0, \infty)$ получаем следующую метрику

$$q = -\frac{1}{a\tau}, \quad (30.74)$$

$$ds^2 = \frac{1}{a\tau^2}(d\tau^2 - d\sigma^2). \quad (30.75)$$

Фундаментальная область в этом случае состоит из двух треугольных конформных блоков, показанных на рис. 30.8 в центре. Чтобы получить универсальную накрывающую поверхность, диффеоморфную диаграмме типа линзы из предыдущего случая, фундаментальную область необходимо продолжить в обе стороны. Если отождествить левую и правую границы фундаментальной области, то получим поверхность, диффеоморфную однополостному гиперболоиду, рассмотренному в разделе 29.3.

Два горизонта. Пусть конформный множитель имеет вид

$$\Phi = -(aq^2 + c), \quad a > 0, \quad c < 0. \quad (30.76)$$

Он имеет два простых нуля в точках $q_{1,2} = \pm\sqrt{-c/a}$, которые соответствуют горизонтам. Поэтому трем интервалам $(-\infty, q_1)$, (q_1, q_2) и (q_2, ∞) соответствует шесть конформных блоков. Из симметрии конформного множителя $q \mapsto -q$ следует, что однородные конформные блоки II, IV и стационарные конформные блоки I, III связаны преобразованием $q \mapsto -q$. Для однородной области типа IV получаем метрику в виде

$$q = -\sqrt{-\frac{c}{a}} \operatorname{cth}(\sqrt{-ac}\tau), \quad (30.77)$$

$$ds^2 = \frac{-c}{\operatorname{sh}^2(\sqrt{-ac}\tau)}(d\tau^2 - d\sigma^2). \quad (30.78)$$

Для стационарной области I метрика имеет вид

$$q = \sqrt{-\frac{c}{a}} \operatorname{th}(\sqrt{-ac}\sigma), \quad (30.79)$$

$$ds^2 = \frac{-c}{\operatorname{ch}^2(\sqrt{-ac}\sigma)}(d\tau^2 - d\sigma^2). \quad (30.80)$$

На рис. 30.8 справа изображены конформные блоки и их склейка (диаграмма Картера–Пенроуза) для поверхностей постоянной кривизны.

Различные виды метрики при отсутствии или наличии горизонтов обусловлены различным выбором систем координат. Выпишем формулы преобразования координат, которые связывают между собой метрики в отсутствие горизонта, а также с одним и двумя горизонтами. Поскольку все метрики являются вейлевски плоскими, то связь между ними осуществляется конформным преобразованием координат. Запишем метрику с одним горизонтом (30.75) в конусных координатах (30.23) и совершим конформное преобразование $\xi \mapsto u(\xi)$, $\eta \mapsto v(\eta)$, тогда метрика примет вид

$$ds^2 = \frac{1}{a\tau^2}(d\tau^2 - d\sigma^2) = \frac{4}{a(\xi + \eta)^2} \frac{d\xi}{du} \frac{d\eta}{dv} du dv. \quad (30.81)$$

Преобразование конусных координат

$$\xi = \frac{2}{\sqrt{ac}} \operatorname{tg} \left(\frac{\sqrt{ac}}{2} u \right), \quad \eta = -\frac{2}{\sqrt{ac}} \operatorname{ctg} \left(\frac{\sqrt{ac}}{2} v \right) \quad (30.82)$$

дает метрику (30.72), которая соответствует отсутствию горизонта. Преобразования конусных координат

$$\xi = \frac{2}{\sqrt{-ac}} \operatorname{th} \left(\frac{\sqrt{-ac}}{2} u \right), \quad \eta = \frac{2}{\sqrt{-ac}} \operatorname{th} \left(\frac{\sqrt{-ac}}{2} v \right) \quad (30.83)$$

и

$$\xi = \frac{2}{\sqrt{-ac}} \operatorname{th} \left(\frac{\sqrt{-ac}}{2} u \right), \quad \eta = -\frac{2}{\sqrt{-ac}} \operatorname{cth} \left(\frac{\sqrt{-ac}}{2} v \right) \quad (30.84)$$

приводят к однородной (30.78) и статической (30.80) метрикам при наличии двух горизонтов.

Рассмотренные примеры показывают, что одной и той же поверхности могут соответствовать различные вейлевски плоские метрики, связанные конформным преобразованием координат. Однако независимо от выбора исходной метрики метод конформных блоков позволяет восстановить всю поверхность единственным образом.

30.6 Координаты Эддингтона–Финкельстейна

В теореме 30.4.1 был описан конструктивный метод конформных блоков построения максимально продолженных поверхностей с одним вектором Киллинга. Дифференцируемость метрики внутри каждого конформного блока следует из построения. Конформные блоки склеиваются вдоль горизонтов q_j (нулей конформного множителя $\Phi(q_j) = 0$) с показателями асимптотики $m = 1$ (простой нуль) и $m \geq 2$. В настоящем разделе мы докажем дифференцируемость метрики на горизонтах после склейки, записав ее в новой системе координат, которая покрывает горизонты.

Пусть внутри интервала (q_-, q_+) есть нуль q_j нечетного порядка. Для определенности, будем считать, что в соседних интервалах (q_{j-1}, q_j) и (q_j, q_{j+1}) конформный множитель положителен и отрицателен, соответственно. В соответствии с правилом 4) стационарный конформный блок типа I должен быть склеен с двумя однородными блоками типа II и IV. Соответствующие горизонты можно накрыть *координатами Эддингтона–Финкельстейна* [230, 231], которые вводятся следующим образом.

Склейка I–II. Для доказательства дифференцируемости метрики в склеенных областях I и II, включая горизонт, введем координаты Эддингтона–Финкельстейна $\tau, \sigma \mapsto \xi, q$. Эти координаты полностью покрывают указанные области, при этом функции перехода внутри областей I и II принадлежат классу $\mathcal{C}^{l+1}$, а метрика в новых координатах – классу $\mathcal{C}^l$.

В области I метрика для $q \in (q_{j-1}, q_j)$ имеет вид

$$I: \quad ds^2 = \Phi(q)(d\tau^2 - d\sigma^2), \quad \frac{dq}{d\sigma} = \Phi(q) > 0. \quad (30.85)$$

перейдем от декартовых координат τ, σ к координатам Эддингтона–Финкельстейна ξ, q :

$$\tau := \xi - \int^q \frac{dr}{\Phi(r)}, \quad \sigma := \int^q \frac{dr}{\Phi(r)}. \quad (30.86)$$

Интегралы в этих формулах на горизонте $q \rightarrow q_j$ расходятся, однако это не важно, т.к. преобразование координат (30.86) рассматривается только во внутренних точках области I. Координата $\xi = \tau + \sigma$ – это обычная “светоподобная” координата на

плоскости τ, σ , а в качестве второй координаты выбрана сама переменная q . Эпитет светоподобная взят в кавычки, т.к. вскоре мы увидим, что координатные линии ξ в действительности не являются светоподобными.

Очевидно, что во внутренних точках функции перехода принадлежат классу $\mathcal{C}^{l+1}$.

Подставляя выражения для дифференциалов

$$d\tau = d\xi - \frac{dq}{\Phi}, \quad d\sigma = \frac{dq}{\Phi}, \quad (30.87)$$

в выражение для метрики (30.85), получим квадратичную форму

$$ds^2 = \Phi d\xi^2 - 2dq d\xi. \quad (30.88)$$

Определитель этой метрики равен $\det g_{\alpha\beta} = -1$, и поэтому метрика (30.88) определена для всех $\xi \in (-\infty, \infty)$ и, что важно, при всех $q \in (q_-, q_+)$, а не только при $q \in (q_{j-1}, q_j)$. Вектор ∂_q , касательный к координатным линиям q , имеет нулевую длину, $(\partial_q, \partial_q) = 0$, и, следовательно, координатные линии q светоподобны. Вектор ∂_ξ , касательный к координатным линиям ξ , не является светоподобным. Его квадрат равен $(\partial_\xi, \partial_\xi) = \Phi$. Он положителен в статичных областях и стремится к нулю при приближении к горизонту.

Метрика (30.88) определена в большей области, чем исходная метрика (30.1). Эта область – диагональная цепочка конформных блоков, идущих сверху слева вниз направо и склеенных вдоль координаты q (см. рис. 30.9 справа). При этом гладкость метрики (30.88) совпадает с гладкостью конформного множителя всюду, включая горизонты.

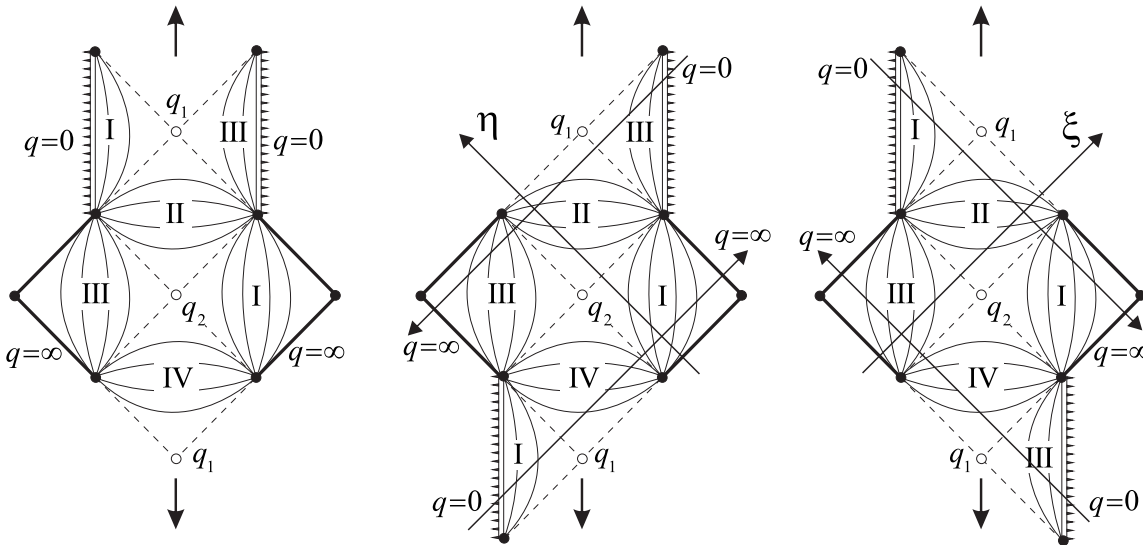


Рис. 30.9: Эквивалентные диаграммы Картера–Пенроуза для решения Рейснера–Нордстрема. Координаты Эддингтона–Финкельштейна покрывают диагональные цепочки конформных блоков, идущие снизу слева направо (в центре) или снизу справа налево (справа).

Отметим, что световые конусы в исходных координатах τ, σ были такими же, как и на плоскости Минковского, т.е. их образующие – это прямые линии, идущие под углом $\pm 45^\circ$. В координатах ξ, q так, как они показаны на рисунке, образующие

световых конусов будут отличаться. Одна из образующих будет параллельна оси q , а вторая будет иметь наклон

$$\frac{d\xi}{dq} = \frac{2}{1 - \frac{2M}{q}}.$$

В однородной области для $q \in (q_j, q_{j+1})$ типа

$$\text{II} : \quad ds^2 = -\Phi(q)(d\tau^2 - dq^2), \quad \frac{dq}{d\tau} = \Phi(q) < 0, \quad (30.89)$$

перейдем к координатам Эддингтона–Финкельстейна ξ, q :

$$\tau := \int^q \frac{dr}{\Phi(r)}, \quad \sigma := \xi - \int^q \frac{dr}{\Phi(r)}. \quad (30.90)$$

Преобразование координат в области II отличается от преобразования (30.86). Тем не менее метрика (30.89) снова принимает вид (30.88). Это означает, что координаты Эддингтона–Финкельстейна и заданная в них метрика (30.88) изометрично покрывают области I и II и, кроме того, метрика определена на их объединении и горизонте. Это доказывает, что на горизонте дифференцируемость метрики на продолженной поверхности совпадает с дифференцируемостью конформного множителя.

Можно также проверить, что горизонт q_j в координатах Эддингтона–Финкельстейна сам является светоподобной экстремалью.

Метрика (30.88) в координатах Эддингтона–Финкельстейна имеет вектор Киллинга ∂_ξ . Квадрат его длины равен конформному множителю $(\partial_\xi, \partial_\xi) = \Phi(q)$. Поэтому на горизонте длина вектора Киллинга равна нулю.

Склейка III–IV. Области III–IV склеиваются аналогично областям I и II. Приведем явные формулы перехода к координатам Эддингтона–Финкельстейна:

$$\text{III} : \quad \tau := \xi + \int^q \frac{dr}{\Phi(r)}, \quad \sigma := - \int^q \frac{dr}{\Phi(r)}. \quad (30.91)$$

и

$$\text{IV} : \quad \tau := - \int^q \frac{dr}{\Phi(r)}, \quad \sigma := \xi + \int^q \frac{dr}{\Phi(r)}. \quad (30.92)$$

При этом в обеих областях метрика имеет одинаковый вид

$$ds^2 = \Phi d\xi^2 + 2dq d\xi. \quad (30.93)$$

Эта метрика определена на всей диагональной цепочке конформных блоков $q \in (q_-, q_+)$, идущей снизу справа наверх налево (см. рис. 30.9 справа).

Между собой метрики (30.88) и (30.93) связаны преобразованием $\xi \mapsto -\xi$.

Склейка I–IV. В области I перейдем к координатам Эддингтона–Финкельстейна η, q :

$$\text{I} : \quad \tau := \eta + \int^q \frac{dr}{\Phi(r)}, \quad \sigma := \int^q \frac{dr}{\Phi(r)}, \quad (30.94)$$

где $\eta = \tau - \sigma$ – светоподобная координата. В новых координатах метрика принимает вид

$$ds^2 = \Phi d\eta^2 + 2dq d\eta. \quad (30.95)$$

Эта метрика определена на всей диагональной цепочке конформных блоков $q \in (q_-, q_+)$, идущей снизу слева наверх направо (см. рис. 30.9 в центре).

К этой же метрике в области IV приводят координаты

$$\text{IV :} \quad \tau := - \int^q \frac{dr}{\Phi(r)}, \quad \sigma := -\eta - \int^q \frac{dr}{\Phi(r)}. \quad (30.96)$$

Аналогичные координаты вводятся при склейке областей II и III, при этом соответствующая метрика связана с метрикой (30.95) преобразованием $\eta \mapsto -\eta$.

Если горизонт q_j имеет четный порядок, и в обоих интервалах (q_{j-1}, q_j) и (q_j, q_{j+1}) конформный множитель, например, положителен, то, согласно правилу 4), склеиваются только однотипные блоки. Для доказательства дифференцируемости склейки на блоках типа I и III вводятся координаты (30.86) и (30.91), соответственно.

Таким образом доказывается дифференцируемость метрики и, следовательно, символов Кристоффеля и тензора кривизны на всех горизонтах. Метрика в координатах Эддингтона–Финкельштейна (30.88) при $q \in (q_-, q_+)$ покрывает всю диагональную цепочку конформных блоков типа I и II для фундаментальной области. Для этой цепочки блоков координата q убывает с возрастанием времени τ . Цепочка блоков представляет собой многообразие класса $\mathcal{C}^{l+1}$, т.к. функции перехода (30.86), (30.90) принадлежат этому классу. Метрика (30.93) покрывает параллельную цепочку блоков типа III и IV. Метрика в виде (30.95) покрывает перпендикулярную диагональную цепочку конформных блоков типа I и IV. Внутри каждой цепочки метрика в координатах Эддингтона–Финкельштейна невырождена и класса $\mathcal{C}^l$.

Если внутри интервала (q_-, q_+) , который соответствует глобальному решению, имеется n нулей, то для получения фундаментальной области необходимо склеить $2(n+1)$ конформных блоков (удвоенное число интервалов (q_j, q_{j+1})). Их можно склеить в две диагональных цепочки блоков, например, идущих снизу слева направо, как показано на рис. 30.9 в центре. В каждой цепочке склеиваются между собой либо только блоки I, IV, либо только блоки II, III. По построению, каждая цепочка является связным многообразием. Если внутри интервала есть хотя бы один нуль нечетного порядка, то вокруг него происходит склейка цепочек, как для решения Шварцшильда. При этом достаточно одного нуля нечетного порядка, чтобы превратить фундаментальную область в связное многообразие. Если же конформный множитель внутри интервала (q_-, q_+) знак не меняет, то диагональные цепочки блоков представляет собой две компоненты связности фундаментальной области. Это доказывает утверждение в правиле 5).

Нетрудно проверить, что переход от декартовых координат τ, σ к координатам Эддингтона–Финкельштейна вырожден на всех горизонтах.

Координаты Эддингтона–Финкельштейна являются естественными в следующем смысле. Прежде всего отметим, что продолжать локальное решение надо по переменной q , т.к. именно она определяет полноту многообразия и сингулярности кривизны. Поскольку для стационарного решения переменная q зависит от σ , то вместо σ можно выбрать само q (30.86), (30.91), (30.94), при этом связь координат определяется уравнением (30.2) взаимно однозначно для каждого типа области с точностью до несущественного сдвига σ . После этого вместо временной координаты вводится светоподобная координата (30.23). При этом существует две возможности, и обе они реализуются, задавая координаты на взаимно перпендикулярных диагональных цепочках конформных блоков. Аналогично вводятся координаты Эддингтона–Финкельштейна на однородных блоках.

30.7 Дифференцируемость метрики в седловой точке

Координаты Эддингтона–Финкельштейна не покрывают седловых точек, расположенных в местах пересечения горизонтов. Седловые точки соответствуют нулям конформного множителя $\Phi(q_j) = 0$, расположенным в конечных точках, $|q_j| < \infty$ с показателем степени $m = 1$ или $m \geq 2$ (см. рис. 30.3). Если точка q_j является нулем кратности два или больше, то седловая точка полна, и доказывать ничего не надо. Если точка q_j – простой нуль, то можно ввести координаты, покрывающие окрестность седловой точки. Поскольку поведение конформного множителя вблизи простого нуля такое же, как и на плоскости Минковского, то перейдем к светоподобным координатам Крускала–Секереша $\tau, \sigma \mapsto u, v$. Явные формулы преобразования в областях I–IV имеют вид (30.58), (30.65), (30.62) и (30.68), соответственно. При этом мы положим $b := \Phi'(q_j)$ (см. раздел 30.5.4).

Рассмотрим область I более подробно. Метрика имеет вид

$$ds^2 = \Phi(d\tau^2 - d\sigma^2) = \Phi d\xi d\eta,$$

где ξ, η – светоподобные координаты (30.23). Преобразуем светоподобные координаты $\xi, \eta \mapsto u, v$ по формулам (30.58). Тогда метрика примет вид

$$ds^2 = \Phi e^{-b\sigma} du dv = \Phi \exp \left(-b \int^q \frac{dp}{\Phi(p)} \right) du dv,$$

где мы учли связь σ с q , которая дается уравнением

$$\frac{dq}{d\sigma} = \Phi.$$

Проведем аналогичное преобразование координат во всех четырех квадрантах и везде положим $b := \Phi'(q_j)$. Тогда во всех четырех квадрантах метрика 30.1) примет одинаковый вид

$$ds^2 = |\Phi| \exp \left(- \int^q dp \frac{\Phi'(q_j)}{\Phi(p)} \right) du dv = \exp \left(\int^q dp \frac{\Phi'(p) - \Phi'(q_j)}{\Phi(p)} \right) du dv. \quad (30.97)$$

Как и в случае плоскости Минковского координаты u, v покрывают окрестность седловой точки, состоящую из областей всех четырех типов, склеенных вдоль горизонтов.

Вблизи седловой точки при $m = 1$ справедливо разложение

$$\Phi(p) = \Phi_1(p - q_j) + \frac{1}{2}\Phi_2(p - q_j)^2 + o((p - q_j)^2), \quad \Phi_{1,2} = \text{const}, \quad \Phi_1 \neq 0.$$

Поэтому

$$\lim_{p \rightarrow q_j} \frac{\Phi'(p) - \Phi'(q_j)}{\Phi(p)} = \frac{\Phi_2}{\Phi_1}.$$

Поскольку в седловой точке предел подынтегрального выражения конечен, то интеграл является достаточно гладкой функцией. Это доказывает невырожденность и дифференцируемость метрики в седловой точке, которая совпадает с дифференцируемостью конформного множителя.

Очевидно, что вблизи каждой седловой точки первого порядка можно ввести координаты Крускала–Секереша. Поскольку функции перехода к координатам Крускала–Секереша класса C^∞ , то класс гладкости универсальной накрывающей поверхности определяется гладкостью функций перехода к координатам Эддингтона–Финкельштейна, C^{l+1} . Тем самым доказательство основной теоремы 30.4.1 завершено.

Таким образом, мы показали, что локальное решение вида (30.1) имеет дифференцируемое максимальное продолжение. Описанная процедура является однозначной. Однако, это не доказывает единственность продолжения. Дело в том, что преобразование координат к координатам Эддингтона–Финкельштейна и Крускала–Секереша вырождено на горизонтах. Поэтому нельзя утверждать, что указанное продолжение решения вдоль экстремалей является единственным.

Метод построения глобальных решений для двумерных метрик лоренцевой сигнатуры, рассмотренный в настоящей главе, является простым и конструктивным. Если уравнения теории гравитации приводят к метрике (30.1) с некоторым конформным множителем, то для построения глобального решения достаточно проанализировать его нули и особенности. При этом восстанавливается универсальная накрывающая поверхность. Остальные глобальные решения являются факторпространствами универсальной накрывающей по группе преобразований, действующей собственно разрывно и свободно. Выше мы отметили возможность построения решений с нетривиальной фундаментальной группой путем отождествления элементов границы (горизонтов) фундаментальной области. В общем случае нахождение групп преобразований, действующих собственно разрывно и свободно, зависит от конкретного вида конформного множителя и является предметом самостоятельного исследования.

Глава 31

Римановы поверхности с одним вектором Киллинга

Римановы поверхности с одним вектором Киллинга включают в себя как частный случай римановы поверхности постоянной кривизны, которые имеют по три вектора Киллинга. Этот более широкий класс поверхностей имеет важное значение для общей теории относительности, т.к. позволяет дать физическую интерпретацию многим точным решениям уравнений Эйнштейна.

31.1 Локальный вид римановой метрики

Чтобы преобразовать лоренцеву метрику (30.1) к метрике евклидовой сигнатуры, совершим поворот в комплексной плоскости той из координат τ или σ , от которой конформный множитель не зависит. Поскольку метрика вида (30.1), по предположению, возникает в результате решения системы уравнений движения для некоторой модели гравитации, то соответствующая риманова метрика также будет являться локальным решением той же системы уравнений после поворота в комплексной плоскости.

В явном виде комплексный поворот задается следующим образом. Введем новую координату $\tau := i\rho$ в статичных областях I, III и $\sigma := i\rho$ – в однородных областях II, IV. В результате получим метрику

$$ds^2 = -\Phi(q)(d\sigma^2 + d\rho^2), \quad (31.1)$$

где в областях II и IV мы переобозначили координату $\tau \mapsto \sigma$. Знак конформного множителя $\Phi(q)$ не фиксирован, т.е. рассматриваются как положительно, так и отрицательно определенные метрики. После комплексного поворота аргумент конформного множителя q зависит только от σ , и эта связь задается обыкновенным дифференциальным уравнением

$$\left| \frac{dq}{d\sigma} \right| = |\Phi(q)|. \quad (31.2)$$

Знаки модуля в этом уравнении раскрываются следующим образом:

$$\begin{array}{lll} \text{I : } & \Phi > 0, & dq/d\sigma > 0, & \text{sign } g_{\alpha\beta} = (--), \\ \text{II : } & \Phi < 0, & dq/d\sigma < 0, & \text{sign } g_{\alpha\beta} = (++) , \\ \text{III : } & \Phi > 0, & dq/d\sigma < 0, & \text{sign } g_{\alpha\beta} = (--), \\ \text{IV : } & \Phi < 0, & dq/d\sigma > 0, & \text{sign } g_{\alpha\beta} = (++) , \end{array} \quad (31.3)$$

где в последней колонке показана сигнатура метрики (31.1).

Как и случае метрики лоренцевой сигнатуры, у нас есть четыре области: две области с положительно определенной и две области с отрицательно определенной метрикой. Метрики в областях I и III, а также II и IV по существу одинаковы, т.к. связаны между собой простым преобразованием $\sigma \mapsto -\sigma$.

Метрика вида (31.1) обладает по крайней мере одним вектором Киллинга $K = \partial_\rho$, квадрат длины которого равен $(\partial_\rho, \partial_\rho) = -\Phi(q)$.

Риманова метрика, заданная уравнениями (31.1), (31.2), и будет предметом исследования в настоящем разделе. Мы допускаем, что конформный множитель может обращаться в нуль и иметь особенности в конечном числе точек q_i , $i = 1, \dots, k$. В эту последовательность включены также бесконечно удаленные точки $q_1 = -\infty$ и $q_k = \infty$. Таким образом, вещественная ось q разбивается точками q_i на интервалы, внутри которых конформный множитель либо строго больше, либо строго меньше нуля. Будем считать, что вблизи граничных точек q_i конформный множитель ведет себя, как и в лоренцевом случае, степенным образом (30.4), (30.5). При конечных значениях q_i показатель степени не равен нулю, $m \neq 0$, т.к. конформный множитель в этой точке, по предположению, либо равен нулю, либо сингулярен.

Символы Кристоффеля имеют следующие нетривиальные компоненты:

$$\text{I, II:} \quad \Gamma_{\sigma\sigma}^\sigma = \Gamma_{\sigma\rho}^\rho = \Gamma_{\rho\sigma}^\rho = -\Gamma_{\rho\rho}^\sigma = \frac{1}{2}\Phi', \quad (31.4)$$

$$\text{III, IV:} \quad \Gamma_{\sigma\sigma}^\sigma = \Gamma_{\sigma\rho}^\rho = \Gamma_{\rho\sigma}^\rho = -\Gamma_{\rho\rho}^\sigma = -\frac{1}{2}\Phi', \quad (31.5)$$

где штрих обозначает дифференцирование по аргументу, $\Phi' := d\Phi/dq$, и никакого суммирования по индексам ρ и σ не проводится. Тензор кривизны одинаков для всех областей и имеет только одну независимую компоненту

$$R_{\sigma\rho\sigma}^\rho = \frac{1}{2}\Phi\Phi'' \quad (31.6)$$

Тензор Риччи диагонален. Его ненулевые компоненты и скалярная кривизна равны

$$R_{\sigma\sigma} = R_{\rho\rho} = \frac{1}{2}\Phi\Phi'', \quad (31.7)$$

$$R = -\Phi''. \quad (31.8)$$

Отсюда следует, что для поверхностей постоянной кривизны, $R = \text{const}$, как и в лоренцевом случае, конформный множитель Φ является квадратичным полиномом.

Из уравнения (31.8) следует, что скалярная кривизна сингулярна вблизи граничной точки q_i при тех же показателях (30.18), (30.19), что и в лоренцевом случае. В граничных точках $q_i = \pm\infty$ скалярная кривизна стремится к отличной от нуля постоянной при $m = 2$ и к нулю при $m < 2$. Отметим, что ненулевое значение кривизны в конечной точке $|q_i| < \infty$ может возникнуть и при $m = 1$ за счет поправок следующего порядка в разложении (30.4).

Значение переменной q и, значит, скалярной кривизны постоянно вдоль траекторий Киллинга $\sigma = \text{const}$. В соответствии с определением областей (31.3) q монотонно возрастает по σ в областях I, IV и монотонно убывает в областях II, III.

Область определения метрики (31.1) на плоскости σ, ρ зависит от вида конформного множителя. Координата $\rho \in \mathbb{R}$ пробегает всю вещественную прямую, поскольку от нее ничего не зависит, а область изменения координаты σ определяется уравнением (31.2). Каждому интервалу переменной $q \in (q_i, q_{i+1})$ соответствует конечный,

полубесконечный или бесконечный интервал координаты σ в зависимости от сходимости или расходимости интеграла

$$\sigma_{i,i+1} \sim \int^{q_i, q_{i+1}} \frac{dq}{\Phi(q)} \quad (31.9)$$

в граничных точках. Сходимость интеграла определяется показателем степени m :

$$\begin{aligned} |q_i| < \infty : & \begin{cases} m < 1, & \text{сходится,} \\ m \geq 1, & \text{расходится,} \end{cases} \\ |q_i| = \infty : & \begin{cases} m \leq 1, & \text{расходится,} \\ m > 1, & \text{сходится,} \end{cases} \end{aligned} \quad (31.10)$$

Если на обоих концах интервала (q_i, q_{i+1}) интеграл расходится, то координата σ пробегает всю вещественную ось, $\sigma \in (-\infty, \infty)$, и метрика определена на всей плоскости $\sigma, \rho \in \mathbb{R}^2$. Если на одном из концов q_{i+1} или q_i интеграл сходится, то метрика определена на полуплоскости $\sigma \in (-\infty, \sigma_{i+1})$ или $\sigma \in (\sigma_i, \infty)$, соответственно. При этом выбор граничных точек σ_{i+1} и σ_i произволен, и, не ограничивая общности, можно положить $\sigma_{i,i+1} = 0$. Если на обоих концах интервала интеграл сходится, то локальное решение определено в полосе $\sigma \in (\sigma_i, \sigma_{i+1})$, при этом можно приравнять нулю только один из концов интервала.

В лоренцевом случае для построения максимально продолженных поверхностей каждому интервалу (q_i, q_{i+1}) ставятся в соответствие конформные блоки определенной формы, которые затем склеиваются между собой. При евклидовой сигнатуре метрики в этом нет необходимости, поскольку “светоподобные” экстремали с асимптотикой $\rho = \pm\sigma + \text{const}$, как будет показано ниже, в общем случае отсутствуют.

31.2 Экстремали

Для того чтобы описать максимально продолженную поверхность для римановой метрики (31.1) с одним вектором Киллинга, необходимо проанализировать поведение экстремалей $\{\zeta^\alpha(t)\} = \{\sigma(t), \rho(t)\}$, $t \in \mathbb{R}$, которые задаются системой нелинейных обыкновенных дифференциальных уравнений (30.25), где точка обозначает дифференцирование по каноническому параметру t .

Для определенности, рассмотрим область I. Используя выражение для символов Кристоффеля (31.4), получаем систему уравнений для экстремалей:

$$\ddot{\sigma} = \frac{1}{2}\Phi'(\dot{\rho}^2 - \dot{\sigma}^2), \quad (31.11)$$

$$\ddot{\rho} = -\Phi'\dot{\sigma}\dot{\rho}. \quad (31.12)$$

Эта система уравнений имеет два первых интеграла:

$$-\Phi(\dot{\sigma}^2 + \dot{\rho}^2) = C_0 = \text{const}, \quad (31.13)$$

$$-\Phi\dot{\rho} = C_1 = \text{const}. \quad (31.14)$$

Существование интеграла движения (31.13) позволяет выбрать длину экстремали в качестве канонического параметра. Второй интеграл (31.14) связан с наличием вектора Киллинга. На этом этапе проявляется важность наличия у метрики одного вектора Киллинга, т.к. существование двух интегралов позволяет проанализировать экстремали в общем случае, не конкретизируя вид конформного множителя $\Phi(q)$.

В системе уравнений для экстремалей (31.11), (31.12) конформный множитель Φ рассматривается, как сложная функция от σ : $\Phi = \Phi(q(\sigma))$, где $q(\sigma)$ – решение уравнения (31.2). Если форма экстремали $\sigma(\rho)$ найдена, то конформный множитель можно рассматривать так же, как сложную функцию от ρ вдоль каждой экстремали: $\Phi = \Phi(\sigma(\rho))$. В дальнейшем анализе это будет подразумеваться.

Теорема 31.2.1. *Любая экстремаль в области I принадлежит одному из следующих четырех классов.*

1. *Прямые экстремали вида (аналог светоподобных экстремалей)*

$$\rho = \pm\sigma + \text{const}, \quad (31.15)$$

существуют только для евклидовой метрики $\Phi = \text{const}$, при этом канонический параметр можно выбрать в виде $\sigma = t$.

2. *Экстремали общего вида, форма которых определена уравнением*

$$\frac{d\rho}{d\sigma} = \pm \frac{1}{\sqrt{-1 - C_2\Phi}}, \quad (31.16)$$

где $C_2 < 0$ – отрицательная постоянная. Соответствующий канонический параметр определяется любым из двух уравнений:

$$\dot{\sigma} = \pm \frac{\sqrt{-1 - C_2\Phi}}{\Phi}, \quad (31.17)$$

$$\dot{\rho} = \frac{1}{\Phi}. \quad (31.18)$$

При этом в уравнениях (31.16) и (31.17) знаки плюс или минус выбираются одновременно.

3. *Прямые вырожденные экстремали, параллельные оси σ и проходящие через каждую точку $\rho = \text{const}$. Канонический параметр определен уравнением*

$$\dot{\sigma} = \frac{1}{\sqrt{\Phi}}. \quad (31.19)$$

4. *Прямые вырожденные экстремали, параллельные оси ρ и проходящие через критические точки $\sigma_0 = \text{const}$, в которых*

$$\Phi'(\sigma_0) = 0. \quad (31.20)$$

Канонический параметр для них можно выбрать в виде

$$t = \rho. \quad (31.21)$$

Доказательство. Почти дословное повторение доказательства теоремы 30.3.1 в лоренцевом случае. $\square$

Постоянная C_2 , как и в лоренцевом случае, определяется интегралами (31.13) и (31.14):

$$C_2 := \frac{C_0}{C_1^2}. \quad (31.22)$$

Для области I $\Phi > 0$ и, следовательно, $C_0 < 0$ и $C_2 < 0$. Из уравнения (31.16) следует, что постоянная C_2 параметризует угол, под которым экстремаль общего вида проходит через заданную точку.

Подчеркнем качественное отличие в поведении экстремалей для метрик евклидовой и лоренцевой сигнатуры. Во-первых, для римановой метрики в общем случае отсутствует аналог светоподобных экстремалей. Это важно, т.к. в лоренцевом случае светоподобные экстремали были неполны на горизонтах и их необходимо было продолжить. Для римановых метрик эта проблема снимается. Во-вторых, уравнения для экстремалей общего вида (31.16) и (31.17) отличаются от соответствующих уравнений в лоренцевом случае знаком перед единицей, стоящей под знаком корня. Это небольшое на первый взгляд отличие приводит к тому, что у экстремалей общего вида уже нет светоподобной асимптотики при $q \rightarrow q_i$ вблизи нулей конформного множителя $\Phi(q_i) = 0$, которые определяют горизонты.

Качественное поведение экстремалей общего вида анализируется в общем случае, для чего достаточно знать поведение конформного множителя $\Phi(q)$ вблизи граничных точек q_i . Поскольку конформный множитель в области I положителен, то экстремали общего вида существуют только при отрицательных значениях постоянной C_2 , т.к. в противном случае правая часть уравнения (31.18) становится мнимой. При этом должно быть выполнено неравенство

$$\Phi(q) \geq -1/C_2.$$

При достаточно больших значениях модуля C_2 это неравенство определяет интервал изменения переменной $q \in (q', q'')$, где граничные точки q' и q'' определяются уравнением $\Phi(q) = -1/C_2$. Точкам q', q'' соответствуют некоторые точки σ' и σ'' . Экстремаль общего вида не может выйти за пределы полосы $\sigma \in (\sigma', \sigma'')$, $\rho \in (-\infty, \infty)$. Несложный анализ уравнения (31.18) показывает, что экстремали общего вида осциллируют между значениями σ' и σ'' , как показано на рис.31.1 a_3, b_3 . Осциллирующие экстремали всегда полны, поскольку правая часть уравнения (31.18) ограничена сверху значением $1/\Phi(q')$ или $1/\Phi(q'')$ и снизу $1/\max \Phi(q)$.

Если конформный множитель $\Phi(q)$ равен бесконечности в точке q_{i+1} , как показано на рис.31.1 b , то экстремаль общего вида может начинаться и заканчиваться на границе, где $\Phi(q_{i+1}) = \infty$. Эта граница соответствует конечному значению σ_{i+1} при $|q_{i+1}| < \infty$, $m < 1$ и $|q_{i+1}| = \infty$, $m > 1$. При этом все экстремали подходят к границе $\Phi = \infty$ под прямым углом, т.к. правая часть уравнения (31.16) стремится к нулю. Полнота экстремалей общего вида, подходящих к границе $\Phi = \infty$, определяется интегралом

$$\lim_{q \rightarrow q_{i+1}} t \rightarrow \int^{q_{i+1}} \frac{dq}{\sqrt{\Phi}}, \quad (31.23)$$

что следует из уравнений (31.17) и (31.2). Тем самым полнота экстремалей общего вида, подходящих к сингулярной границе, такая же, как и прямых экстремалей, параллельных оси σ (31.19). Они полны только при $|q_{i+1}| = \infty$ и $1 < m \leq 2$.

В лоренцевом случае в области I к границе $\Phi = \infty$ подходят только прямые и пространственноподобные экстремали (см. раздел 30.3). Их полнота связана с показателем степени m так же, как и для римановой метрики.

На рис.31.1 в верхнем ряду (a_1, b_1) показано типичное поведение конформного множителя с одним экстремумом между двумя нулями и с двумя локальными экстремумами между нулем и особенностью. В последнем случае мы предполагаем, что $|q_{i+1}| < \infty$ и $m < 0$, что соответствует сингулярной кривизне. В среднем ряду (a_2, b_2) представлена зависимость этих конформных множителей от координаты σ . На рис.31.1, b_2 значение координаты σ_{i+1} конечно, что следует из сходимости интеграла (31.9). В нижнем ряду (a_3, b_3) показано качественное поведение экстремалей на

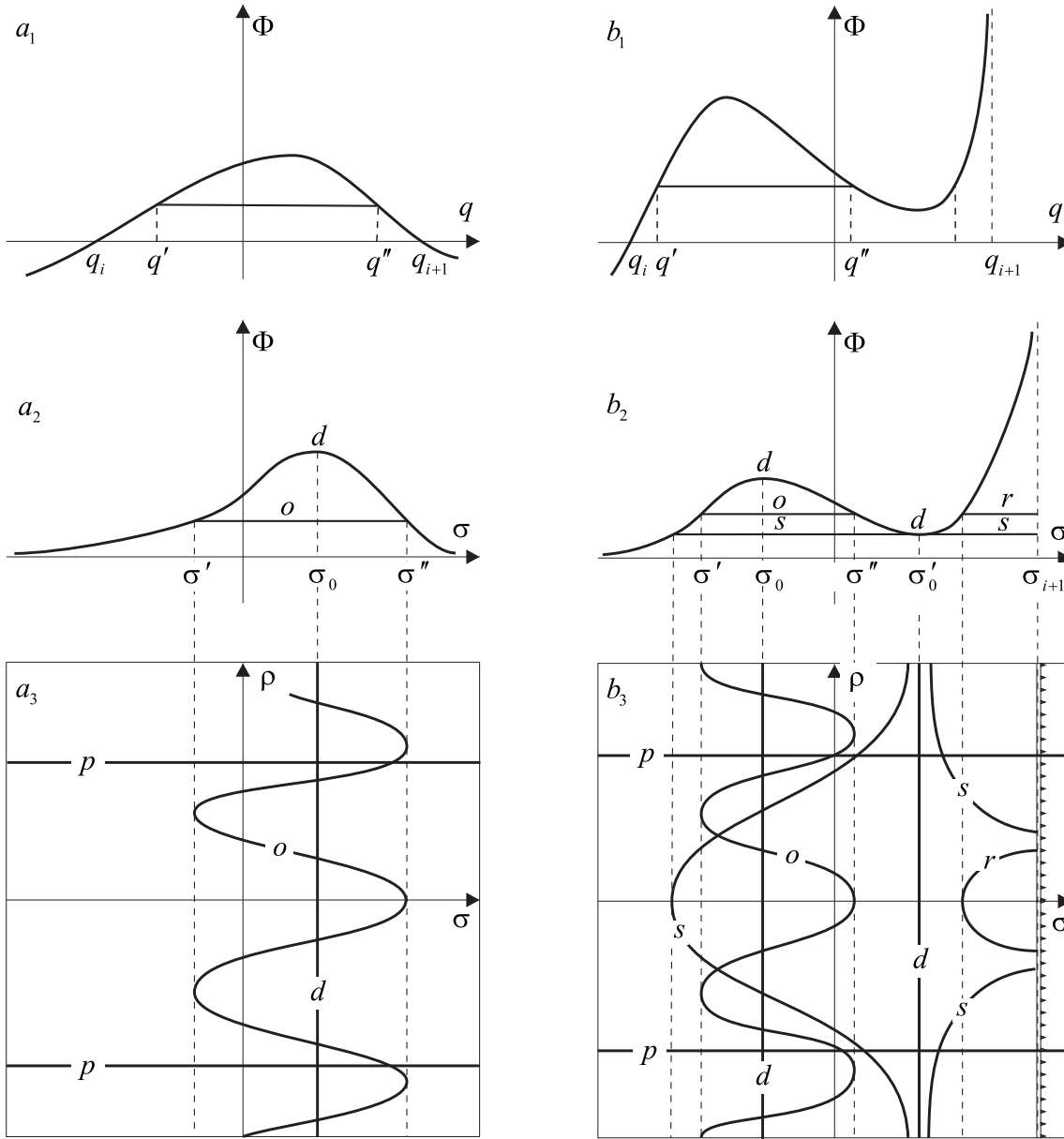


Рис. 31.1: Верхний ряд (a_1, b_1): типичное поведение конформного множителя $\Phi(q)$ между двумя нулями и нулем и сингулярностью. Средний ряд (a_2, b_2): зависимость конформного множителя $\Phi(\sigma)$ от координаты σ . Точка σ_{i+1} может находиться как на конечном, так и на бесконечном расстоянии от начала координат. Нижний ряд (a_3, b_3): типичные экстремали общего вида при разных значениях постоянной C_2 . Вблизи локального максимума экстремали (o) осциллируют между σ' и σ'' . Через каждый локальный экстремум проходит вырожденная экстремаль (d). Существуют экстремали (s), которые асимптотически приближаются к вырожденной экстремали в локальном минимуме σ'_0 и могут заканчиваться на сингулярной границе σ_{i+1} . Экстремали (r) конечной длины начинаются и заканчиваются на сингулярной границе. Все экстремали можно произвольно сдвигать вдоль оси ρ . Через каждую точку ρ проходит также экстремаль (p), параллельная оси σ .

σ, ρ плоскости. Через каждый локальный экстремум конформного множителя проходит вырожденная экстремаль (d). Экстремали общего вида (o) осциллируют вблизи локального максимума $\Phi(\sigma_0)$ между значениями σ' и σ'' , которые определяются постоянной C_2 . Экстремали общего вида (r) могут также начинаться и заканчиваться на границе $\Phi(\sigma_{i+1}) = \infty$, имея при этом конечную длину. Существуют также экстремали общего вида (s), которые асимптотически приближаются при $\rho \rightarrow \pm\infty$ к вырожденной экстремали, проходящей через локальный минимум σ'_0 . Часть из этих экстремалей заканчивается на границе $\Phi(\sigma_{i+1}) = \infty$ при конечном значении канонического параметра.

Все экстремали можно произвольно сдвигать вдоль оси ρ . Через каждую точку ρ проходит также экстремаль (p), параллельная оси σ .

Если обе граничные точки интервала q_i и q_{i+1} являются нулями, то из непрерывности конформного множителя Φ следует, что у него есть по крайней мере один экстремум, через который проходит вырожденная экстремаль (d). Вырожденные экстремали всегда полны, т.е. имеют бесконечную длину, т.к. канонический параметр совпадает с координатой ρ (31.21).

Проведенный анализ экстремалей показывает, что неполнота экстремалей в полсе $\sigma \in (\sigma_i, \sigma_{i+1})$, $\rho \in (-\infty, \infty)$ полностью определяется поведением экстремалей, параллельных оси σ , при подходе к граничным точкам σ_i, σ_{i+1} . Что, в свою очередь, определяется сходимостью интеграла (31.23). Эти экстремали неполны в конечных точках $|q_i| < \infty$ при $m < 2$. В бесконечно удаленных точках $|q_i| = \infty$ они неполны при $m > 2$ и полны во всех остальных случаях. Учитывая, что продолжать поверхность необходимо только при конечных значениях кривизны, приходим к выводу, что продолжение поверхности необходимо только в случае простого нуля, $m = 1$, у конформного множителя в конечной точке $|q_i| < \infty$. Напомним, что простой нуль конформного множителя в лоренцевом случае соответствует горизонту.

31.3 Построение глобальных решений

Чтобы построить максимальное продолжение поверхностей с римановой метрикой (31.1), имеющей по крайней мере один вектор Киллинга, произведем следующее построение, которое будет также полезно для наглядного изображения поверхностей. Поскольку конформный множитель не зависит от ρ , то можно отождествить прямые ρ и $\rho + L$, где $L > 0$ – произвольная положительная постоянная. При таком отождествлении плоскость σ, ρ сворачивается в цилиндр. Длина направляющей окружности P определяется конформным множителем и на границе $q \rightarrow \pm\infty$ равна либо постоянной, либо бесконечности:

$$P^2 = -L^2\Phi(q) \xrightarrow{q \rightarrow \pm\infty} \begin{cases} 0, & \Phi(\pm\infty) = 0, \\ \text{const} \neq 0, & \Phi(\pm\infty) = \text{const} \neq 0, \\ \mp\infty & \Phi(\pm\infty) = \pm\infty. \end{cases} \quad (31.24)$$

В конечных точках $q \rightarrow q_i \neq \pm\infty$ квадрат длины направляющей окружности может принимать только два значения:

$$P^2 = -L^2\Phi(q) \xrightarrow{q \rightarrow q_i} \begin{cases} 0, & \Phi(q_i) = 0, \\ \mp\infty & \Phi(q_i) = \pm\infty. \end{cases} \quad (31.25)$$

При этом плоскость σ, ρ является универсальным накрывающим пространством цилиндра.

$ q_i < \infty$						
	$m < 0$	$0 < m < 1$	$m = 1$	$1 < m < 2$	$m = 2$	$m > 2$
R	∞	∞	const	∞	const	0
σ_i	const	const	∞	∞	∞	∞
P^2	∞	0	0	0	0	0
Полнота	—	—	—	—	+	+

$ q_i = \infty$						
	$m < 0$	$m = 0$	$0 < m \leq 1$	$1 < m < 2$	$m = 2$	$m > 2$
R	0	0	0	0	const	∞
σ_i	∞	∞	∞	const	const	const
P^2	0	const	∞	∞	∞	∞
Полнота	+	+	+	+	+	—

Таблица 31.1: Свойства граничных точек в зависимости от показателя степени m . Символ **const** в строках для скалярной кривизны и длины направляющей окружности обозначает отличную от нуля константу.

Подытожим в таблице 31.1 свойства граничных точек q_i . В зависимости от показателя степени m здесь указаны значения скалярной кривизны R на соответствующей границе, конечность значения координаты σ_i в точке q_i , квадраты длин направляющих окружностей цилиндров P^2 (с точностью до знака) и полнота экстремалей, параллельных оси σ .

На рис. 31.2 показан вид поверхностей вблизи граничной точки q_{i+1} после отождествления $\rho \sim \rho + L$. Поверхности вблизи точки q_i имеют такой же вид, но повернуты в другую сторону. Поверхность, соответствующая всему интервалу (q_i, q_{i+1}) , получается после склейки двух таких поверхностей для граничных точек q_i и q_{i+1} . Из таблицы 31.1 следует, что точка q_i является неполной по экстремали, и при этом кривизна несингулярна в единственном случае, когда $|q_i| < \infty$ при $m = 1$. Таким образом, мы видим, что необходимо продолжить экстремали только вблизи точки $|q_i| < \infty$ при $m = 1$. Поскольку эта точка при лоренцевой сигнатуре метрики соответствует горизонту, то и в евклидовом случае оставим это же название.

Продолжение поверхности через горизонт проводится следующим образом. Прежде всего отметим, что горизонт в евклидовом случае представляет собой точку, так как длина направляющей окружности стремится к нулю. Кроме того, эта “бесконечно удаленная” точка $\sigma = \infty$ в плоскости ρ, σ на самом деле находится на конечном расстоянии, т.к. все экстремали достигают этой точки при конечном значении канонического параметра. В окрестности точки q_i с точностью до членов более высокого порядка конформный множитель имеет вид

$$\Phi(q) = \Phi'_i(q - q_i), \quad (31.26)$$

где $\Phi'_i := \Phi'(q_i) = \text{const} \neq 0$. В области I при $\Phi'_i > 0$ и $q > q_i$ уравнение (31.2) легко интегрируется

$$q - q_i = e^{\Phi'_i \sigma},$$

где отброшена несущественная постоянная интегрирования, соответствующая сдвигу σ . Таким образом, граничная точка $|q_i| < \infty$ достигается при $\sigma \rightarrow -\infty$. В координатах q, ρ , которые играют роль шварцшильдовских координат в лоренцевом случае,

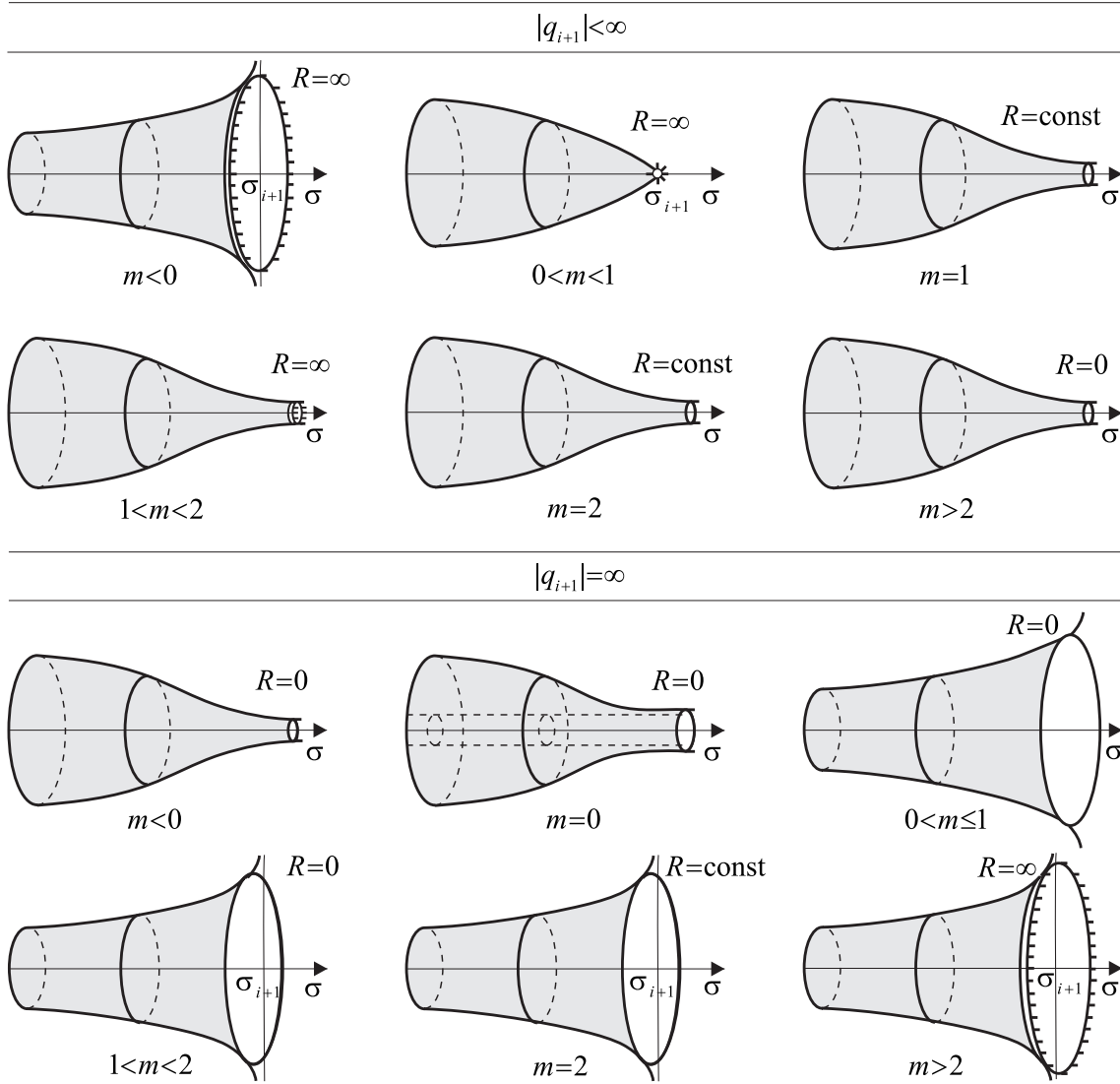


Рис. 31.2: Вид поверхности вблизи граничной точки q_{i+1} после отождествления $\rho \sim \rho + L$. Мы предполагаем, что координата σ возрастает слева направо и $\sigma_{i+1} > \sigma_i$. Поверхности вблизи точки q_i имеют такой же вид, но повернуты в другую сторону. Поверхность для интервала q_i, q_{i+1} получается после склейки двух таких поверхностей для точек q_i и q_{i+1} .

метрика (31.1) принимает вид

$$-ds^2 = \frac{dq^2}{\Phi'_i(q - q_i)} + \Phi'_i(q - q_i)d\rho^2. \quad (31.27)$$

В полярных координатах r, φ , определенных формулами

$$q - q_i := \frac{\Phi'_i}{4}r^2, \quad \rho := \frac{2}{\Phi'_i}\varphi, \quad (31.28)$$

метрика становится евклидовой

$$-ds^2 = dr^2 + r^2d\varphi^2.$$

При этом полярный угол, что важно, меняется в интервале $\varphi \in (0, L\Phi'_i/2)$ при $\rho \in [0, L]$, а радиус r определен в окрестности граничной точки q_i соотношением (31.28). Рассмотренное преобразование координат отображает “бесконечно удаленную” точку $\sigma_i = -\infty$ в начало координат евклидовой плоскости. Поскольку полярный угол меняется в интервале, который в общем случае отличен от интервала $(0, 2\pi)$, то в начале координат возможно наличие конической сингулярности. Соответствующий угол дефицита равен

$$2\pi\theta := \frac{L\Phi'_i}{2} - 2\pi. \quad (31.29)$$

Таким образом, мы получили евклидову метрику на плоскости $\mathbb{R}^2$ с конической сингулярностью. При

$$L = 4\pi/\Phi_i \quad (31.30)$$

угол дефицита равен нулю, коническая сингулярность отсутствует, и мы получаем плоскую евклидову метрику, гладкость которой в начале координат очевидна. В общем случае конформный множитель (31.26) вблизи граничной точки q_i содержит поправки более высокого порядка, и переход к полярным координатам при нулевом угле дефицита дает метрику того же класса гладкости, что и конформный множитель.

Таким образом, в общем случае продолжение решения через точку $|q_i| < \infty$ при $t = 1$ не имеет смысла, т.к. эта точка соответствует конической сингулярности. Мы считаем, что точка, в которой сосредоточена коническая сингулярность, как и носитель всякой другой сингулярности, не принадлежит многообразию. Поэтому универсальной накрывающей для такого решения будет плоскость σ, ρ или ее часть с метрикой (31.1). В исключительном случае отсутствия конической сингулярности при $L = 4\pi/\Phi'_i$ продолжение необходимо. При этом прямые экстремали, параллельные оси σ и проходящие через точки ρ и $\rho + L/2$, представляют собой две половины одной экстремали, как показано на рис.31.3. При отсутствии конической сингулярности фундаментальная группа поверхности тривиальна, и поэтому соответствующая поверхность сама представляет собой универсальную накрывающую.

Если на обоих концах интервала (q_i, q_{i+1}) конформный множитель имеет асимптотики $\Phi \sim (q - q_i)$ и $\Phi \sim (q_{i+1} - q)$, то поверхность необходимо продолжить в обеих точках $\sigma = \pm\infty$. После отождествления $\rho \sim \rho + L$ в общем случае в граничных точках возникают конические сингулярности. Если $L\Phi'_i = 4\pi$ и $L\Phi'_{i+1} = 4\pi$, то конические сингулярности отсутствуют. На рис.31.4 показаны три возможных вида глобальных поверхностей, соответствующих наличию или отсутствию конических сингулярностей. Эти поверхности имеют топологию цилиндра, плоскости и сферы, соответственно.

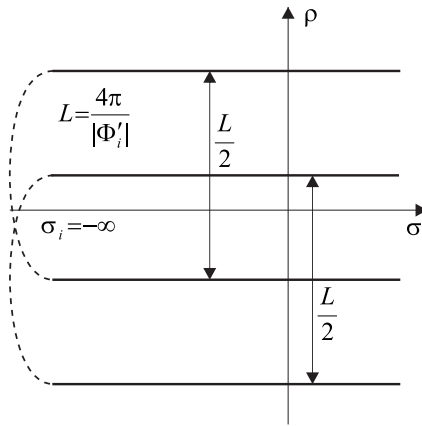


Рис. 31.3: Продолжение прямых экстремалей, проходящих через точки ρ и $\rho + L/2$ в отсутствие конической сингулярности $L = 4\pi/|\Phi'_i|$. Отождествление происходит в точке $\sigma_i = -\infty$.

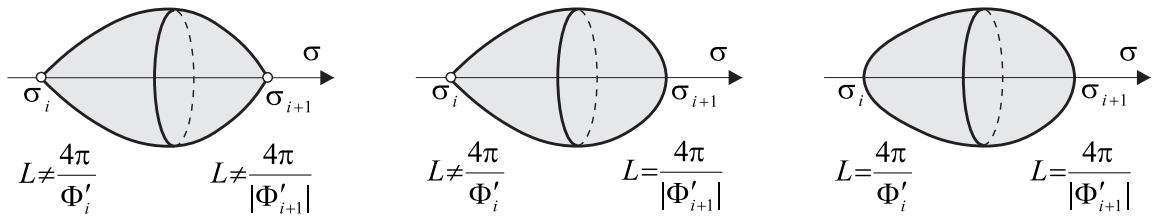


Рис. 31.4: Три возможных вида поверхностей, соответствующих интервалу (q_i, q_{i+1}) , когда конформный множитель имеет асимптотику $\Phi \sim q - q_{i,i+1}$ в граничных точках. При отождествлении $\rho \sim \rho + L$ в общем случае возникают две конические сингулярности. При $L|\Phi'_{i,i+1}| = 4\pi$ конические сингулярности отсутствуют.

Сформулируем правила построения максимально продолженных решений с метрикой вида (31.1)

1. После отождествления $\rho \sim \rho + L$ каждому интервалу (q_i, q_{i+1}) соответствует максимально продолженное решение, которое получается склеиванием двух поверхностей, изображенных на рис.31.2 и соответствующих граничным точкам q_i и q_{i+1} .
2. Во всех случаях, кроме отсутствия конической сингулярности $|q_i| < \infty$, $L\Phi'_i \neq 4\pi$ или $|q_{i+1}| < \infty$, $L|\Phi'_{i+1}| \neq 4\pi$, полоса $\sigma \in (\sigma_i, \sigma_{i+1})$, $\rho \in \mathbb{R}$ с метрикой (31.1) является универсальным накрывающим пространством для соответствующего максимально продолженного решения.
3. При отсутствии одной из конических сингулярностей, $|q_i| < \infty$, $L\Phi'_i = 4\pi$, или $|q_{i+1}| < \infty$, $L|\Phi'_{i+1}| = 4\pi$, поверхность, полученная из плоскости σ, ρ путем отождествления $\rho \sim \rho + L$, представляет собой максимально продолженное решение с тривиальной фундаментальной группой.

Отметим, что в общем случае при построении максимально продолженных решений нам не приходилось склеивать карты. Это значит, что соответствующие поверхности принадлежат классу C^∞ . При отсутствии конической сингулярности переход к полярным координатам (31.28) не использует явный вид конформного множителя,

и возникающая поверхность так же, как и евклидова плоскость является бесконечно дифференцируемым многообразием. Тем самым мы доказали теорему, оправдывающую приведенные выше правила построения глобальных решений.

Теорема 31.3.1. *Универсальное накрывающее пространство, построенное по правилам 1–3, является максимально продолженным гладким многообразием класса C^∞ с римановой метрикой класса C^l , $l \geq 2$, таким, что область вне горизонта изометричная (части) плоскости ρ, σ с метрикой (31.1).*

Таким образом построены глобальные решения для широкого класса двумерных метрик евклидовой сигнатуры, которые имеют один вектор Киллинга. Доказана гладкость возникающих двумерных поверхностей и дифференцируемость метрики. Этот метод конструктивен и является аналогом метода конформных блоков для двумерных метрик лоренцевой сигнатуры. Почти всегда каждый конформный блок представляет собой универсальное накрывающее пространство для максимально продолженной римановой поверхности. Исключения составляют поверхности с горизонтом, который в евклидовом случае представляет собой точку.

31.4 Решение Шварцшильда

В качестве примера рассмотрим решение Шварцшильда. Для этого решения t, r часть метрики имеет вид

$$ds^2 = |\Phi(q)|(d\tau^2 - d\sigma^2), \quad (31.31)$$

где

$$\Phi(q) = 1 - \frac{2M}{q}, \quad q = \begin{cases} r, & \Phi(q) > 0 \\ t, & \Phi(q) < 0. \end{cases}$$

При этом связь переменной q с одной из координат τ или σ задается дифференциальным уравнением (30.2). Максимально продолженное решение Шварцшильда имеет вид топологического произведения двух поверхностей $\mathbb{U} \times \mathbb{S}^2$, где $\mathbb{S}^2$ – сфера и $\mathbb{U}$ – двумерная лоренцева поверхность, которая наглядно изображается в виде диаграммы Картера–Пенроуза, рис.31.5 слева.

Диаграмма Картера–Пенроуза после изменения сигнатуры метрики распадается на четыре несвязные между собой поверхности, которые показаны на рис.31.5 справа. Областям вне горизонта черной дыры соответствуют две римановы поверхности I и III с отрицательно определенной метрикой. Четырехмерная метрика является отрицательно определенной, $\text{sign } g_{\mu\nu} = (- - - -)$. Отрицательная определенность евклидовой метрики связана с выбором сигнатуры для метрики Шварцшильда (31.31) и может быть изменена. Это решение обычно рассматривается, как евклидова версия решения Шварцшильда. Области под горизонтами черной и белой дыр соответствуют поверхностям II и IV, соответственно. При этом сигнатура полной метрики равна $\text{sign } g_{\mu\nu} = (+ + - -)$.

Ситуация с евклидовой версией решения Шварцшильда типична. Связная диаграмма Картера–Пенроуза в лоренцевом случае разрезается вдоль всех горизонтов и переходит в несвязные римановы поверхности с отрицательно или положительно определенной метрикой. При этом переход через горизонт с нечетным m соответствует изменению знака римановой метрики.

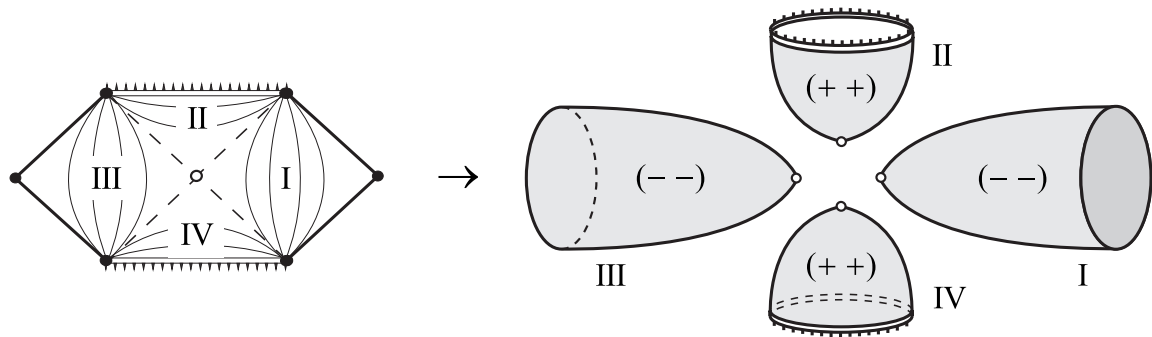


Рис. 31.5: Диаграмма Картера–Пенроуза для решения Шварцшильда при евклидовой сигнатуре метрики распадается на четыре несвязных между собой поверхности. Областям вне горизонта соответствуют две поверхности с отрицательно определенной метрикой I, III, а областям внутри горизонта – две поверхности с положительно определенной метрикой.

Глава 32

Двумерная гравитация

Модели двумерной гравитации привлекают большое внимание в последнее время по трем причинам. Во-первых, эти модели являются эффективными моделями гравитации в двух измерениях после редукции многомерных теорий гравитации. Например, сферически симметричная общая теория относительности сводится к двумерной гравитации, где в качестве координат выступают время и радиус. Во-вторых, модели двумерной гравитации относительно просты и поддаются аналитическому анализу, что позволяет глубже понять многомерные модели. В-третьих, интерес к двумерным моделям гравитации возрос после появления теории струн, в которых метрика на мировой поверхности струны становится новой динамической переменной.

В настоящей главе мы опишем модель двумерной гравитации довольно общего вида. Проведем гамильтонов анализ и проинтегрируем уравнения движения. В качестве примера мы построим все глобальные решения в квадратичной модели гравитации с кручением. В заключительных разделах будет рассмотрена пуассонова сигма модель и двумерная теория полей Янга–Миллса.

32.1 Двумерная гравитация с кручением

Рассмотрим двумерное многообразие (поверхность) $\mathbb{M}$. Пусть x^α , $\alpha = 0, 1$, – локальная система координат. Предположим, что на $\mathbb{M}$ задана геометрия Римана–Картана, т.е. метрика $g_{\alpha\beta}$ и кручение $T_{\alpha\beta}{}^\gamma := \Gamma_{\alpha\beta}{}^\gamma - \Gamma_{\beta\alpha}{}^\gamma$, где $\Gamma_{\alpha\beta}{}^\gamma$ – компоненты аффинной связности (см. раздел 6.2). Предположим также, что метрика имеет лоренцеву сигнатуру, $\text{sign } g_{\alpha\beta} = (+-)$.

Метрика и кручение полностью определяют геометрию на $\mathbb{M}$. В геометрии Римана–Картана они однозначно задают аффинную связность $\Gamma_{\alpha\beta}{}^\gamma$ (6.21) и соответствующий тензор кривизны $R_{\alpha\beta\gamma}{}^\delta$ (6.74). Конечно, поскольку на поверхности задана метрика, то определены также символы Кристоффеля $\tilde{\Gamma}_{\alpha\beta}{}^\gamma$ и построенный по ним тензор кривизны $\tilde{R}_{\alpha\beta\gamma}{}^\delta$. Эти геометрические объекты, построенные при нулевом тензоре кручения, мы будем, как и раньше, помечать знаком тильды.

В двумерном случае геометрия существенно упрощается ввиду наличия полностью антисимметричного тензора второго ранга $\varepsilon_{\alpha\beta} = -\varepsilon_{\beta\alpha}$. Действительно, поскольку тензор кривизны в геометрии Римана–Картана $R_{\alpha\beta\gamma\delta}$ со всеми опущенными индексами антисимметричен как по первой, так и по второй паре индексов, то он всегда представим в виде

$$R_{\alpha\beta\gamma\delta} = -\frac{1}{2}\varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta}R, \quad (32.1)$$

где $R := R_{\alpha\beta\gamma}{}^\beta g^{\alpha\gamma}$ – скалярная кривизна. Такому же равенству удовлетворяет тензор кривизны, построенный только по метрике

$$\tilde{R}_{\alpha\beta\gamma\delta} = -\frac{1}{2}\varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta}\tilde{R}. \quad (32.2)$$

Поскольку тензор кручения антисимметричен по первой паре индексов, то он всегда представим виде

$$T_{\alpha\beta}{}^\gamma = -\frac{1}{2}\varepsilon_{\alpha\beta}T^{*\gamma}, \quad (32.3)$$

где $T^{*\gamma} := T_{\alpha\beta}{}^\gamma \varepsilon^{\alpha\beta}$ – псевдослед тензора кручения. Приставка “псевдо-” в данном случае означает, что вектор $T^{*\gamma}$ относительно общих преобразований координат меняет знак при пространственных отражениях $x^1 \mapsto -x^1$, когда координаты выбраны соответствующим образом.

Таким образом, в геометрии Римана–Картана на поверхности тензор кривизны имеет только одну независимую компоненту, в качестве которой мы выбираем скалярную кривизну R , а тензор кручения – две $T^{*\alpha}$.

Хорошо известно, что в действии Гильберта–Эйнштейна

$$S_{\text{HE}} = \int_{\mathbb{M}} dx \sqrt{|g|} \tilde{R}, \quad (32.4)$$

подынтегральное выражение $\sqrt{|g|}\tilde{R} = \partial_\alpha(\dots)^\alpha$ представляет собой полную дивергенцию (в дальнейшем мы это увидим), и интеграл является топологическим инвариантом, который пропорционален эйлеровой характеристике поверхности. Отсюда следует, что вариация действия Гильберта–Эйнштейна (32.4) по метрике в двумерном случае не дает никаких уравнений движения. В этом также можно убедиться непосредственно. Действительно, из выражения для тензора кривизны (32.2) следует, что равенство

$$\tilde{R}_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\tilde{R} = 0$$

в двумерном случае является тождеством. В этом смысле общая теория относительности в двумерном пространстве-времени является тривиальной.

Мы видим, что для того, чтобы построить нетривиальную модель гравитации в двумерном пространстве-времени требуется какое-либо обобщение. Простейшим геометрическим обобщением является рассмотрение на $\mathbb{M}$ не псевдоримановой геометрии, на которой основана общая теория относительности, а геометрии более общего вида. Ниже мы покажем, что нетривиальную двумерную модель гравитации можно построить в рамках геометрии Римана–Картана.

Чтобы выполнить построение, перейдем от переменных метрика $g_{\alpha\beta}$, кручение $T_{\alpha\beta}{}^\gamma = -T_{\beta\alpha}{}^\gamma$ к переменным Картана: репер (диада) $e_\alpha{}^a$, лоренцева связность $\omega_\alpha{}^{ab} = -\omega_\alpha{}^{ba}$. Эти переменные определяются следующими равенствами (см. раздел 6.9):

$$g_{\alpha\beta} = e_\alpha{}^a e_\beta{}^b \eta_{ab}, \quad (32.5)$$

$$\nabla_\alpha e_\beta{}^a = \partial_\alpha e_\beta{}^a - \Gamma_{\alpha\beta}{}^\gamma e_\gamma{}^a - \omega_\alpha{}^{ab} e_{\beta b} = 0, \quad (32.6)$$

где $\eta_{ab} = \text{diag}(+,-)$ – двумерная метрика Лоренца. Первое равенство определяет репер с точностью до локальных лоренцевых вращений, а второе – взаимно однозначное соответствие между тензором кручения и лоренцевой связностью (при заданном репере).

Напомним, что переход от греческих индексов к латинским и наоборот всегда осуществляется с помощью репера и его обратного. Подъем и опускание греческих индексов производится с помощью метрики $g_{\alpha\beta}$, а латинских – с помощью метрики Лоренца η_{ab} . Кроме того, переход от греческих индексов к латинским можно выполнять под знаком ковариантного дифференцирования в силу условия метричности (32.6).

Тензоры кривизны и кручения в переменных Картана имеют вид (5.56), (5.57),

$$R_{\alpha\beta}{}^{ab} = \partial_\alpha \omega_\beta{}^{ab} - \omega_\alpha{}^{ac} \omega_{\beta c}{}^b - (\alpha \leftrightarrow \beta), \quad (32.7)$$

$$T_{\alpha\beta}{}^a = \partial_\alpha e_\beta{}^a - \omega_\alpha{}^{ab} e_{\beta b} - (\alpha \leftrightarrow \beta). \quad (32.8)$$

В двумерном случае лоренцева связность взаимно однозначно определяется псевдо вектором ω_α :

$$\omega_\alpha{}^{ab} = \omega_\alpha \epsilon^{ab}. \quad (32.9)$$

Поскольку собственная ортохронная двумерная группа Лоренца $\mathbb{SO}_0(1,1)$ является абелевой, то нелинейные слагаемые в выражении для тензора кривизны (32.7) сокращаются. Прямые вычисления приводят к следующему выражению

$$R_{\alpha\beta}{}^{ab} = F_{\alpha\beta} \epsilon^{ab}, \quad F_{\alpha\beta} := \partial_\alpha \omega_\beta - \partial_\beta \omega_\alpha, \quad (32.10)$$

которое напоминает выражение для напряженности электромагнитного поля.

Нетрудно также вычислить скалярную кривизну и псевдослед тензора кручения, умноженные на элемент объема, в новых переменных:

$$\hat{R} := \sqrt{|g|} R = 2\hat{\epsilon}^{\alpha\beta} \partial_\alpha \omega_\beta, \quad (32.11)$$

$$\hat{T}^{*a} = e T^{*a} = 2\hat{\epsilon}^{\alpha\beta} (\partial_\alpha e_\beta{}^a - \omega_\alpha \epsilon^a{}_b e_\beta{}^b), \quad (32.12)$$

где $\hat{\epsilon}^{\alpha\beta} = \sqrt{|g|} \epsilon^{\alpha\beta} = \sqrt{|g|} \epsilon^{ab} e_a^\alpha e_b^\beta$ – полностью антисимметричная тензорная плотность второго ранга, $\hat{\epsilon}^{01} = -\hat{\epsilon}^{10} = -1$. Из выражения (32.11) следует, что действие Гильберта–Эйнштейна в двумерном случае, как уже отмечалось, представляет собой дивергенцию $\sqrt{|g|} \hat{R} = 2\hat{\epsilon}^{\alpha\beta} \partial_\alpha \tilde{\omega}_\beta$.

Теперь все готово для построения действия для двумерной гравитации с кручением. Из выражений для тензора кривизны и кручения в переменных Картана (32.7), (32.8) следует, что для того, чтобы уравнения движения для репера и лоренцевой связности были второго порядка, достаточно рассмотреть инварианты, квадратичные по тензору кривизны и кручения. Более высокие степени этих геометрических объектов в общем случае будут давать уравнения более высокого порядка. Поскольку тензор кривизны имеет только одну независимую компоненту, то среди квадратичных инвариантов только один является независимым. Аналогично, среди инвариантов, квадратичных по тензору кручения, также только один является независимым. Поэтому наиболее общий квадратичный лагранжиан можно записать в следующем виде

$$L = -\sqrt{|g|} \left(\frac{1}{4} \gamma R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} + \frac{1}{4} \beta T^{\alpha\beta\gamma} T_{\alpha\beta\gamma} + \lambda \right), \quad (32.13)$$

где γ, β – некоторые константы связи и λ – космологическая постоянная. Это есть лагранжиан двумерной гравитации с кручением.

Из равенств (32.1) и (32.3) следует, что лагранжиан (32.13) можно переписать в терминах неприводимых компонент тензоров кривизны и кручения:

$$L = -\sqrt{|g|} \left(\frac{1}{4} \gamma R^2 - \frac{1}{8} \beta T^{*\alpha} T_\alpha^* + \lambda \right). \quad (32.14)$$

Действие для лагранжиана (32.13) инвариантно относительно общих преобразований координат и локальных лоренцевых вращений.

Лагранжиан (32.13) был предложен в [222, 232, 233] в качестве обобщения теории струн, когда метрика на мировой поверхности бозонной струны становится независимой динамической переменной. Вскоре было найдено общее решение уравнений движения, вытекающих из лагранжиана (32.13) в конформной калибровке для репера [234, 235]. Был также проведен анализ экстремалей и геодезических [236] и построены все глобальные (максимально продолженные вдоль геодезических) решения квадратичной двумерной гравитации с кручением [214].

В следующих разделах мы рассмотрим действие для двумерной гравитации общего вида и решим явно соответствующие гамильтоновы уравнения движения без полей материи и без фиксирования какой-либо калибровки.

32.2 Действие и уравнения движения

Действие для двумерной гравитации с кручением общего вида задается интегралом

$$S = \int_{\mathbb{M}} dx (L_G + L_X), \quad (32.15)$$

где гравитационную часть лагранжиана запишем в формализме первого порядка

$$L_G := -\frac{1}{2}(\pi \hat{R} + p_a \hat{T}^{*a}) - \sqrt{|g|} \left(\frac{1}{2} p_a p^a U + V \right). \quad (32.16)$$

Здесь плотности скалярной кривизны $\hat{R}$ и псевдоследа тензора кручения $\hat{T}^{*a}$ заданы равенствами (32.11), (32.12). Этот лагранжиан зависит от двух произвольных дифференцируемых функций $U = U(\pi)$ и $V = V(\pi)$ от π . В качестве независимых переменных, по которым производится варьирование, мы рассматриваем репер $e_\alpha^a(x)$, лоренцеву связность $\omega_\alpha(x)$, а также поля $\pi(x)$ и $p_a(x)$.

В действии (32.15) мы добавили к лагранжиану двумерной гравитации скалярное поле $X(x)$ с лагранжианом

$$L_X := -\frac{1}{2} \rho \sqrt{|g|} (\partial X^2 - m^2 X^2), \quad (32.17)$$

где $m = \text{const}$ – масса скалярного поля, введено сокращенное обозначение для кинетической части

$$\partial X^2 := g^{\alpha\beta} \partial_\alpha X \partial_\beta X,$$

и в общем случае $\rho = \rho(\pi)$ – некоторая функция от π . Если $\rho = 1$, то скалярное поле связано с гравитацией минимальной связью. Тогда лагранжиан легко обобщается на случай бозонной струны, движущейся в D -мерном пространстве-времени. Для этого достаточно заменить одно скалярное поле на несколько: $X(x) \mapsto X^\mu(x)$, $\mu = 0, 1, \dots, D-1$ и положить $m = 0$. При метрике объемлющего пространства $\eta_{\mu\nu} = \text{diag}(+ - \dots -)$, в котором распространяется струна, общий знак минус в лагранжиане (32.17) обеспечивает положительную определенность канонического гамильтониана для пространственных компонент струны X^μ . В любом случае общий знак лагранжиана можно изменить, переопределив ρ . Если $\rho(\pi)$ – нетривиальная функция, то мы имеем неминимальную связь скалярного поля с гравитацией. Такая связь возникает, например, при редукции многомерных моделей гравитации.

Мы также предполагаем, что π является скалярным, а p^a – псевдовекторным полями на $\mathbb{M}$. Тогда действие для двумерной гравитации (32.15) инвариантно относительно общих преобразований координат, локальных лоренцевых вращений и пространственных отражений.

Уравнения движения для действия (32.15) имеют вид

$$\frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta \pi} : \quad -\frac{1}{2}R - \left(\frac{1}{2}p^a p_a U' + V' \right) - \frac{1}{2}\rho'(\partial X^2 - m^2 X^2) = 0, \quad (32.18)$$

$$\frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta p_a} : \quad -\frac{1}{2}T^{*a} - p^a U = 0, \quad (32.19)$$

$$\hat{\varepsilon}_{\beta\alpha} \frac{\delta S}{\delta \omega_\beta} : \quad \partial_\alpha \pi - p_a \varepsilon^a_b e_\alpha^b = 0, \quad (32.20)$$

$$\hat{\varepsilon}_{\beta\alpha} \frac{\delta S}{\delta e_\beta^a} : \quad \nabla_\alpha p_a + \varepsilon_{\alpha a} \left(\frac{1}{2}p^b p_b U + V \right) + \rho T^\beta_{\alpha} \varepsilon_{\beta\alpha} = 0, \quad (32.21)$$

$$\frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta X} : \quad \rho(g^{\alpha\beta} \tilde{\nabla}_\alpha \tilde{\nabla}_\beta X + m^2 X) + \rho' g^{\alpha\beta} \partial_\alpha \pi \partial_\beta X = 0, \quad (32.22)$$

где $U' := dU/d\pi$ и $V' := dV/d\pi$, введено обозначение для ковариантной производной

$$\nabla_\alpha p_a := \partial_\alpha p_a - \omega_\alpha \varepsilon_a^b p_b$$

и

$$T_{\alpha\beta} = \partial_\alpha X \partial_\beta X - \frac{1}{2}g_{\alpha\beta}(\partial X^2 - m^2 X^2) \quad (32.23)$$

– тензор энергии-импульса скалярного поля.

Вариации действия по $\pi(x)$ и $p_a(x)$ приводит к алгебраическим уравнениям (32.18), (32.19) для $\pi(x)$ и $p_a(x)$. Если эти уравнения можно решить в явном виде, то после подстановки найденных решений в исходное действие получим действие для двумерной гравитации в формализме второго порядка. Соответствующее действие в формализме второго порядка будет некоторой функцией от скалярной кривизны и квадрата тензора кручения, которая определяется функциями $U(\pi)$ и $V(\pi)$.

Случай $U = 0$ и $V = 0$ в отсутствие полей материи, $X = 0$, описывает поверхности нулевой кривизны и кручения, и не представляет интереса с геометрической точки зрения. Поэтому мы предположим, что по крайней мере одна из функций U или V отлична от нуля.

Лагранжиан (32.16) представляет собой лагранжиан очень общего вида, т.к. зависит от двух произвольных функций U и V .

Пример 32.2.1. Если $U = 0$, то кручение равно нулю, что следует из уравнения (32.19). Тогда уравнение (32.18) неявно определяет функцию $\pi = \pi(R)$ при отсутствии скалярного поля, $X = 0$, или минимальной связи, $\rho' = 0$. После подстановки этого решения в исходный лагранжиан получим действие в формализме второго порядка. Соответствующий гравитационный лагранжиан представляет собой произвольную функцию от скалярной кривизны, определяемую функцией $V(\pi)$. $\square$

Пример 32.2.2. Если

$$U = \frac{1}{\beta}, \quad V = -\frac{1}{4\gamma}\pi^2 + \lambda, \quad (32.24)$$

то уравнения движения (32.18) и (32.19) при $X = 0$ или $\rho' = 0$ решаются в явном виде:

$$\pi = \gamma R, \quad p^a = -\frac{\beta}{2}T^{*a}.$$

после подстановки этого решения в исходное действие, получим лагранжиан квадратичной двумерной гравитации с кручением в формализме второго порядка (32.14). $\square$

К модели двумерной гравитации (32.15) можно подойти с другой стороны. При произвольных функциях $U(\pi)$ и $V(\pi)$ лоренцеву связность ω_α можно исключить из рассмотрения, решив алгебраические уравнения движения (см. далее раздел 32.4). Это приводит к дилатонной двумерной гравитации общего вида, причем функция $\pi(x)$ играет роль поля дилатона. Среди этих моделей важную роль играет сферически редуцированная общая теория относительности со скалярным полем (см. раздел ??), для которой

$$U = \frac{1}{2\pi}, \quad V = -2\kappa K - \frac{\Lambda\pi}{\kappa}, \quad \rho = \frac{\pi}{2\kappa}, \quad (32.25)$$

где $K = 1, 0, -1$ для сферической, планарной и гиперболической редукции, соответственно, [237] (см. раздел 33), Λ – четырехмерная космологическая постоянная и κ – гравитационная постоянная в четырех измерениях.

Отметим, что при произвольных функциях U и V алгебраические уравнения (32.18) и (32.19) не всегда можно решить в явном виде относительно π и p^a . В таких случаях действие нельзя записать в формализме второго порядка. С другой стороны, действие в формализме второго порядка всегда можно записать в формализме первого порядка. С этой точки зрения действие (32.15) в формализме первого порядка является более общим действием для двумерной гравитации.

В разделе 32.6.1 при отсутствии скалярного поля, $X = 0$, мы решим в явном виде полную систему уравнений движения (32.18)–(32.21) для произвольных функций U и V и без фиксирования какой-либо калибровки.

Остановимся на общих свойствах системы уравнений движения (32.18)–(32.22). Действие (32.15) инвариантно относительно локальных лоренцевых вращений и общих преобразований координат. Согласно второй теореме Нетер эта локальная инвариантность приводит к зависимости уравнений движения (см. раздел 20.3).

Бесконечно малые локальные вращения на угол $\omega(x) \ll 1$ запишем в виде

$$\begin{aligned} \delta e_\alpha^a &= -e_\alpha^b \varepsilon_b^a \omega, & \delta \omega_\alpha &= \partial_\alpha \omega, \\ \delta p_a &= \varepsilon_a^b p_b \omega, & \delta \pi &= 0, & \delta X &= 0. \end{aligned} \quad (32.26)$$

Согласно второй теореме Нетер из инвариантности действия относительно этих преобразований вытекает зависимость уравнений движения:

$$\tilde{\nabla}_\alpha \frac{\delta S}{\delta \omega_\alpha} + \frac{\delta S}{\delta e_\alpha^a} e_\alpha^b \varepsilon_b^a - \frac{\delta S}{\delta p_a} \varepsilon_a^b p_b = 0, \quad (32.27)$$

где $\tilde{\nabla}_\alpha$ обозначает ковариантную производную с символами Кристоффеля.

Кроме этого действие (32.15) инвариантно относительно общих преобразований координат, которые в инфинитезимальной форме имеют вид

$$\begin{aligned} \delta e_\alpha^a &= -\partial_\alpha \epsilon^\beta e_\beta^a - \epsilon^\beta \partial_\beta e_\alpha^a, & \delta \omega_\alpha &= -\partial_\alpha \epsilon^\beta \omega_\beta - \epsilon^\beta \partial_\beta \omega_\alpha, \\ \delta p_a &= -\epsilon^\beta \partial_\beta p_a, & \delta \pi &= -\epsilon^\beta \partial_\beta \pi, & \delta X &= -\epsilon^\beta \partial_\beta X, \end{aligned} \quad (32.28)$$

где $\epsilon^\alpha(x)$ – инфинитезимальный параметр преобразования $x^\alpha \mapsto x^\alpha + \epsilon^\alpha(x)$. Вместе с тождеством (32.27) это приводит к двум линейным соотношениям между уравнениями движения:

$$e_\alpha^a \nabla_\beta \frac{\delta S}{\delta e_\beta^a} - \frac{\delta S}{\delta e_\beta^a} T_{\alpha\beta}^a + \frac{1}{2} \frac{\delta S}{\delta \omega_\beta} \varepsilon_{\alpha\beta} R - \frac{\delta S}{\delta \pi} \partial_\alpha \pi - \frac{\delta S}{\delta p_a} \nabla_\alpha p_a - \frac{\delta S}{\delta X} \partial_\alpha X = 0, \quad (32.29)$$

где ковариантная производная действует с символами Кристоффеля на греческие индексы и лоренцевой связностью – на латинские.

32.3 Каноническая формулировка

Анализ уравнений движения двумерной гравитации с действием (32.15) оказывается значительно проще в канонической формулировке. Более того, общее решение уравнений движения для двумерной гравитации без полей материи удастся выписать только в каноническом формализме. При этом обобщенные импульсы играют роль произвольных функций, параметризующих общие преобразования координат и локальные лоренцевы вращения.

Будем считать, что координаты на $\mathbb{M}$ выбраны таким образом, что координата x^0 является временем, т.е. векторное поле ∂_0 времениподобно, и все сечения $x^0 = \text{const}$ пространственноподобны. Для краткости введем обозначения

$$x^0 := \tau, \quad x^1 := \sigma.$$

Двумерная гравитация представляет собой калибровочную модель, т.к. инвариантна относительно локальных преобразований. Мы увидим, что эта инвариантность приводит к возникновению в гамильтоновом формализме связей первого рода. Общий подход к гамильтоновой формулировке моделей со связями описан в разделе 22.

Обобщенные импульсы, сопряженные лоренцевой связности ω_α , реперу e_α^a и скалярным полям X имеют вид

$$\pi^0 := \frac{\partial(L_G + L_X)}{\partial(\partial_0 \omega_0)} = 0, \quad (32.30)$$

$$\pi^1 := \frac{\partial(L_G + L_X)}{\partial(\partial_0 \omega_1)} = \pi, \quad (32.31)$$

$$p^0_a := \frac{\partial(L_G + L_X)}{\partial(\partial_0 e_0^a)} = 0, \quad (32.32)$$

$$p^1_a := \frac{\partial(L_G + L_X)}{\partial(\partial_0 e_1^a)} = p_a, \quad (32.33)$$

$$P := \frac{\partial(L_G + L_X)}{\partial(\partial_0 X)} = \frac{\rho}{\sqrt{|g|}} (g_{11} \partial_0 X - g_{01} \partial_1 X). \quad (32.34)$$

Мы видим, что гравитационный лагранжиан в формализме первого порядка (32.16) уже записан в гамильтоновой форме. Прямые вычисления приводят к следующему гамильтониану

$$\begin{aligned} H &:= \int d\sigma (p \partial_0 \omega_1 + p_a \partial_0 e_1^a + P \partial_0 X - L_G - L_X) = \\ &= \int d\sigma (\omega_0 G + e_0^a G_a), \end{aligned} \quad (32.35)$$

где

$$G := -\partial_1 \pi + p_a \varepsilon^a_b e_1^b, \quad (32.36)$$

$$G_a := -\partial_1 p_a - \omega_1 p_b \varepsilon^b_a + e_1^b \varepsilon_{ab} \left(\frac{1}{2} p^c p_c U + V - \frac{\rho}{2} m^2 X^2 \right) + \frac{e_1^b \varepsilon_{ab}}{g_{11}} \left(\frac{1}{2\rho} P^2 + \frac{\rho}{2} \partial_1 X^2 \right) + \frac{e_{1a}}{g_{11}} P \partial_1 X. \quad (32.37)$$

Функции G и G_a являются лоренцевым скаляром и вектором, соответственно. Для полиномиальных функций $U(\pi)$ и $V(\pi)$ гамильтониан (32.35) является полиномиальным при отсутствии скалярных полей. Добавление скалярных полей нарушает полиномиальность из-за присутствия компоненты метрики g_{11} в знаменателе и возможной неминимальной связи $\rho(\pi)$.

По определению, одновременные скобки Пуассона имеют вид

$$\begin{aligned} [e_1^a, p'_b] &= \delta_b^a \delta(\sigma, \sigma'), \\ [\omega_1, \pi'] &= \delta(\sigma, \sigma'), \\ [X, P'] &= \delta(\sigma, \sigma'), \end{aligned} \quad (32.38)$$

где штрих обозначает, что соответствующее поле рассматривается в точке σ' .

Из определения обобщенных импульсов следует, что равенства (32.30) и (32.32) представляют собой первичные связи в двумерной гравитации. Для самосогласованности уравнений движения первичные связи (32.30) должны сохраняться со временем. Это приводит к следующим условиям:

$$\partial_0 \pi^0 = [\pi^0, H] = -G = 0, \quad (32.39)$$

$$\partial_0 p^0_a = [p^0_a, H] = -G_a = 0. \quad (32.40)$$

Полученные равенства представляют собой вторичные связи. Таким образом, гамильтониан (32.35) представляет собой линейную комбинацию вторичных связей.

Прямые довольно громоздкие вычисления показывают, что вторичные связи образуют замкнутую алгебру относительно скобок Пуассона:

$$[G_a, G'_b] = \varepsilon_{ab} \left[U p^c G_c + \left(\frac{1}{2} p^c p_c U' + V' - \frac{\rho'}{\rho \sqrt{|g|}} L_x \right) G \right] \delta(\sigma, \sigma'), \quad (32.41)$$

$$[G_a, G'] = \varepsilon_a^b G_b \delta(\sigma, \sigma'), \quad (32.42)$$

$$[G, G'] = 0, \quad (32.43)$$

где U' , V' и ρ' обозначают производные по аргументу π и L_x – лагранжиан скалярного поля (32.17), выраженный через канонические переменные:

$$L_x = -\frac{1}{2} \sqrt{|g|} \left(-\frac{1}{\rho g_{11}} P^2 + \frac{\rho}{g_{11}} \partial_1 X^2 - m^2 X^2 \right).$$

Скобки Пуассона первичных связей между собой и с вторичными связями, очевидно, равны нулю. Поэтому двумерная гравитация с действием (32.15) содержит шесть связей первого рода и других связей нет. Таким образом, гамильтониан (32.35) представляет собой линейную комбинацию трех вторичных связей первого рода.

На этом этапе мы в состоянии вычислить число независимых физических (распространяющихся) степеней свободы согласно общим правилам, сформулированным

в разделе 22. Предположим, что имеется D скалярных полей. Гравитационное поле в формализме второго порядка описывается репером e_α^a (четыре компоненты) и лоренцевой связностью ω_α (две компоненты). Поэтому число физических степеней свободы равно $D + 4 + 2 - 6 = D$. Чистая же двумерная гравитация (без полей материи) вообще не имеет физических степеней свободы. Это обстоятельство является причиной интегрируемости двумерной гравитации.

Заметим, что алгебра скобок Пуассона (32.41)–(32.43) замыкается на δ -функции, и не содержит их производных в отличие от алгебры скобок Пуассона в общей теории относительности (24.71)–(24.73). Кроме того “структурные функции” зависят не только от полей, но и от функций $U(\pi)$, $V(\pi)$ и $\rho(\pi)$, которые входят в действие. В этом смысле алгебра скобок Пуассона зависит от динамики полей. Напомним, что в обычной гамильтоновой формулировке моделей гравитации, инвариантных относительно общих преобразований координат, алгебра скобок Пуассона замыкается на первые производные от δ -функции.

Алгебра скобок Пуассона (32.41)–(32.43) имеет (не инвариантную) подалгебру, генерируемую связью с векторным индексом

$$\tilde{G}_a := G_a + k_a \omega_1 G, \quad (32.44)$$

где мы ввели светоподобный вектор

$$k_a := \frac{e_{1a} + e_1^b \varepsilon_{ba}}{g_{11}}, \quad k^a k_a = 0, \quad (32.45)$$

с компонентами

$$k_0 = k_1 = \frac{e_{10} + e_{11}}{g_{11}}.$$

Прямые вычисления приводят к следующим скобкам Пуассона

$$[\tilde{G}_a, \tilde{G}'_b] = \varepsilon_{ab}(Up^c + \omega_1 k^c) \tilde{G}_c \delta(\sigma, \sigma'), \quad (32.46)$$

$$[\tilde{G}_a, G'] = \varepsilon_a^b \tilde{G}_b \delta(\sigma, \sigma') - k_a G \delta'(\sigma, \sigma'). \quad (32.47)$$

где введена производная δ -функции

$$\delta'(\sigma', \sigma) := \frac{\partial}{\partial \sigma'} \delta(\sigma', \sigma).$$

Первая скобка Пуассона (32.46) показывает, что связи $\tilde{G}_a$ генерируют подалгебру, а вторая (32.47) – что подалгебра не инвариантна.

Напомним, что конформная алгебра скобок Пуассона генерируется двумя скалярными функциями (относительно лоренцевых вращений) H_0 и H_1 и имеет вид

$$\begin{aligned} [H_0, H'_0] &= -(H_1 + H'_1) \delta'(\sigma', \sigma), \\ [H_0, H'_1] &= -(H_0 + H'_0) \delta'(\sigma', \sigma), \\ [H_1, H'_1] &= -(H_1 + H'_1) \delta'(\sigma', \sigma). \end{aligned} \quad (32.48)$$

Эта алгебра связей возникает, например, в теории струн. В двумерной гравитации конформной алгебре (32.48) удовлетворяют связи

$$H_0 := -e_1^a \varepsilon_a^b \tilde{G}_b = -e_1^a \varepsilon_a^b G_b + \omega_1 G, \quad (32.49)$$

$$H_1 := e_1^a \tilde{G}_a = e_1^a G_a + \omega_1 G. \quad (32.50)$$

Связи H_1 и H_0 являются лоренц инвариантными проекциями векторной связи $\tilde{G}_a$ на направления, параллельное и перпендикулярное вектору e_1^a , соответственно. Оставшиеся скобки Пуассона имеют вид

$$[H_0, G'] = [H_1, G'] = -G\delta'(\sigma', \sigma). \quad (32.51)$$

Связи H_0 и H_1 образуют хорошо известную конформную алгебру (32.48). Полная же алгебра связей (32.48)–(32.51) представляет собой “полупрямое произведение” конформной алгебры с инвариантной абелевой подалгеброй, генерируемой связью G , которая соответствует локальным лоренцевым вращениям [238].

Преобразование связей, обратное к формулам (32.49) и (32.50) выглядит следующим образом

$$G_0 = \frac{1}{g_{11}} [-e_1^1(H_0 - \omega_1 G) + e_1^0(H_1 - \omega_1 G)], \quad (32.52)$$

$$G_1 = \frac{1}{g_{11}} [e_1^0(H_0 - \omega_1 G) - e_1^1(H_1 - \omega_1 G)]. \quad (32.53)$$

В дальнейшем нам понадобится явное выражение для связей H_0 и H_1 через канонические переменные

$$\begin{aligned} H_0 = & -\partial_1 p_a \varepsilon^a_b e_1^b - \omega_1 p_a e_1^a + \omega_1 (-\partial_1 \pi + p_a \varepsilon^a_b e_1^b) - \\ & - g_{11} \left(\frac{1}{2} p^a p_a U + V - \frac{\rho}{2} m^2 X^2 \right) - \frac{1}{2\rho} P^2 - \frac{\rho}{2} \partial_1 X^2, \end{aligned} \quad (32.54)$$

$$H_1 = -e_1^a \partial_1 p_a - \omega_1 \partial_1 \pi + P \partial_1 X. \quad (32.55)$$

Заметим, что третье слагаемое в H_0 пропорционально связи G (32.36).

Из выражений для связей (32.54), (32.55) следует их геометрический смысл. Рассмотрим вариации полей, которые генерируются связью H_1 . По определению, эта связь является генератором калибровочного преобразования

$$\delta F := -[F, \int d\sigma \epsilon^1 H_1],$$

где $F(\omega_1, \pi, \dots)$ – дифференцируемая функция канонических переменных и $\epsilon^1(x) \ll 1$ – параметр преобразования. Простые вычисления приводят к равенствам

$$\begin{aligned} \delta e_1^a &= -\partial_1 \epsilon^1 e_1^a - \epsilon^1 \partial_1 e_1^a, & \delta p_a &= -\epsilon^1 \partial_1 p_a, & \delta X &= -\epsilon^1 \partial_1 X, \\ \delta \omega_1 &= -\partial_1 \epsilon^1 \omega_1 - \epsilon^1 \partial_1 \omega_1, & \delta \pi &= -\epsilon^1 \partial_1 \pi, & \delta P &= -\partial_1 \epsilon^1 P - \epsilon^1 \partial_1 P. \end{aligned}$$

Сравнение этих преобразований с формулами (32.28) показывает, что связь H_1 генерирует общие преобразования координат вдоль пространственного направления σ .

Аналогично, связь G (32.36) генерирует калибровочные преобразования

$$\delta F := [F, \int d\sigma \omega G]$$

с параметром $\omega(x) \ll 1$ (знак правой части не играет роли, т.к. его можно отнести к параметру преобразования). Простые вычисления дают следующие калибровочные преобразования:

$$\begin{aligned} \delta e_1^a &= -e_1^b \varepsilon_b^a \omega, & \delta p_a &= \varepsilon_a^b p_b \omega, & \delta X &= 0, \\ \delta \omega_1 &= \partial_1 \omega, & \delta \pi &= 0, & \delta P &= 0. \end{aligned}$$

Из сравнения полученных формул с преобразованиями (32.26) следует, что связь G генерирует локальные лоренцевы вращения.

Таким образом, двумерная гравитация содержит три вторичные связи первого рода. Связь H_0 является динамической и, как мы увидим в дальнейшем, по существу определяет динамику системы. Связи H_1 и G являются кинематическими и генерируют калибровочные преобразования.

Если рассматривается только бозонная струна ($\rho = 1, m = 0$), то связи принимают хорошо известный вид

$$\begin{aligned} H_0 &\mapsto -\frac{1}{2}P^2 - \frac{1}{2}\partial_1 X^2, \\ H_1 &\mapsto P\partial_1 X, \end{aligned}$$

которые удовлетворяют конформной алгебре (32.48).

32.3.1 Каноническое преобразование

На классическом уровне две гамильтоновы модели, связанные каноническим преобразованием являются эквивалентными. На квантовом уровне это в общем случае не так: существуют канонические преобразования, которые после квантования приводят к неэквивалентным квантовым теориям. При этом особое внимание следует уделять нелинейным каноническим преобразованиям. Это значит, что правильный выбор канонических переменных чрезвычайно важен. До сих пор не известен правильный выбор канонических переменных, в которых гравитацию можно было бы проквантовать. Поэтому мы вправе выбирать те или иные канонические переменные из соображений удобства. В настоящем разделе мы введем новые канонические переменные, в которых связи решены в общем виде в разделе ??.

Совершим каноническое преобразование репера, оставив другие канонические переменные без изменения,

$$e_1^a, p_a \mapsto q, q_\perp, p, p_\perp,$$

которое в явном виде выделяет лоренцев угол поворота $q_\perp$ и упрощает многие формулы [238]. Рассмотрим производящую функцию канонического преобразования, зависящую от старых координат и новых импульсов (см. раздел 21.12),

$$F := \frac{1}{2} \int d\sigma \left(p \ln |g_{11}| + p_\perp \ln \left| \frac{e_1^0 + e_1^1}{e_1^0 - e_1^1} \right| \right). \quad (32.56)$$

Вариация F по старым координатам приводит к соотношению между старыми и новыми импульсами:

$$p_a = p \frac{e_{1a}}{g_{11}} + p_\perp \frac{e_1^b \varepsilon_{ba}}{g_{11}}. \quad (32.57)$$

Отсюда вытекают выражения для новых импульсов

$$p = e_1^a p_a, \quad p_\perp = p_a \varepsilon^a_b e_1^b,$$

т.е. функции p и $p_\perp$ являются, соответственно, проекциями импульса p_a на вектор e_1^a и перпендикулярное направление. Вариация производящего функционала (32.56) по импульсам дает соотношения между координатами:

$$\begin{aligned} e_1^0 &= e^q \operatorname{sh} q_\perp, \\ e_1^1 &= e^q \operatorname{ch} q_\perp. \end{aligned} \quad (32.58)$$

Для того, чтобы отбросить знак модуля во втором слагаемом в равенстве (32.56) мы предположим, для определенности, что

$$e_1^1 > 0 \quad \text{и} \quad e_1^1 > e_1^0. \quad (32.59)$$

Тогда пространственная компонента метрики равна

$$g_{11} = -e^{2q}$$

и всегда отрицательна, что соответствует нашему выбору координат на $\mathbb{M}$.

Мы видим, что новая координата $q_\perp$ совпадает с углом лоренцева поворота, а q параметризует длину вектора e_1^a .

Квадрат старого импульса в новых переменных имеет вид

$$p^a p_a = (p_\perp^2 - p^2) e^{-2q}. \quad (32.60)$$

Вторичные связи двумерной гравитации в новых канонических переменных выглядят следующим образом

$$\begin{aligned} H_0 = & -\partial_1 p_\perp + p \partial_1 q_\perp + p_\perp \partial_1 q - \omega_1 p + \omega_1 (-\partial_1 \pi + p_\perp) + \\ & + \frac{1}{2} (p_\perp^2 - p^2) U + e^{2q} \left(V - \frac{\rho}{2} m^2 X^2 \right) - \frac{1}{2\rho} P^2 - \frac{\rho}{2} \partial_1 X^2, \end{aligned} \quad (32.61)$$

$$H_1 = -\partial_1 p - \omega_1 \partial_1 \pi + p \partial_1 q + p_\perp \partial_1 q_\perp + P \partial_1 X, \quad (32.62)$$

$$G = -\partial_1 \pi + p_\perp. \quad (32.63)$$

Если $U = \text{const}$, то квадратичная часть импульсов $\frac{1}{2}(p_\perp^2 - p^2)$ в динамической связи H_0 имеет тот же вид, что и для двух свободных частиц со знаконеопределенной метрикой. Это значит, что каноническое преобразование “сдвигает” неполиномиальность с кинетического слагаемого в потенциал.

32.4 Дилатонизация

В предыдущих разделах мы рассматривали нетривиальные модели гравитации в двух измерениях, основанные на геометрии Римана–Картана. Однако существуют и распространены другие нетривиальные модели двумерной гравитации, которые основаны на псевдоримановой геометрии, но содержат дополнительное скалярное поле – поле *дилатона*. В настоящем разделе мы докажем эквивалентность двумерной гравитации с кручением и двумерной дилатонной гравитации [224, 239, 240].

Если лоренцева связность не входит в лагранжиан полей материи, тогда ее можно исключить из теории, решив соответствующие ей уравнения движения. В результате мы получим обобщенную дилатонную модель двумерной гравитации, в которой роль дилатона играет импульс $\pi(x)$, который сопряжен к пространственной компоненте лоренцевой связности ω_1 .

Прежде всего заметим, что тождество (6.96) в двумерном случае принимает вид

$$\sqrt{|g|} R = \sqrt{|g|} \tilde{R} + \partial_\alpha (\sqrt{|g|} \varepsilon^\alpha{}_\beta T^{*\beta}), \quad (32.64)$$

где $\tilde{R}$ – скалярная кривизна поверхности, построенная только по символам Кристоффеля (при нулевом кручении). После замены $R \mapsto \tilde{R}$ и интегрирования по частям гравитационная часть лагранжиана (32.16) принимает вид

$$L_G = -\frac{1}{2} \sqrt{|g|} \pi \tilde{R} + \frac{1}{2} \partial_\alpha \pi \varepsilon^\alpha{}_a \hat{T}^{*a} - \frac{1}{2} p_a \hat{T}^{*a} - \sqrt{|g|} \left(\frac{1}{2} p_a p^a U + V \right). \quad (32.65)$$

Теперь лоренцева связность входит в лагранжиан только через псевдослед тензора кручения $\hat{T}^{*a}$, и они связаны между собой обратимым алгебраическим соотношением

$$\omega_\alpha = e_{\alpha a} \left(\frac{1}{2} T^{*a} - \varepsilon^{\beta\gamma} \partial_\beta e_\gamma^a \right). \quad (32.66)$$

Поэтому вместо лоренцевой связности можно выбрать компоненты кручения T^{*a} в качестве новых переменных, по которым производится варьирование. В результате возникает алгебраическое уравнение на импульс p_a , которое имеет решение

$$p_a = \partial_\alpha \pi \varepsilon^\alpha_a.$$

Подстановка этого решения в исходное действие приводит к лагранжиану двумерной дилатонной гравитации

$$L_D := -\frac{1}{2} \sqrt{|g|} \pi \tilde{R} + \frac{1}{2} \sqrt{|g|} \partial \pi^2 U - \sqrt{|g|} V, \quad (32.67)$$

где использовано сокращенное обозначение $\partial^2 \pi := g^{\alpha\beta} \partial_\alpha \pi \partial_\beta \pi$. Мы видим, что поле π является полем дилатона, и лагранжиан зависит от двух произвольных функций $U(\pi)$ и $V(\pi)$, которые определяют кинетическую и потенциальную части лагранжиана дилатонного поля.

Отметим, что первое слагаемое в лагранжиане (32.67) не равно дивергенции, т.к. скалярная кривизна умножается на поле дилатона, что соответствует неминимальной связи дилатона с гравитационным полем.

Таким образом, если лагранжиан полей материи не содержит лоренцевой связности, то двумерная гравитация с кручением эквивалентна двумерной дилатонной гравитации. На самом деле мы доказали, что если исходные переменные e_α^a , ω_α , p_a , π и X удовлетворяют уравнениям движения (32.18)–(32.22), то переменные $g_{\alpha\beta}$, π и X удовлетворяют уравнениям движения, вытекающим из лагранжиана $L_D + L_X$. Обратное утверждение следующее. Пусть переменные $g_{\alpha\beta}$, π и X удовлетворяют уравнениям движения для лагранжиана $L_D + L_X$. После этого решим алгебраическое уравнение $e_\alpha^a e_\beta^b \eta_{ab} = g_{\alpha\beta}$ для репера (решение определено с точностью до локальных лоренцевых вращений), построим импульс p_a , используя уравнение (32.20), и решим уравнения (32.19) и (32.66) для T^{*a} и ω_α , соответственно. Тогда исходные уравнения (32.18)–(32.22) будут удовлетворены.

Эквивалентность моделей гравитации является не только локальной, но и глобальной, поскольку преобразование переменных (32.66) невырождено, что следует из невырожденности репера.

Доказанная эквивалентность двумерных моделей гравитации дает геометрический смысл дилатонному полю: это импульс, сопряженный пространственной компоненте лоренцевой связности ω_1 .

Чтобы получить уравнения движения для двумерной дилатонной гравитации удобно воспользоваться формулой (23.21). Вариация лагранжиана (32.67) по метрике и дилатонному полю дает следующие уравнения движения:

$$\frac{1}{\sqrt{|g|}} \frac{\delta L_D}{\delta g_{\alpha\beta}} : -\frac{1}{2} (g^{\alpha\beta} \tilde{\square} \pi - \tilde{\nabla}^\alpha \tilde{\nabla}^\beta \pi) - \frac{1}{2} \tilde{\nabla}^\alpha \pi \tilde{\nabla}^\beta \pi U + \frac{1}{2} g^{\alpha\beta} \left(\frac{1}{2} \partial \pi^2 U - V \right) = 0, \quad (32.68)$$

$$\frac{1}{\sqrt{|g|}} \frac{\delta L_D}{\delta \pi} : -\frac{1}{2} R - \tilde{\square} \pi U - \frac{1}{2} \partial \pi^2 U' - V' = 0, \quad (32.69)$$

where $\tilde{\square} := g^{\alpha\beta} \tilde{\nabla}_\alpha \tilde{\nabla}_\beta$ – двумерный волновой оператор. В настоящее время неизвестно, как решить непосредственно данную систему нелинейных уравнений для произвольных функций U и V и без фиксирования системы координат. В следующем разделе мы решим эквивалентную систему уравнений двумерной гравитации с кручением.

32.5 Размерности полей

Вычисления, проведенные в следующих разделах, сложны и громоздки. Поэтому мы припишем в настоящем разделе всем полям и константам связи определенные размерности. Это позволяет контролировать вычисления на каждом этапе, подсчитывая и сравнивая размерности получаемых выражений.

По определению, все координаты имеют размерность длины, а компоненты репера, метрики и действие безразмерны:

$$[x^\alpha] := l, \quad [g_{\alpha\beta}] := 1, \quad [e_\alpha^a] := 1, \quad [S] := 1.$$

Кроме этого мы предполагаем, что компоненты лоренцевой связности имеют ту же размерность, что и частная производная ∂_α :

$$[\omega_\alpha] := [\partial_\alpha] = l^{-1}.$$

Отсюда вытекают размерности компонент тензора кручения и скалярной кривизны:

$$[T^{*a}] = l^{-1}, \quad [R] = l^{-2}.$$

Мы предположим также, что скалярное поле имеет ту же размерность, что и в четырехмерном пространстве-времени,

$$[X] := l^{-1}, \tag{32.70}$$

потому что в разделе ?? будет проанализирована сферически симметричная общая теория относительности. После этого размерности всех остальных полей и констант связи однозначно определены.

Лагранжиан в двумерном случае имеет размерность $[L] = l^{-2}$. Поэтому размерность обобщенных импульсов равна

$$[\pi] = 1, \quad [p_a] = l^{-1}. \tag{32.71}$$

Произвольные функции $U(\pi)$ и $V(\pi)$, входящие в лагранжиан (32.16), имеют размерности

$$[U] = 1, \quad [V] = l^{-2}. \tag{32.72}$$

Поэтому можно было бы ввести размерную константу перед $V(\pi)$. Размерность массы скалярного (32.17) поля одинакова во всех размерностях пространства-времени

$$[m] = l^{-1}.$$

Размерность функции ρ определена размерностью скалярного поля (32.70) и безразмерностью действия

$$[\rho] = l^2.$$

Импульс скалярного поля безразмерен, $[P] = 1$.

Четырехмерная гравитационная константа связи κ и космологическая постоянная Λ имеют размерности

$$[\kappa] = l^{-2}, \quad [\Lambda] = l^{-4}.$$

Минимально взаимодействующие скалярные поля в двух измерениях имели бы размерность

$$[X] = 1, \quad [\rho] = 1.$$

Эти размерности удобны при анализе бозонной струны с динамической геометрией, когда гравитационный лагранжиан (32.16) добавляется к действию бозонной струны.

32.6 Общее решение без полей материи

Важным свойством двумерной модели гравитации, описываемой только лагранжианом (32.16), является ее интегрируемость. Решение уравнений движения имеет длинную историю. Первоначально, уравнения движения были решены для квадратичной модели в конформной калибровке [241, 235] и калибровке светового конуса [242]. В статьях [243, 244, 245] решение для квадратичной модели было найдено без фиксирования калибровки. После этого решение уравнений движения было упрощено и обобщено в статьях [246, 240]. В настоящем разделе мы приведем конечный результат и выпишем общее решение уравнений движения для произвольных функций U и V в отсутствие полей материи. Решение естественным образом записывается в канонической формулировке модели. Это решение имеет один вектор Киллинга (аналог теоремы Бирхгоффа). Поэтому, используя метод конформных блоков, описанный в главе 31, можно построить все глобальные (максимально продолженные вдоль геодезических) решения двумерных моделей гравитации. В настоящем разделе мы полагаем $X = 0$.

32.6.1 Локальное решение

Интегрирование уравнений движения наиболее просто проводится для конусных компонент векторов в касательном пространстве

$$p_{\pm} := \frac{1}{\sqrt{2}}(p_0 \pm p_1), \quad e_{\alpha}^{\pm} := \frac{1}{\sqrt{2}}(e_{\alpha}^0 \pm e_{\alpha}^1). \quad (32.73)$$

Метрика Лоренца и антисимметричный тензор второго ранга имеют следующие конусные компоненты

$$\eta_{\pm\pm} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \varepsilon_{\pm\pm} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (32.74)$$

Подъем и опускание конусных индексов проводится по правилам $p^+ = p_-$ и $p^- = p_+$. Тогда полная система уравнений движения (32.18)–(32.21) примет вид

$$-\frac{1}{2}R - p_+p_-U' - V' = 0, \quad (32.75)$$

$$-\frac{1}{2}T^{*+} - p^+U = 0, \quad (32.76)$$

$$-\frac{1}{2}T^{*-} - p^-U = 0, \quad (32.77)$$

$$\partial_\alpha \pi - e_\alpha^+ p_+ + e_\alpha^- p_- = 0, \quad (32.78)$$

$$\partial_\alpha p_+ + \omega_\alpha p_+ + e_\alpha^-(p_+p_-U + V) = 0, \quad (32.79)$$

$$\partial_\alpha p_- - \omega_\alpha p_- - e_\alpha^+(p_+p_-U + V) = 0, \quad (32.80)$$

где

$$T^{*\pm} = 2\varepsilon^{\alpha\beta}(\partial_\alpha \mp \omega_\alpha)e_\beta^\pm. \quad (32.81)$$

Любое локальное решение приведенной системы уравнений для $U \neq 0$ и $V \neq 0$ принадлежит одному из двух классов. Первый класс соответствует нулевому импульсу $p_- = 0$ или, что эквивалентно, $p_+ = 0$.

Теорема 32.6.1. *Если в области $\mathbb{U} \subset \mathbb{M}$ импульс $p_- = 0$ или, что эквивалентно $p_+ = 0$, то общее решение уравнений движения (32.75)–(32.81) описывает поверхность нулевого кручения, $T^{*\alpha} = 0$, и постоянной кривизны, $R = \text{const}$. Поле дилатона π при этом постоянно в $\mathbb{U}$, и его значение определяется уравнением*

$$V(\pi) = 0. \quad (32.82)$$

Значение скалярной кривизны определяется равенством

$$R = -2V'(\pi). \quad (32.83)$$

Доказательство. Для $p_- = 0$ поле дилатона постоянно $\pi = \text{const}$ причем постоянная определяется уравнением (32.82), что следует из уравнения (32.80). Тогда $p_+ = 0$, благодаря уравнению (32.78), и уравнение (32.79) выполнено. То есть условия $p_- = 0$ и $p_+ = 0$ эквивалентны. Оставшиеся уравнения (32.75)–(32.77) показывают, что поверхность $\mathbb{U} \subset \mathbb{M}$ (пространство-время) является поверхностью с нулевым кручением и постоянной кривизной

$$T^{*\pm} = 0, \quad R = -2V'(\pi) = \text{const}. \quad \square \quad (32.84)$$

Лоренцевы поверхности постоянной кривизны были описаны в разделе (29.3). Для этих поверхностей можно восстановить репер e_α^a и лоренцеву связность ω_α , которые входят в систему уравнений (32.75)–(32.81), однозначно с точностью до локальных лоренцевых вращений.

Из теоремы 32.6.1 следует, что двумерные модели гравитации, в основу которых положено постоянство кривизны [247, 248, 249], образуют один из классов решений уравнений движения двумерной гравитации с кручением.

Другой класс решений соответствует нетривиальным импульсам $p_- \neq 0$ или $p_+ \neq 0$ и отличному от нуля кручению. Для определенности, предположим, что $p_- \neq 0$ в некоторой области $\mathbb{U} \subset \mathbb{M}$. Случай $p_+ \neq 0$ рассматривается аналогично.

Интегрируемость уравнений движения обязана двум нетривиальным наблюдениям.

Предложение 32.6.1. Уравнения движения (32.75)–(32.80) имеют интеграл движения

$$M := p_+ p_- e^{-Q} - W = \frac{1}{2}(p_\perp^2 - p^2) e^{-2q-Q} - W = \text{const}, \quad (32.85)$$

где Q и W являются первообразными:

$$Q(\pi) := \int^\pi ds U(s), \quad W(\pi) := \int^\pi ds V(s) e^{-Q(s)}. \quad (32.86)$$

Доказательство. Равенство

$$\partial_\alpha M = 0.$$

доказывается прямыми вычислениями с использованием уравнений движения (32.78)–(32.80). Постоянные интегрирования в определении (32.86) несущественны, т.к. M является интегралом движения. $\square$

Замечание. Обозначение интеграла движения буквой M не случайно. Для сферически редуцированной общей теории относительности интеграл движения M совпадает с массой в решении Шварцшильда с точностью до постоянного множителя (см. раздел ??). $\square$

Для контроля дальнейших вычислений выпишем размерности

$$[Q] = 1, \quad [W] = [M] = l^{-2},$$

что является следствием размерностей полей (32.72) и (32.71).

Заметим, что интеграл движения зависит только от импульсов p_a и π , и его сохранение является следствием трех уравнений (32.78)–(32.80).

Предложение 32.6.2. 1-форма

$$dx^\alpha f_\alpha := \sqrt{\kappa} \frac{dx^\alpha e_\alpha^+}{p_-} e^Q, \quad p_- \neq 0 \quad (32.87)$$

является замкнутой на уравнениях движения.

В определении (32.87) мы ввели размерную постоянную $\kappa > 0$, $[\kappa] = l^{-2}$, для того, чтобы компоненты формы f_α были безразмерными, $[f_\alpha] = 1$.

Доказательство. Чтобы доказать замкнутость формы (32.87), необходимо доказать равенство

$$\varepsilon^{\alpha\beta} \partial_\alpha f_\beta = 0.$$

Его нетрудно проверить, используя уравнения движения (32.76), (32.78) и (32.80). $\square$

Согласно лемме Пуанкаре из замкнутости 1-формы (32.87) следует, что в области $\mathbb{U}$ форму можно записать в виде градиента

$$f_\alpha = \partial_\alpha f \quad (32.88)$$

от некоторой достаточно гладкой функции f .

Теперь можно выписать общее решение уравнений движения.

Теорема 32.6.2. Если в области $\mathbb{U} \subset \mathbb{M}$ импульс $p_- \neq 0$, то общее решение уравнений движения (32.75)–(32.80) имеет вид

$$e_\alpha^+ = \frac{1}{\sqrt{\kappa}} p_- e^{-Q} \partial_\alpha f, \quad (32.89)$$

$$e_\alpha^- = \frac{1}{p_-} \left[\frac{1}{\sqrt{\kappa}} (M + W) \partial_\alpha f - \partial_\alpha \pi \right], \quad (32.90)$$

$$\omega_\alpha = \frac{\partial_\alpha p_-}{p_-} - \frac{1}{\sqrt{\kappa}} [(M + W)U + e^{-Q}V] \partial_\alpha f, \quad (32.91)$$

$$p_+ = \frac{1}{p_-} (M + W) e^Q, \quad (32.92)$$

где M – постоянная; функции Q , W определены равенствами (32.86), f – произвольная функция с отличным от нуля градиентом, $\partial_\alpha f \neq 0$ и функции π и $p_- \neq 0$ произвольны.

Доказательство. Выражение компонент репера (32.89) является следствием уравнений (32.87) и (32.88). Выражение (32.90) следует из уравнений (32.78) и (32.85). Выражение для лоренцевой связности (32.91) вытекает из уравнений (32.80) и (32.85). Выражение для импульса (32.92) является простым переписыванием равенства (32.85). Последний шаг доказательства заключается в проверке того, что оставшиеся уравнения движения (32.75)–(32.77) и (32.79) удовлетворяются для произвольных функций f , π и p_- . Это проверяется прямыми вычислениями и является следствием линейной зависимости уравнений движения (32.27) и (32.29). $\square$

Второй класс решений уравнений движения, который описан теоремой 32.6.2, соответствует лоренцевым поверхностям с нетривиальным кручением и непостоянной кривизной, которые проще всего вычислить, воспользовавшись уравнениями (32.75)–(32.77). На этих поверхностях метрика определена репером (32.89), (32.90):

$$ds^2 = 2e_\alpha^+ e_\beta^- dx^\alpha dx^\beta = 2e^{-Q} \left[\frac{1}{\kappa} (M + W) df^2 - \frac{1}{\sqrt{\kappa}} df d\pi \right]. \quad (32.93)$$

Мы видим, что функции f и π аналогичны координатам Эддингтона–Финкельштейна [230, 231] для решения Шварцшильда (см. раздел 34.2). При этом поле дилатона $\pi(x)$ является светоподобной координатой.

Общее решение уравнений движения (32.89)–(32.92) было получено без фиксирования какой-либо калибровки и содержит три произвольные функции $f(x)$, $\pi(x)$ и $p_-(x)$. Из вида решения (32.89)–(32.92) и метрики (32.93) следует, что первые две функции соответствуют инвариантности модели относительно преобразования координат. Поэтому на них необходимо наложить условие

$$d\pi \wedge df \neq 0 \quad \Leftrightarrow \quad \varepsilon^{\alpha\beta} \partial_\alpha \pi \partial_\beta f \neq 0. \quad (32.94)$$

Третья произвольная функция p_- соответствует локальным лоренцевым вращениям и должна быть либо строго положительной $p_- > 0$, либо строго отрицательной $p_- < 0$, что соответствует предположению $p_- \neq 0$ в условии теоремы.

Компоненты метрики (32.93) зависят только от поля дилатона π и поэтому все решения уравнений движения из второго класса имеют векторное поле Киллинга

$$K = \frac{1}{\sqrt{\kappa}} \frac{\partial}{\partial f}. \quad (32.95)$$

Квадрат вектора Киллинга равен

$$K^2 = \frac{2}{\kappa^2} e^{-Q} (M + W) \quad (32.96)$$

и может быть времениподобным или пространственноподобным в зависимости от знака $M + W$. Горизонты определяются равенством $M + W = 0$, где вектор Киллинга светоподобен.

Существование вектора Киллинга для решений уравнений движения из второго класса является обобщением хорошо известной теоремы Бирхгоффа в общей теории относительности, утверждающей, что у любого сферически симметричного решения вакуумных уравнений Эйнштейна существует дополнительный вектор Киллинга, не связанный с вращательной симметрией.

Общее решение (32.89)–(32.92) в новых канонических переменных, введенных в разделе 32.3.1, имеет вид

$$p = \frac{2}{\sqrt{\kappa}} (M + W) \partial_1 f - \partial_1 \pi, \quad (32.97)$$

$$p_{\perp} = \partial_1 \pi, \quad (32.98)$$

$$e^{2q} = \frac{2}{\kappa} \partial_1 f [\sqrt{\kappa} \partial_1 \pi - (M + W) \partial_1 f] e^{-Q}, \quad (32.99)$$

$$e^{2q_{\perp}} = \frac{p_{\perp}^2 \partial_1 f}{\sqrt{\kappa} \partial_1 \pi - (M + W) \partial_1 f} e^{-Q}. \quad (32.100)$$

Это решение определено только для тех функций π и f , которые обеспечивают положительность правой части равенств (32.99) и (32.100) в соответствии со сделанным ранее предположением (32.59).

Решение уравнений движения в теореме 32.6.2 получено при условии $p_{-} \neq 0$. Если в области $\mathbb{U} \subset \mathbb{M}$ выполнено другое условие $p_{+} \neq 0$, то функция (32.85) все еще будет интегралом движения, однако 1-форму (32.87) необходимо заменить, т.к. функция p_{-} может иметь нули. Нетрудно проверить, что 1-форма

$$df := \sqrt{\kappa} \frac{dx^{\alpha} e_{\alpha}^{-}}{p_{+}} e^Q$$

является замкнутой на уравнениях движения (32.77), (32.78) и (32.79). Аналогичные вычисления показывают, что в случае $p_{+} \neq 0$ общее решение имеет вид

$$e_{\alpha}^{-} = \frac{1}{\sqrt{\kappa}} p_{+} e^{-Q} \partial_{\alpha} f, \quad (32.101)$$

$$e_{\alpha}^{+} = \frac{1}{p_{+}} \left[\frac{1}{\sqrt{\kappa}} (M + W) \partial_{\alpha} f + \partial_{\alpha} \pi \right], \quad (32.102)$$

$$\omega_{\alpha} = -\frac{\partial_{\alpha} p_{+}}{p_{+}} - \frac{1}{\sqrt{\kappa}} [(M + W)U + e^{-Q}V] \partial_{\alpha} f, \quad (32.103)$$

$$p_{-} = \frac{1}{p_{+}} (M + W) e^Q, \quad (32.104)$$

где произвольная функция $p_{+}(x) \neq 0$ параметризует решение (вместо p_{-}). Соответствующая метрика отличается от метрики (32.93) знаком перед вторым слагаемым

$$ds^2 = 2 e^{-Q} \left[\frac{1}{\kappa} (M + W) df^2 + \frac{1}{\sqrt{\kappa}} df d\pi \right]. \quad (32.105)$$

Различие можно устранить переопределением $f \mapsto -f$.

Второй класс решений уравнений движения описывает поверхности непостоянной кривизны и нетривиального кручения. Из уравнения (32.75) следует выражение для скалярной кривизны в этом случае

$$R = -2(p^+ p^- U' + V') = -2[(M + W) e^Q U' + V'] . \quad (32.106)$$

Квадрат тензора кручения, который является геометрическим инвариантом, определяется уравнениями (32.76), (32.77):

$$T^{*a} T^*_{\ a} = 2T^{*-+} T^{*-} = 8p^+ p^- U^2 = 8(M + W) e^Q U^2 . \quad (32.107)$$

Хотя общее решение уравнений движения из второго класса определяется тремя произвольными функциями π , f и p_- (или p_+), в правых частях равенств (32.106), (32.107) стоят функции только от π . Это является следствием наличия векторного поля Киллинга $\partial/\partial f$.

В заключение раздела рассмотрим отношение интеграла движения M к вторичным связям, имеющимся в теории. Уравнение

$$M = \text{const} \quad (32.108)$$

представляет собой связь первого рода. В этом можно убедиться после вычисления пространственной производной:

$$\partial_1 M = e^{-Q} \left[-p^a G_a + \left(\frac{1}{2} p^a p_a U + V \right) G \right] . \quad (32.109)$$

Скобка Пуассона связи (32.108) с вторичными связями обращается в нуль

$$[M, G'_a] = [M, G'] = 0.$$

Следовательно, связь M принадлежит центру алгебры (32.41)–(32.43). Прямые вычисления приводят к следующим скобкам Пуассона связи M с другими связями:

$$\begin{aligned} [M, \tilde{G}'_a] &= \left[k^c p_c \omega_1 + \frac{1}{2} p^c p_c U + V \right] k_a e^{-Q} G \delta, \\ [M, \tilde{G}'_0] &= [M, \tilde{G}'_1] = \left[(p_\perp - p) e^{-2q} \omega_1 + \frac{1}{2} (p_\perp^2 - p^2) e^{-2q} U + V \right] e^{-(q+q_\perp+Q)} G \delta, \\ [M, H'_0] &= \left[-\frac{e_1^a p_a}{g_{11}} H_0 - \frac{e_1^a \varepsilon_a^b p_b}{g_{11}} H_1 + \left(k^a p_a \omega_1 + \frac{1}{2} p^a p_a U + V \right) G \right] e^{-Q} \delta = \\ &= \left[p e^{-2q} H_0 - p_\perp e^{-2q} H_1 + \left((p_\perp - p) e^{-2q} \omega_1 + \frac{1}{2} (p_\perp^2 - p^2) e^{-2q} U + V \right) G \right] e^{-Q} \delta, \\ [M, H'_1] &= \left[\frac{e_1^a \varepsilon_a^b p_b}{g_{11}} H_0 + \frac{e_1^a p_a}{g_{11}} H_1 + \left(-k^a p_a \omega_1 + \frac{1}{2} p^a p_a U + V \right) G \right] e^{-Q} \delta = \\ &= \left[p_\perp e^{-2q} H_0 - p e^{-2q} H_1 + \left(-(p_\perp - p) e^{-2q} \omega_1 + \frac{1}{2} (p_\perp^2 - p^2) e^{-2q} U + V \right) G \right] e^{-Q} \delta, \end{aligned}$$

где, для краткости, мы опустили аргументы δ -функции в правых частях. Выше мы выписали скобки Пуассона в переменных до и после канонического преобразования для сравнения. Мы видим, что связь (32.108) не образует замкнутой подалгебры со связями H_0 и H_1 .

32.6.2 Глобальные решения

В предыдущем разделе мы нашли все локальные решения двумерной гравитации без полей материи. Решения из первого класса описывают поверхности постоянной кривизны и нулевого кручения. Глобальная структура этих поверхностей хорошо известна: это либо однополостный гиперboloид, описанный в разделе 29.3, либо плоскость Минковского $\mathbb{R}^{1,1}$. Решения из второго класса описывает лоренцевы поверхности ненулевого кручения и непостоянной кривизны. На этих поверхностях определена метрика, имеющая один вектор Киллинга. Поэтому для построения глобальных решений (поверхностей, максимально продолженных вдоль экстремалей и геодезических) можно воспользоваться методом конформных блоков, описанным в главе 30. Однако для его применения необходимо преобразовать полученную ранее метрику к нужному виду. Этой задачей мы займемся в настоящем разделе.

Для применения метода конформных блоков двумерную метрику необходимо переписать в конформно плоском виде. Это необходимо сделать в каждой области, где определено решение уравнений движения.

Запишем сначала метрику (32.93) в диагональном виде, который будет полезен для сравнения со случаем наличия скалярного поля. В области, где $M + W > 0$ мы оставим координату π без изменения, но заменим координату f :

$$(\pi, f) \mapsto (\pi, \tau), \quad f := \frac{1}{2}\tau + g(\pi), \quad (32.110)$$

где функция $g(\pi)$ определена уравнением

$$g' = \frac{\sqrt{\kappa}}{2(M + W)}. \quad (32.111)$$

Это преобразование определено по крайней мере между горизонтами, которые определяются уравнением $M + W = 0$. Тогда метрика (32.93) станет диагональной

$$ds^2 = \frac{1}{2} e^{-Q} \left[\frac{1}{\kappa} (M + W) d\tau^2 - \frac{d\pi^2}{M + W} \right]. \quad (32.112)$$

В области $M + W > 0$ координата τ является времениподобной, а π – пространственноподобной.

Эту метрику легко переписать в конформно плоском виде, который нужен для построения глобальных решений. С этой целью введем пространственную координату σ , определенную уравнением

$$\frac{d\pi}{d\sigma} = \pm \frac{M + W}{\sqrt{\kappa}},$$

и оставив координату τ без изменения. В результате получим конформно плоскую метрику

$$ds^2 = \frac{1}{2\kappa} e^{-Q} (M + W) (d\tau^2 - d\sigma^2), \quad M + W > 0. \quad (32.113)$$

Теперь введем инвариантную переменную $\hat{q}(\pi)$, связанную с π обыкновенным дифференциальным уравнением

$$\frac{d\hat{q}}{d\pi} = \frac{1}{\sqrt{\kappa}} e^{-Q}. \quad (32.114)$$

Тогда метрика (32.113) примет следующую форму

$$ds^2 = \Phi(\hat{q}) (d\tau^2 - d\sigma^2), \quad \Phi(\hat{q}) > 0, \quad (32.115)$$

где

$$\pm \frac{d\hat{q}}{d\sigma} = \Phi(\hat{q}), \quad (32.116)$$

и конформный множитель имеет вид

$$\Phi(\hat{q}) = \frac{1}{2\kappa} e^{-Q}(M + W). \quad (32.117)$$

Это и есть необходимый вид метрики в статической области $M + W > 0$. Данная область является статической, т.к. конформный множитель Φ не зависит от времени τ .

В области $M + W < 0$ мы вводим новую пространственную координату σ :

$$(\pi, f) \mapsto (\pi, \sigma), \quad f := \frac{1}{2}\sigma + g(\pi),$$

где функция $g(\pi)$ определена прежним уравнением (32.111). Тогда метрика (32.93) примет диагональный вид

$$ds^2 = -\frac{1}{2} e^{-Q} \left[\frac{d\pi^2}{M + W} - \frac{1}{\kappa} (M + W) d\sigma^2 \right].$$

Поскольку в рассматриваемом случае $M + W < 0$, то координата π является времениподобной, а σ – пространственноподобной. Теперь введем новую временную координату τ с помощью уравнения

$$\frac{d\pi}{d\tau} = \pm \frac{M + W}{\sqrt{\kappa}}.$$

Тогда метрика примет конформно плоский вид

$$ds^2 = -\frac{1}{2\kappa} e^{-Q}(M + W)(d\tau^2 - d\sigma^2), \quad M + W < 0.$$

В терминах инвариантной переменной $\hat{q}(\pi)$ (32.114) эта метрика запишется в виде

$$ds^2 = -\Phi(\hat{q})(d\tau^2 - d\sigma^2), \quad \Phi(\hat{q}) < 0, \quad (32.118)$$

где конформный множитель определен прежним уравнением (32.117) и

$$\pm \frac{d\hat{q}}{d\tau} = -\Phi(\hat{q}).$$

Эта метрика однородна, т.к. конформный множитель зависит только от времени τ .

Решения для метрики в статическом $M + W > 0$ и $M + W < 0$ однородном случае можно объединить

$$ds^2 = |\Phi(\hat{q})| (d\tau^2 - d\sigma^2),$$

$$\left| \frac{d\hat{q}}{d\zeta} \right| = \pm \Phi(\hat{q}),$$

где $\zeta = \sigma$ и $\zeta = \tau$, соответственно, для положительных и отрицательных значений конформного множителя $\Phi(\hat{q})$. Метод конформных блоков был разработан для

метрик именно такого вида. Конечно, глобальная структура решений зависит от функций U и V , определяющих модель. В любом случае можно построить максимально продолженные решения путем анализа нулей и особенностей конформного множителя $\Phi(\hat{q})$.

Глобальные решения строятся путем анализа нулей и особенностей конформного множителя $\Phi(\hat{q})$. Для упрощения вида конформного множителя переменную $\hat{q}$ можно сдвинуть $\hat{q} \mapsto \hat{q} + \text{const}$. При этом вид метрики не изменится. Кроме того, переменную $\hat{q}$ и координаты τ, σ можно одновременно умножить на отличную от нуля постоянную a :

$$\hat{q} \mapsto \hat{q}' := a\hat{q} \quad \text{и} \quad \tau, \sigma \mapsto \tau' := a\tau, \quad \sigma' := a\sigma.$$

В результате метрика умножится на a^2 . Конечно, такое преобразование не изменит глобальной структуры пространства-времени. В частности, можно выбрать $a = -1$. Таким образом, без ограничения общности можно совершить линейное преобразование переменной $\hat{q}$ для упрощения вида конформного множителя $\Phi(\hat{q})$.

32.7 Квадратичная модель

В качестве примера интегрирования уравнений движения двумерной гравитации рассмотрим квадратичную модель (32.13) без скалярного поля [222, 232, 233].

Начнем с локального решения. Для квадратичной модели произвольные функции, входящие в лагранжиан, имеют вид (32.24). Теорема 32.6.1 дает первый класс решений уравнений движения. Условие (32.82) определяет значение поля дилатона

$$V(\pi) = 0 \quad \Leftrightarrow \quad \pi = \pm 2\sqrt{\lambda\gamma}.$$

Далее, из равенств (32.84) находим кручение и кривизну:

$$T^{*a} = 0, \quad R = \pm 2\sqrt{\frac{\lambda}{\gamma}}. \quad (32.119)$$

Таким образом, квадратичная модель описывает поверхности нулевого кручения и постоянной кривизны. Такие решения существуют только при $\lambda/\gamma > 0$.

Второй класс решений уравнений движения описывает поверхности с нетривиальным кручением и непостоянной кривизной. Эти решения зависят от двух первообразных (32.86):

$$Q := \int^\pi ds \frac{1}{\beta} = \frac{\pi}{\beta}, \quad (32.120)$$

$$W := \int^\pi ds \left(-\frac{s^2}{4\gamma} + \lambda \right) e^{-\frac{\pi}{\beta}} = \left(\frac{\beta}{4\gamma} \pi^2 + \frac{\beta^2}{2\gamma} \pi + \frac{\beta^3}{2\gamma} - \beta\lambda \right) e^{-\frac{\pi}{\beta}}. \quad (32.121)$$

После этого общее решение из второго класса принимает вид (32.89)–(32.92), куда надо подставить найденные функции Q , W и постоянную M . Это решение записано в произвольной системе координат (функции f , π) и при произвольном лоренцевом повороте (функция p_-). Соответствующая метрика на лоренцевой поверхности принимает вид (32.93).

Скалярную кривизну и псевдослед тензора кручения находим из уравнений (32.75)–(32.77):

$$R = \frac{\pi}{\gamma}, \quad T^{*\pm} = -\frac{2p^\pm}{\beta}. \quad (32.122)$$

Мы видим, что поле дилатона в квадратичной модели с точностью до множителя совпадает со скалярной кривизной.

Вид решения можно упростить, зафиксировав калибровку. Положим, например,

$$f = \tau, \quad \pi = \sigma, \quad p_- = 1.$$

Тогда общее локальное решение уравнений движения примет вид (для простоты положим также $\kappa = 1$)

$$\begin{aligned} e_0^+ &= 0, \\ e_1^+ &= e^{-Q}, \\ e_0^- &= M + W, \\ e_1^- &= -1, \\ \omega_0 &= -\frac{M + W}{\beta} + \left(\frac{\sigma^2}{4\gamma} - \lambda \right) e^{-Q}, \\ \omega_1 &= 0, \\ p_+ &= (M + W) e^Q, \end{aligned} \tag{32.123}$$

где

$$Q = \frac{\sigma}{\beta}, \quad W = \left(\frac{\beta}{4\gamma} \sigma^2 + \frac{\beta^2}{2\gamma} \sigma + \frac{\beta^3}{2\gamma} - \beta\lambda \right) e^{-\frac{\sigma}{\beta}}.$$

В этой калибровке скалярная кривизна не зависит от времени τ и пропорциональна σ :

$$R = \frac{\sigma}{\gamma}.$$

Псевдослед тензора кручения имеет следующие компоненты:

$$T^{*+} = -\frac{2}{\beta}, \quad T^{*-} = -\frac{2}{\beta} (M + W) e^Q.$$

Метрика на этой поверхности имеет вид

$$ds^2 = 2 e^{-Q} [(M + W) d\tau^2 - d\tau d\sigma]. \tag{32.124}$$

Нахождение общего локального решения уравнений движения является важным этапом исследования модели, но это далеко не все. Используя метод конформных блоков, изложенный в разделе 30, мы в состоянии описать глобальную структуру лоренцевых поверхностей двумерной гравитации. Чтобы это сделать, метрику надо переписать в конформно плоском виде, как было указано в разделе 32.6.2. Для простоты, в дальнейшем анализе мы положим $\kappa = 0$.

Переменная $\hat{q}$ связана с полем дилатона уравнением (32.114). В рассматриваемом случае его решение имеет вид

$$\hat{q} = -\beta e^{-\frac{\pi}{\beta}} \quad \Leftrightarrow \quad \pi = -\beta \ln \left(-\frac{\hat{q}}{\beta} \right). \tag{32.125}$$

Теперь можно выписать конформный множитель (32.117)

$$\Phi = \frac{\beta}{8\gamma} \hat{q} \left\{ \left[\ln^2 \left(-\frac{\hat{q}}{\beta} \right) - 2 \ln \left(-\frac{\hat{q}}{\beta} \right) + 2 - \Lambda \right] \hat{q} - \frac{4\gamma}{\beta^2} M \right\},$$

где

$$\Lambda := \frac{4\lambda\gamma}{\beta^2}.$$

Для упрощения анализа нулей и особенностей конформного множителя мы, не ограничивая общности, можем совершить линейное преобразование переменной $\hat{q}$. Положив

$$\hat{q} := -\frac{\beta q}{e}, \quad M \mapsto \frac{\beta^3 e}{4\gamma} M$$

и отбросив несущественный постоянный множитель, получим следующий конформный множитель

$$\Phi = q \left[(\ln^2 q + 1 - \Lambda)q + M \right]. \quad (32.126)$$

Глобальная структура решений уравнений движения двумерной гравитации соответствуют интервалу $q \in (0, \infty)$ и определяются нулями конформного множителя. В статье [214] классифицированы все глобальные решения квадратичной модели двумерной гравитации и, в частности, построены соответствующие диаграммы Картера–Пенроуза.

32.8 Пуассоновы сигма модели

Пуассоновы сигма модели представляют собой широкий класс интегрируемых моделей в двумерном пространстве. Они были введены в [250] под названием нелинейные калибровочные модели. Смысл этого названия будет пояснен далее в разделе 32.8.3. В дальнейшем модель была сформулирована на языке дифференциальных форм [251], и теперь известна как *пуассонова сигма модель*.

Пусть дано два многообразия $\mathbb{M}$ и $\mathbb{N}$ размерностей: $\dim \mathbb{M} = 2$ и $\dim \mathbb{N} = n \geq 2$. Двумерное многообразие $\mathbb{M}$ (поверхность) является произвольным, а многообразие $\mathbb{N}$ предполагается пуассоновым (см. раздел 19.3). Обозначим локальные координаты на $\mathbb{M}$ и $\mathbb{N}$, соответственно,

$$(x^\alpha) \in \mathbb{M}, \quad \alpha = 0, 1; \quad (X^A) \in \mathbb{N}, \quad A = 1, \dots, n.$$

Будем рассматривать многообразие $\mathbb{N}$ в качестве пространства-мишени:

$$X : \mathbb{M} \ni x \mapsto X(x) \in \mathbb{N}.$$

Локально это отображение задается n достаточно гладкими функциями $X^A(x)$.

Можно представлять себе многообразие $\mathbb{M}$ как двумерное пространство-время, хотя для формулировки пуассоновых сигма моделей задание метрики на $\mathbb{M}$ необязательно.

Пуассоновость пространства-мишени $\mathbb{N}$ означает, что на нем задана пуассонова структура

$$J = \frac{1}{2} J^{AB} \partial_A \wedge \partial_B,$$

которая была определена в разделе 19.3. Компоненты антисимметричного контравариантного тензора второго ранга $J^{AB}(X)$, которые задают пуассонову структуру в координатах, зависят от точки многообразия пространства-мишени $\mathbb{N}$ и являются сложными функциями на $\mathbb{M}$, поскольку $X = X(x)$.

Будем также предполагать, что на $\mathbb{M}$ задан набор 1-форм $A_A = dx^\alpha A_{\alpha A}$, которые по индексу A образуют ковектор относительно преобразований координат в $\mathbb{N}$. Тогда пуассоновой сигма модель определяется следующим действием

$$S := \int_{\mathbb{M}} \left(A_A \wedge dX^A + \frac{1}{2} J^{AB} A_A \wedge A_B \right). \quad (32.127)$$

Подынтегральное выражение представляет собой 2-форму и поэтому интеграл определен. Конечно, мы также предполагаем, что интеграл сходится. Подчеркнем, что действие (32.127) вообще не зависит от какой-либо метрики ни на $\mathbb{M}$, ни на $\mathbb{N}$.

Действие (32.127) можно переписать в координатах:

$$S = \int_{\mathbb{M}} dx L, \quad (32.128)$$

где лагранжиан пуассоновой сигма модели имеет вид

$$L = -\hat{\varepsilon}^{\alpha\beta} \left(A_{\alpha A} \partial_\beta X^A + \frac{1}{2} J^{AB} A_{\alpha A} A_{\beta B} \right), \quad (32.129)$$

и $\hat{\varepsilon}^{\alpha\beta}$ — полностью антисимметричная тензорная плотность второго ранга: $\hat{\varepsilon}^{01} = -\hat{\varepsilon}^{10} = -1$.

Здесь видна специфика двумерия, т.к. для построения пуассоновой сигма модели существенно используется существование полностью антисимметричного тензора второго ранга. Попытки обобщения пуассоновой сигма модели на многообразия $\mathbb{M}$ более высоких размерностей пока ни к чему не привели.

По построению, действие (32.127) инвариантно относительно общих преобразований координат на $\mathbb{M}$ и не зависит от того, задана на $\mathbb{M}$ метрика или нет. Действие (32.127) инвариантно также относительно преобразований координат в $\mathbb{N}$, т.к. при преобразовании координат $X^A \mapsto X^{A'} = X^{A'}(X)$ по правилу дифференцирования сложной функции справедливы следующие равенства

$$dX^{A'} = dx^\alpha \partial_\alpha X^A \partial_A X^{A'} = dX^A \partial_A X^{A'}.$$

Пуассонова структура J^{AB} и 1-формы A_A при этом преобразуются по тензорному закону:

$$J^{A'B'} = J^{AB} \partial_A X^{A'} \partial_B X^{B'}, \quad A_{A'} = \partial_{A'} X^A A_A,$$

что следует из их определения. Отметим, что поля X^A сами по себе не образуют вектора относительно преобразований координат в $\mathbb{N}$, т.к. являются координатными функциями.

Уравнения Эйлера–Лагранжа пуассоновой сигма модели имеют вид

$$\frac{\delta S}{\delta A_{\alpha A}} = -\hat{\varepsilon}^{\alpha\beta} (\partial_\beta X^A + J^{AB} A_{\beta B}) = 0, \quad (32.130)$$

$$\frac{\delta S}{\delta X^A} = -\hat{\varepsilon}^{\alpha\beta} \left(\partial_\alpha A_{\beta A} + \frac{1}{2} \partial_A J^{BC} A_{\alpha B} A_{\beta C} \right) = 0. \quad (32.131)$$

Используя обозначения дифференциальных форм и внешнего дифференцирования d , эти уравнения можно переписать в эквивалентном более компактном виде

$$dX^A + J^{AB} A_B = 0, \quad (32.132)$$

$$dA_A + \frac{1}{2} \partial_A J^{BC} A_B \wedge A_C = 0. \quad (32.133)$$

Внешний дифференциал от уравнения (32.132) равен нулю:

$$d(32.132) = -\frac{1}{2}(J^{AD}\partial_D J^{BC} + J^{BD}\partial_D J^{CA} + J^{CD}\partial_D J^{AB})A_B \wedge A_C = 0$$

в силу определения пуассоновой структуры. Внешний дифференциал от второго уравнения также равен нулю, $d(32.133) = 0$, т.к. любая 3-форма на двумерном многообразии равна нулю. Тем самым условия интегрируемости системы уравнений в частных производных (32.132), (32.133) выполнены.

Действие пуассоновой сигма модели (32.127) при локальных инфинитезимальных преобразованиях с малым параметром $\epsilon_A(x)$:

$$\begin{aligned}\delta X^A &= J^{AB}\epsilon_B, \\ \delta A_{\alpha A} &= -\partial_\alpha \epsilon_A - \partial_A J^{BC} A_{\alpha B} \epsilon_C,\end{aligned}\tag{32.134}$$

в линейном приближении по ϵ меняется на полную производную:

$$\delta L = -\partial_\alpha(\hat{\epsilon}^{\alpha\beta}\epsilon_A\partial_\beta X^A).$$

В этом можно убедиться прямой проверкой. Согласно второй теореме Нетер наличие локальной инвариантности приводит к зависимости уравнений движения (см. раздел 20.3)

$$\partial_\alpha \frac{\delta S}{\delta A_{\alpha A}} - \partial_B J^{CA} A_{\alpha C} \frac{\delta S}{\delta A_{\alpha B}} + J^{AB} \frac{\delta S}{\delta X^B} = 0.$$

Преобразования симметрии (32.134) содержат в себе общие преобразования координат в инфинитезимальной форме. Чтобы убедиться в этом, рассмотрим инфинитезимальные преобразования (32.134) специального вида

$$\epsilon_A := \xi^\alpha A_{\alpha A}.$$

Тогда для преобразования полей с учетом уравнений Эйлера–Лагранжа (32.130), (32.131) получим следующие равенства:

$$\delta X^A = -\xi^\alpha \partial_\alpha X^A, \quad \delta A_{\alpha A} = -\xi^\beta \partial_\beta A_{\alpha A} - \partial_\alpha \xi^\beta A_{\beta A}.$$

Это и есть общие преобразования координат на $\mathbb{M}$, записанные в инфинитезимальной форме: $x^\alpha \mapsto x^\alpha + \xi^\alpha$, где $\xi^\alpha(x)$ – компоненты векторного поля, параметризующего малые преобразования координат.

Рассмотрим коммутатор двух локальных преобразований (32.134) с параметрами ϵ_1 и ϵ_2 . Прямые вычисления приводят к равенствам:

$$[\delta_2, \delta_1]X^A = \delta_3 X^A, \tag{32.135}$$

$$[\delta_2, \delta_1]A_{\alpha A} = \delta_3 A_{\alpha A} + \partial_{AB}^2 J^{CD} \epsilon_{2C} \epsilon_{1D} (\partial_\alpha X^B + J^{BE} A_{\alpha E}), \tag{32.136}$$

где параметр результирующего преобразования (32.134) равен

$$\epsilon_{3A} := \partial_A J^{BC} \epsilon_{2B} \epsilon_{1C}.$$

Мы видим, что “алгебра” локальных преобразований (32.134) замыкается только на уравнениях движения, если пуассонова структура нелинейна по координатам.

32.8.1 Гамильтонова формулировка

Сформулируем пуассонову сигма модель на гамильтоновом языке. С этой целью проинтегрируем первое слагаемое в действии (32.128) по частям. Тогда лагранжиан пуассоновой сигма модели (32.129) примет вид

$$L = \partial_0 A_{1A} X^A - \partial_1 A_{0A} X^A + J^{AB} A_{0A} A_{1B}.$$

Теперь вычислим импульсы, сопряженные полям $A_{\alpha A}$,

$$P^{0A} := \frac{\partial L}{\partial(\partial_0 A_{0A})} = 0, \quad (32.137)$$

$$P^{1A} := \frac{\partial L}{\partial(\partial_0 A_{1A})} = X^A. \quad (32.138)$$

Отсюда следует, что поля $X^A(x)$ являются импульсами, сопряженными пространственным компонентам A_{1A} , а действие пуассоновой сигма модели после интегрирования по частям уже записано в гамильтоновом виде.

Переобозначим для краткости координаты: $x^0 = \tau$, $x^1 = \sigma$. Говорить о том, что координаты τ и σ являются временной и пространственной не имеет никакого смысла, т.к. на $\mathbb{M}$ отсутствует метрика. Мы рассматриваем координату τ просто как некоторый параметр эволюции.

Одновременные скобки Пуассона для координат и импульсов имеют канонический вид:

$$\begin{aligned} [A_{\alpha A}, P'^{\beta B}] &= \delta_\alpha^\beta \delta_A^B \delta(\sigma, \sigma'), \\ [A_{\alpha A}, A'_{\beta B}] &= 0, \quad [P^{\alpha A}, P'^{\beta B}] = 0, \end{aligned}$$

где штрих у поля означает, что оно рассматривается в точке σ' .

Нетрудно вычислить гамильтониан пуассоновой сигма модели

$$H = - \int d\sigma A_{0A} G^A, \quad (32.139)$$

где

$$G^A := \partial_1 X^A + J^{AB} A_{1B}. \quad (32.140)$$

Равенства нулю импульсов (32.137), сопряженных временным компонентам A_{0A} , являются первичными связями пуассоновой сигма модели. Условие их сохранения во времени приводит к вторичным связям:

$$\dot{P}^{0A} = [P^{0A}, H] = G^A = 0,$$

где точка обозначает дифференцирование по τ . Прямые вычисления показывают, что вторичные связи удовлетворяют замкнутой алгебре скобок Пуассона:

$$[G^A, G'^B] = G^C \partial_C J^{AB} \delta(\sigma, \sigma'). \quad (32.141)$$

Скобки Пуассона первичных связей (32.137) между собой и с вторичными связями, очевидно, равны нулю. Это значит, что пуассонова сигма модель содержит n первичных и n вторичных связей первого рода, а гамильтониан (32.139) равен линейной комбинации связей. Временные компоненты A_{0A} играют при этом роль множителей Лагранжа.

Как известно, все связи первого рода являются генераторами локальных преобразований симметрии действия (см. раздел 22.3). Преобразования симметрии для первичных связей задаются интегралом

$$T_\zeta = \int d\sigma \zeta_A P^{0A},$$

где $\zeta(x)$ – параметр преобразования. Эти преобразования действуют только на 0-компоненты полей. Явный вид преобразований задается скобками Пуассона:

$$\begin{aligned}\delta P^{0A} &:= [P^{0A}, T_\zeta] = 0, \\ \delta A_{0A} &:= [A_{0A}, T_\zeta] = \zeta_A.\end{aligned}$$

Аналогично, преобразования симметрии для вторичных связей с параметром ϵ_A задаются интегралом

$$T_\epsilon = \int d\sigma \epsilon_A G^A.$$

В явном виде имеем следующие инфинитезимальные преобразования:

$$\begin{aligned}\delta X^A &:= [X^A, T_\epsilon] = J^{AB} \epsilon_B, \\ \delta A_{1A} &:= [A_{1A}, T_\epsilon] = -\partial_1 \epsilon_A - \partial_A J^{BC} A_{1B} \epsilon_C.\end{aligned}$$

Эти преобразования симметрии совпадают с инфинитезимальными преобразованиями (32.134) для пространственных компонент полей.

Следуя общим правилам, описанным в разделе 22.3, посчитаем число физических (распространяющихся) степеней свободы. Модель содержит $4n$ полей: $A_{\alpha A}$ и $P^{\alpha A}$. Пуассонова сигма модель является калибровочной и содержит $2n$ связей первого рода. Чтобы однозначно зафиксировать решение уравнений движения необходимо наложить такое же число калибровочных условий. Поскольку $4n - 2n - 2n = 0$, то пуассонова сигма модель не описывает ни одной физической степени свободы. Такие модели часто называют *топологическими*.

32.8.2 Общее решение уравнений движения

Докажем, что пуассоновы сигма модели являются интегрируемыми. Допустим, что пуассонова структура (т.е. матрица J^{AB}) имеет постоянный ранг $2m < n$ на $\mathbb{N}$. Поскольку матрица J^{AB} антисимметрична, то ранг пуассоновой структуры должен быть четным. Согласно теореме Дарбу (см. раздел 3.4) на пуассоновом многообразии существует такая система координат, что пуассонова структура имеет канонический вид. Обозначим эти координаты через

$$\{X^A\} \mapsto \{X^M, X^I\}, \quad M = 1, \dots, 2m < n, \quad I = 2m + 1, \dots, n.$$

Тогда в координатах Дарбу пуассонова структура примет канонический вид

$$J^{AB} = \begin{pmatrix} \varpi^{-1MN} & 0 \\ 0 & 0 \end{pmatrix}, \quad (32.142)$$

где $\varpi^{-1} = -\varpi$ – матрица, обратная к канонической симплектической форме,

$$\varpi := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

Здесь $\mathbb{1}$ обозначает единичную $m \times m$ матрицу. Координатные функции X^I являются $n - 2m$ функциями Казимира пуассонова многообразия, а $2m$ -мерное подмногообразие, определяемое уравнениями $X^I = \text{const}$, – симплектическим.

Теорема 32.8.1. *Общее решение уравнений движения (32.132), (32.133) пуассоновой сигма модели в некоторой области $\mathbb{U} \subset \mathbb{M}$ имеет вид*

$$A_M = -\varpi_{MN} dX^N, \quad A_I = df_I, \quad X^I = \text{const}, \quad (32.143)$$

где $X^M(x)$ и $f_I(x)$ – произвольные функции, а X^I – произвольный набор постоянных.

Доказательство. Поскольку пуассонова сигма модель инвариантна относительно общих преобразований координат в пространстве-мишени, то перейдем к координатам Дарбу (они всегда существуют в некоторой области $\mathbb{U} \subset \mathbb{M}$). Тогда уравнения движения (32.132), (32.133) примут простой вид

$$dX^M + \varpi^{-1MN} A_N = 0, \quad (32.144)$$

$$dX^I = 0, \quad (32.145)$$

$$dA_M = 0, \quad (32.146)$$

$$dA_I = 0. \quad (32.147)$$

Уравнение (32.144) решим относительно 1-форм $A_M = -\varpi_{MN} dX^N$. Тогда уравнение (32.146) тождественно удовлетворяется. Уравнение (32.145) имеет решение $X^I = \text{const}$. Последнее уравнение (32.147) означает, что 1-формы A_I должны быть замкнуты. Тогда из леммы Пуанкаре следует, что в области $\mathbb{U}$ существуют такие функции $f_I(x)$, что $A_I = df_I$. Отсюда вытекает, что общее решение уравнений Эйлера–Лагранжа пуассоновой сигма модели имеет вид (32.143). $\square$

Найденное решение было получено без фиксирования калибровки (системы координат) и параметризуется n произвольными функциями $\{X^M, f_I\}$. Единственное условие на произвольные функции состоит в том, что они должны определять переход к координатам Дарбу, т.е. якобиан преобразования координат должен быть отличен от нуля:

$$\frac{\partial(X^M, f_I)}{\partial(X^A)} \neq 0.$$

Мы видим, что пуассонова сигма модель является интегрируемой, поскольку координаты Дарбу локально всегда существуют. Для того, чтобы выписать общее решение в явном виде при заданной пуассоновой структуре $J^{AB}(X)$, необходимо найти функции перехода к координатам Дарбу. Эта задача сложна, однако сам факт их существования очень важен. Фактически, нахождение координат Дарбу эквивалентно решению исходной системы уравнений (32.132), (32.133).

Сделаем важное замечание. Общее решение (32.143) было получено в предположении, что ранг Пуассоновой структуры постоянен на всем $\mathbb{N}$. Однако в пространстве-мишени могут быть подмногообразия, на которых ранг пуассоновой структуры меньше $2m$. Уравнения движения на этих подмногообразиях (32.132), (32.133) необходимо рассматривать, как особые случаи. Например, двумерную гравитацию с кручением можно рассматривать, как пуассонову сигма модель с $n = 3$ (см. далее раздел 32.8.4). Ранг соответствующей пуассоновой структуры равен 2 всюду на пространстве-мишени за исключением отдельных точек, где он обращается в нуль. Этим точкам соответствует важный класс вырожденных решений, описывающий поверхности нулевого кручения и постоянной кривизны.

32.8.3 Двумерные поля Янга–Миллса

Пусть пространством-мишенью является полупростая группа Ли $\mathbb{N} = \mathbb{G}$ размерности n . Базис $\{L_A\}$ соответствующей алгебры Ли $\mathfrak{g}$ удовлетворяет коммутационным соотношениям

$$[L_A, L_B] = f_{AB}^{C} L_C, \quad (32.148)$$

где f_{AB}^{C} – структурные константы. Форма Киллинга–Картана

$$\eta_{AB} := -f_{AC}^{D} f_{BD}^{C}$$

для полупростой алгебры Ли $\mathfrak{g}$ невырождена и задает двусторонне инвариантную метрику на групповом многообразии $\mathbb{G}$.

Базис алгебры $\mathfrak{g}^*$, дуальной к $\mathfrak{g}$, задается коммутационными соотношениями

$$[L^A, L^B] = f^{AB}_{C} L^C,$$

где подъем и опускание индексов производится с помощью формы Киллинга–Картана. Изоморфизм алгебр $\mathfrak{g}$ и $\mathfrak{g}^*$ естественным образом устанавливается формой Киллинга–Картана: $L^A := \eta^{AB} L_B$.

Если обозначить координаты группового многообразия $\mathbb{G}$ через X^A , то линейный по координатам антисимметричный тензор второго ранга

$$J^{AB} := f^{AB}_{C} X^C \quad (32.149)$$

задает пуассонову структуру на групповом многообразии. При этом тождество Якоби для пуассоновой структуры является следствием тождеств Якоби для структурных констант.

Рассмотрим действие Янга–Миллса для калибровочной алгебры Ли $\mathfrak{g}^*$, дуальной к $\mathfrak{g}$, в двумерном пространстве-времени Минковского

$$S_{\text{YM}} = \int_{\mathbb{M}} dx L_{\text{YM}} = -\frac{1}{4} \int_{\mathbb{M}} dx F_{\alpha\beta A} F^{\alpha\beta A}, \quad (32.150)$$

где

$$F_{\alpha\beta A} := \partial_\alpha A_{\beta A} - \partial_\beta A_{\alpha A} + A_{\alpha B} A_{\beta C} f^{BC}_{A}$$

– компоненты тензора кривизны для локальной формы связности с компонентами $A_{\alpha A}$. Подъем и опускание греческих индексов производится с помощью метрики Минковского $\eta_{\alpha\beta} := \text{diag}(+-)$.

Запишем действие Янга–Миллса в гамильтоновой форме. Импульсы, сопряженные к полям $A_{\alpha A}$, равны

$$P^{0A} := \frac{\partial L_{\text{YM}}}{\partial \dot{A}_{0A}} = 0, \quad (32.151)$$

$$P^{1A} := \frac{\partial L_{\text{YM}}}{\partial \dot{A}_{1A}} = -F^{01A}, \quad (32.152)$$

где точка обозначает дифференцирование по $x^0 := \tau$. Нетрудно вычислить гамильтониан модели

$$\mathcal{H}_{\text{YM}} = \int d\sigma H_{\text{YM}} = \int d\sigma \left(\frac{1}{2} P^{1A} P^1_{A} - A_{0A} G^A \right),$$

где

$$G^A = \partial_1 P^{1A} + f^{AB}_{C} A_{1B} P^{1C}. \quad (32.153)$$

Равенства нулю импульсов (32.151), соответствующих нулевым компонентам полей, представляют собой первичные связи. Их сохранение во времени приводит к вторичным связям:

$$\dot{P}^{0A} = [P^{0A}, H_{\text{YM}}] = G^A = 0.$$

Одновременные скобки Пуассона вторичных связей между собой образуют замкнутую алгебру:

$$[G^A, G'^B] = f^{AB}{}_C G^C \delta(\sigma, \sigma').$$

Таким образом, теория Янга–Миллса содержит $2n$ связей первого рода.

Запишем действие Янга–Миллса в канонически сопряженных переменных $A_{\alpha A}$ и $P^{\alpha A}$. Простые вычисления с интегрированием по частям показывают, что действие равно сумме двух слагаемых:

$$S_{\text{YM}} = \int dx (P^{\alpha A} \dot{A}_{\alpha A} - H_{\text{YM}}) = S_{\text{YM}}^{\text{PSM}} + S_{\text{YM}}^{\text{P}^2},$$

где

$$S_{\text{YM}}^{\text{PSM}} := \int dx (A_{0A} \partial_1 P^{1A} - A_{1A} \partial_0 P^{1A} + A_{0A} A_{1B} f^{AB}{}_C P^{1C}), \quad (32.154)$$

$$S_{\text{YM}}^{\text{P}^2} := -\frac{1}{2} \int dx P^{1A} P^1{}_A. \quad (32.155)$$

Первое слагаемое представляет собой пуассонову сигма модель (32.128), если положить $X^A := P^{1A}$. Второе слагаемое представляет собой интеграл от квадратичной формы импульсов

$$C = \frac{1}{2} X^A X_A, \quad (32.156)$$

которая является функцией Казимира для пуассоновой структуры (32.149). В этом легко убедиться, вычислив скобку Пуассона с произвольной гладкой функцией $f \in C^\infty(\mathbb{G})$:

$$[C, f] = f^{ABC} X_A X_C \partial_B f = 0,$$

ввиду полной антисимметрии структурных констант.

Важным фактом является то, что не только исходное действие (32.150) калибровочно инвариантно, но и оба слагаемых (32.154) и (32.155) в отдельности (с точностью до дивергенции). Более того, слагаемое $S_{\text{YM}}^{\text{PSM}}$ инвариантно также относительно общих преобразований координат, а эта инвариантность отсутствует в исходном действии.

Пуассонова сигма модель для полей Янга–Миллса (32.154) соответствует линейной по координатам пуассоновой структуре (32.149). Поэтому ее обобщение на произвольную пуассонову структуру [250] сначала было названо нелинейной калибровочной моделью.

32.8.4 Двумерная гравитация как пуассонова сигма модель

Двумерная гравитация с кручением общего вида, рассмотренная в разделе 32.2, представляет собой нетривиальную пуассонову сигма модель. В настоящем разделе эта модель будет описана с этой точки зрения.

В моделях двумерной гравитации предполагается, что на базовом двумерном пространстве-времени $\mathbb{M}$ задана метрика лоренцевой сигнатуры. Независимыми переменными в двумерной гравитации с кручением без полей материи являются репер

$e_{\alpha a}$, $a = 0, 1$, и лоренцева связность, которая в двумерном случае параметризуется псевдоковекторным полем ω_α . Как было показано в разделе 32.8.1, пуассонова сигма модель представляет собой формулировку теории в канонической форме. Поэтому сформулируем модель сразу на этом языке, что было сделано в разделе 32.3. Напомним, что импульсы, сопряженные пространственным компонентам репера e_{1a} и лоренцевой связности ω_α , были обозначены через p^a и π .

Действие двумерной гравитации с кручением имеет вид

$$S = \int_{\mathbb{M}} dx L_G,$$

где

$$L_G = -\frac{1}{2}(\pi \hat{R} + p^a \hat{T}_a^*) - \sqrt{|g|} \left(\frac{1}{2} p_a p^a U + V \right). \quad (32.157)$$

Лагранжиан (32.157) зависит от двух произвольных функций $U(\pi)$ и $V(\pi)$. Напомним, что плотности скалярной кривизны и псевдоследа тензора кручения выражаются через репер и лоренцеву связность с помощью следующих соотношений:

$$\hat{R} := \sqrt{|g|} R = 2\hat{\varepsilon}^{\alpha\beta} \partial_\alpha \omega_\beta, \quad (32.158)$$

$$\hat{T}_a^* := \sqrt{|g|} T_a^* = 2\hat{\varepsilon}^{\alpha\beta} (\partial_\alpha e_{\beta a} - \omega_\alpha \varepsilon^{ab} e_\beta{}^b). \quad (32.159)$$

В разделе 32.4 было показано, что модель двумерной гравитации (32.157) эквивалентна двумерной дилатонной гравитации. При этом импульс π отождествляется с полем дилатона. Однако только в виде (32.157) двумерная гравитация имеет вид пуассоновой сигма модели.

Пространство-мишень пуассоновой сигма модели $\mathbb{N}$, соответствующей двумерной гравитации, трехмерно: $\{X^A\} = \{p^a, \pi\} \in \mathbb{R}^3$. Пуассонова структура на $\mathbb{R}^3$ задается антисимметричной 3×3 матрицей

$$J^{AB} := \begin{pmatrix} -\varepsilon^{ab} \left(\frac{1}{2} p_c p^c U + V \right) & p^c \varepsilon_c{}^a & \\ & 0 & \frac{1}{2} p_c p^c U + V - p^0 \\ -p^c \varepsilon_c{}^b & 0 & -p^1 \\ & p^0 & p^1 & 0 \end{pmatrix} \quad (32.160)$$

Эта матрица, как легко проверить, удовлетворяет тождеству Якоби (19.24).

Набор 1-форм пуассоновой сигма модели задается репером и лоренцевой связностью: $\{A_{\alpha a}\} = \{e_{\alpha a}, \omega_\alpha\}$. Теперь нетрудно убедиться, что с точностью до дивергенции лагранжиан (32.157) совпадает с лагранжианом пуассоновой сигма модели

$$L_G \stackrel{\text{div}}{=} -\hat{\varepsilon}^{\alpha\beta} \left(A_{\alpha A} \partial_\beta X^A + \frac{1}{2} J^{AB} A_{\alpha A} A_{\beta B} \right).$$

Если функцию $\frac{1}{2} p_c p^c U + V$ в пуассоновой структуре заменить на более общую функцию $W(p_c p^c, \pi)$, зависящую от двух аргументов: инвариантной суммы $p_c p^c$ и π , то можно проверить, что соответствующая матрица J^{AB} также будет задавать пуассонову структуру. Согласно общей схеме раздела 32.8.2 такая модель является интегрируемой. Однако найти координаты Дарбу и, следовательно, выписать общее решение не удастся. Для двумерной гравитации с кручением вида (32.157) общее решение уравнений Эйлера–Лагранжа можно выписать в явном виде, что и является причиной конкретного выбора пуассоновой структуры.

Пуассонова структура (32.160) является вырожденной: $\det J^{AB} = 0$. Почти всюду на пространстве-мишени $N \approx \mathbb{R}^3$ ее ранг равен двум $\text{rank } J^{AB} = 2$. Исключение составляют лишь точки $\pi = \text{const}$, $p^a = 0$, где значения π являются корнями уравнения $V(\pi) = 0$. В этих точках пуассонова структура обращается в нуль. Общее решение уравнений движения двумерной гравитации без полей материи было приведено в разделе 32.6. В вырожденных точках модель описывает поверхности постоянной кривизны и нулевого кручения. Остальным точкам пространства-мишени соответствуют поверхности непостоянной кривизны и ненулевого кручения.

Глава 33

Сплетенные решения в общей теории относительности

В настоящей главе построены глобальные решения вакуумных уравнений общей теории относительности (23.5) с космологической постоянной в предположении, что четырехмерное пространство-время является *сплетенным произведением* двух поверхностей. При этом не делается никаких предположений о симметрии решений. Как следствие уравнений движения по крайней мере одна из двух поверхностей должна быть поверхностью постоянной кривизны. Отсюда вытекает, что метрика имеет по крайней мере три вектора Киллинга. Другими словами, свойства симметрии решений при таком подходе являются следствием самих уравнений движения. Построенные решения включают, в частности, сферически симметричные решения, которые соответствуют произведению некоторой лоренцевой поверхности на сферу. Многие глобальные решения имеют интересную физическую интерпретацию. В частности, построены решения, описывающие кротовые норы, доменные стенки сингулярностей кривизны, космические струны, космические струны, окруженные доменными стенками, решения с замкнутыми времениподобными кривыми и др. [237]

33.1 Сплетенное произведение

Определение. Пусть задано два многообразия $\mathbb{M}_1$ и $\mathbb{M}_2$ с метриками g и h , соответственно. Касательное пространство в каждой точке топологического произведения $(x_1, x_2) \in \mathbb{M}_1 \times \mathbb{M}_2$ разлагается в прямую сумму:

$$\mathbb{T}_{(x_1, x_2)}(\mathbb{M}_1 \times \mathbb{M}_2) = \mathbb{T}_{x_1}(\mathbb{M}_1) \oplus \mathbb{T}_{x_2}(\mathbb{M}_2).$$

Сплетенным произведением (warped product) двух многообразий называется их топологическое произведение $\mathbb{M}_1 \times \mathbb{M}_2$ с метрикой $\hat{g}$, которая определена следующим соотношением

$$\hat{g}(X, Y) := k(x_2)g(X_1, Y_1) + m(x_1)h(X_2, Y_2), \quad (33.1)$$

где $k(x_2) \in \mathcal{C}^{n_2}(\mathbb{M}_2)$ и $m(x_1) \in \mathcal{C}^{n_1}(\mathbb{M}_1)$ – достаточно гладкие отличные от нуля функции на многообразиях $\mathbb{M}_2$ и $\mathbb{M}_1$, и

$$\begin{aligned} \mathbb{T}(\mathbb{M}_1 \times \mathbb{M}_2) \ni \quad X &= X_1 \oplus X_2 && \in \mathbb{T}(\mathbb{M}_1) \oplus \mathbb{T}(\mathbb{M}_2), \\ \mathbb{T}(\mathbb{M}_1 \times \mathbb{M}_2) \ni \quad Y &= Y_1 \oplus Y_2 && \in \mathbb{T}(\mathbb{M}_1) \oplus \mathbb{T}(\mathbb{M}_2), \end{aligned}$$

– разложение векторных полей X, Y , касательных к $\mathbb{M}_1 \times \mathbb{M}_2$, в прямую сумму. $\square$

Предположим, что четырехмерное пространство-время является сплетенным произведением двух поверхностей: $\mathbb{M} = \mathbb{U} \times \mathbb{V}$, где $\mathbb{U}$ – поверхность с лоренцевой метрикой g и $\mathbb{V}$ – поверхность с римановой метрикой h . Обозначим локальные координаты на $\mathbb{U}$ и $\mathbb{V}$, соответственно, через x^α ($\alpha, \beta, \dots = 0, 1$) и y^μ ($\mu, \nu, \dots = 2, 3$). Тогда топологическому произведению $\mathbb{U} \times \mathbb{V}$ соответствуют координаты $\{x^i\} := \{x^\alpha, y^\mu\}$ ($i, j, \dots = 0, 1, 2, 3$). В этой системе координат четырехмерная метрика имеет блочно-диагональный вид:

$$\hat{g}_{ij} = \begin{pmatrix} k(y)g_{\alpha\beta}(x) & 0 \\ 0 & m(x)h_{\mu\nu}(y) \end{pmatrix}, \quad (33.2)$$

где $k(y)$ и $m(x)$ – достаточно гладкие отличные от нуля функции на $\mathbb{V}$ и $\mathbb{U}$, соответственно.

В настоящей главе шляпка над символом означает, что соответствующий геометрический объект относится ко всему четырехмерному пространству-времени $\mathbb{M}$, а символы без шляпки относятся к двумерным поверхностям $\mathbb{U}$ или $\mathbb{V}$. Соответственно $g_{\alpha\beta}$ и $h_{\mu\nu}$ являются метриками на $\mathbb{U}$ и $\mathbb{V}$. Греческие буквы из начала ($\alpha, \beta, \dots$) и середины ($\mu, \nu, \dots$) алфавита всегда относятся к координатам на первой и второй поверхностям, соответственно.

В физике функции $k(y)$ и $m(x)$ часто называют (*дилатонными*) полями на поверхностях $\mathbb{V}$ и $\mathbb{U}$.

Для определенности будем считать поверхность $\mathbb{U}$ псевдоримановым многообразием с метрикой лоренцевой сигнатуры, а поверхность $\mathbb{V}$ – римановым многообразием с положительно определенной метрикой. Тогда с точностью до перестановки первых двух координат сигнатура метрики на $\mathbb{M}$ будет либо $(+ - - -)$, либо $(- + + +)$ в зависимости от знака m . Эти метрики связаны между собой инверсией $\hat{g}_{ij} \mapsto -\hat{g}_{ij}$, относительно которой уравнения Эйнштейна при отсутствии полей материи и космологической постоянной инвариантны. Предположим также, что обе поверхности являются ориентируемыми.

Отметим, что относительно вида метрики (33.2) не делается никаких дополнительных предположений, связанных с симметрией. Однако в дальнейшем мы увидим, что уравнения Эйнштейна и требование полноты многообразий приводят к тому, что по крайней мере одна из поверхностей $\mathbb{U}$ или $\mathbb{V}$ должна быть поверхностью постоянной кривизны. То есть любое максимально продолженное решение уравнений Эйнштейна вида (33.2) допускает по крайней мере три вектора Киллинга. Следовательно, в рассматриваемом случае симметрия решений является следствием уравнений движения. В частном случае будут получены сферически симметричные решения, когда поверхность $\mathbb{V}$ является сферой S^2 .

Все решения, рассмотренные в настоящей главе относятся к решениям типа D по классификации Петрова [122]. Физическая интерпретация решений опирается на глобальную структуру пространства-времени.

33.2 Двумерная редукция

Вид метрики (33.2) позволяет решить явно четырехмерные вакуумные уравнения Эйнштейна с космологической постоянной Λ

$$\hat{R}_{ij} = \Lambda \hat{g}_{ij}, \quad (33.3)$$

и построить глобальные (максимально продолженные) решения.

Мы увидим, что уравнения Эйнштейна существенно ограничивают дилатонные поля: по крайней мере одно дилатонное поле должно быть постоянно. Поэтому все решения делятся на три основных класса: оба дилатонных поля постоянны (случай А), только $k = \text{const}$ (случай В) или $m = \text{const}$ (случай С). В первом случае решение уравнений Эйнштейна представляет собой топологическое произведение двух поверхностей постоянной кривизны. В случае В риманова поверхность $\mathbb{V}$ должна быть поверхностью постоянной кривизны. Сюда входят сферически симметричные решения, а также другие решения, когда поверхность $\mathbb{V}$ представляет собой евклидову плоскость или плоскость Лобачевского (двуполостный гиперboloид). Последние решения соответствуют кротовым норам. В случае С поверхность $\mathbb{U}$ должна быть поверхностью постоянной кривизны. Эти решения описывают космические струны и доменные стенки сингулярности кривизны.

Приступим к решению уравнений Эйнштейна (33.3). Метрика, обратная к (33.2), имеет вид

$$\hat{g}^{ij} = \begin{pmatrix} \frac{1}{k} g^{\alpha\beta} & 0 \\ 0 & \frac{1}{m} h^{\mu\nu} \end{pmatrix}, \quad (33.4)$$

где $g^{\alpha\beta}$ и $h^{\mu\nu}$ матрицы, обратные, соответственно, к $g_{\alpha\beta}$ и $h_{\mu\nu}$. Символы Кристоффеля (6.24) равны

$$\begin{aligned} \hat{\Gamma}_{\alpha\beta}^{\gamma} &= \Gamma_{\alpha\beta}^{\gamma}, \\ \hat{\Gamma}_{\alpha\beta}^{\mu} &= -\frac{1}{2} g_{\alpha\beta} \frac{h^{\mu\nu} \partial_{\nu} k}{m}, \\ \hat{\Gamma}_{\alpha\mu}^{\gamma} &= \hat{\Gamma}_{\mu\alpha}^{\gamma} = \frac{1}{2} \delta_{\alpha}^{\gamma} \frac{\partial_{\mu} k}{k}, \\ \hat{\Gamma}_{\alpha\mu}^{\nu} &= \hat{\Gamma}_{\mu\alpha}^{\nu} = \frac{1}{2} \delta_{\mu}^{\nu} \frac{\partial_{\alpha} m}{m}, \\ \hat{\Gamma}_{\mu\nu}^{\alpha} &= -\frac{1}{2} h_{\mu\nu} \frac{g^{\alpha\beta} \partial_{\beta} m}{k}, \\ \hat{\Gamma}_{\mu\nu}^{\rho} &= \Gamma_{\mu\nu}^{\rho}. \end{aligned} \quad (33.5)$$

Вычислим компоненты полного тензора кривизны для сплетенного произведения двух многообразий. Это немного больше, чем необходимо для решения уравнений Эйнштейна, т.к. последние содержат только тензор Риччи и скалярную кривизну, однако приведенные ниже формулы необходимы для анализа получаемых решений.

Прямые вычисления приводят к следующим выражениям для линейно независимых компонент тензора кривизны со всеми опущенными индексами:

$$\begin{aligned} \hat{R}_{\alpha\beta\gamma\delta} &= k R_{\alpha\beta\gamma\delta} + \frac{1}{4} (g_{\alpha\gamma} g_{\beta\delta} - g_{\beta\gamma} g_{\alpha\delta}) \frac{(\partial k)^2}{m}, \\ \hat{R}_{\alpha\beta\gamma\mu} &= \frac{1}{4} (g_{\beta\gamma} \partial_{\alpha} m - g_{\alpha\gamma} \partial_{\beta} m) \frac{\partial_{\mu} k}{m}, \\ \hat{R}_{\alpha\beta\mu\nu} &= 0, \\ \hat{R}_{\alpha\mu\beta\nu} &= \frac{1}{2} h_{\mu\nu} \nabla_{\alpha} \nabla_{\beta} m - \frac{1}{4} h_{\mu\nu} \frac{\partial_{\alpha} m \partial_{\beta} m}{m} + \frac{1}{2} g_{\alpha\beta} \nabla_{\mu} \nabla_{\nu} k - \frac{1}{4} g_{\alpha\beta} \frac{\partial_{\mu} k \partial_{\nu} k}{k}, \\ \hat{R}_{\alpha\mu\nu\rho} &= \frac{1}{4} (h_{\mu\nu} \partial_{\rho} k - h_{\mu\rho} \partial_{\nu} k) \frac{\partial_{\alpha} m}{k}, \\ \hat{R}_{\mu\nu\rho\sigma} &= m R_{\mu\nu\rho\sigma} + \frac{1}{4} (h_{\mu\rho} h_{\nu\sigma} - h_{\nu\rho} h_{\mu\sigma}) \frac{(\partial m)^2}{k}, \end{aligned} \quad (33.6)$$

где ∇_α и ∇_μ – ковариантные производные с символами Кристоффеля $\Gamma_{\alpha\beta}^\gamma$ и $\Gamma_{\mu\nu}^\rho$, соответственно.

Число линейно независимых компонент тензора кривизны с одним контравариантным индексом больше. Они имеют вид

$$\begin{aligned}
\widehat{R}_{\alpha\beta\gamma}{}^\delta &= R_{\alpha\beta\gamma}{}^\delta + \frac{1}{4}(g_{\alpha\gamma}\delta_\beta^\delta - g_{\beta\gamma}\delta_\alpha^\delta)\frac{(\partial k)^2}{km}, \\
\widehat{R}_{\alpha\beta\mu}{}^\delta &= \frac{1}{4}(\delta_\alpha^\delta\partial_\beta m - \delta_\beta^\delta\partial_\alpha m)\frac{\partial_\mu k}{km}, \\
\widehat{R}_{\mu\alpha\beta}{}^\delta &= \frac{1}{4}(g_{\alpha\beta}\partial^\delta m - \delta_\alpha^\delta\partial_\beta m)\frac{\partial_\mu k}{km}, \\
\widehat{R}_{\alpha\mu\nu}{}^\delta &= -\frac{1}{2}h_{\mu\nu}\frac{\nabla_\alpha\nabla^\delta m}{k} + \frac{1}{4}h_{\mu\nu}\frac{\partial_\alpha m\partial^\delta m}{km} - \frac{1}{2}\delta_\alpha^\delta\frac{\nabla_\mu\nabla_\nu k}{k} + \frac{1}{4}\delta_\alpha^\delta\frac{\partial_\mu k\partial_\nu k}{k^2}, \\
\widehat{R}_{\mu\nu\alpha}{}^\delta &= 0, \\
\widehat{R}_{\mu\nu\rho}{}^\delta &= \frac{1}{4}(h_{\nu\rho}\partial_\mu k - h_{\mu\rho}\partial_\nu k)\frac{\partial^\delta m}{k^2}, \\
\widehat{R}_{\alpha\beta\gamma}{}^\sigma &= \frac{1}{4}(g_{\beta\gamma}\partial_\alpha m - g_{\alpha\gamma}\partial_\beta m)\frac{\partial^\sigma k}{m^2}, \\
\widehat{R}_{\alpha\beta\mu}{}^\sigma &= 0, \\
\widehat{R}_{\alpha\mu\beta}{}^\sigma &= \frac{1}{2}\delta_\mu^\sigma\frac{\nabla_\alpha\nabla_\beta m}{m} - \frac{1}{4}\delta_\mu^\sigma\frac{\partial_\alpha m\partial_\beta m}{m^2} + \frac{1}{2}g_{\alpha\beta}\frac{\nabla_\mu\nabla^\sigma k}{m} - \frac{1}{4}g_{\alpha\beta}\frac{\partial_\mu k\partial^\sigma k}{km}, \\
\widehat{R}_{\alpha\mu\nu}{}^\sigma &= \frac{1}{4}(g_{\mu\nu}\partial^\sigma k - \delta_\mu^\sigma\partial_\nu k)\frac{\partial_\alpha m}{km}, \\
\widehat{R}_{\mu\nu\alpha}{}^\sigma &= \frac{1}{4}(\delta_\mu^\sigma\partial_\nu k - \delta_\nu^\sigma\partial_\mu k)\frac{\partial_\alpha m}{km}, \\
\widehat{R}_{\mu\nu\rho}{}^\sigma &= R_{\mu\nu\rho}{}^\sigma + \frac{1}{4}(h_{\mu\rho}\delta_\nu^\sigma - h_{\nu\rho}\delta_\mu^\sigma)\frac{(\partial m)^2}{km},
\end{aligned} \tag{33.7}$$

где введены обозначения: $\partial^\delta m := g^{\delta\gamma}\partial_\gamma m$ и $\partial^\sigma k := h^{\sigma\rho}\partial_\rho k$.

Полученные выше формулы для компонент тензора кривизны не зависят от размерностей многообразий $\mathbb{M}_1$ и $\mathbb{M}_2$. Положим $n_1 := \dim \mathbb{M}_1$ и $n_2 := \dim \mathbb{M}_2$. Тогда для компонент тензора Риччи получим следующие выражения:

$$\begin{aligned}
\widehat{R}_{\alpha\beta} &= R_{\alpha\beta} + \frac{1}{2}g_{\alpha\beta}\frac{\nabla^2 k}{m} + \frac{1}{4}(n_1 - 2)g_{\alpha\beta}\frac{(\partial k)^2}{km} + \frac{1}{2}n_2\frac{\nabla_\alpha\nabla_\beta m}{m} - \frac{1}{4}n_2\frac{\partial_\alpha m\partial_\beta m}{m^2}, \\
\widehat{R}_{\alpha\mu} &= \widehat{R}_{\mu\alpha} = -\frac{1}{4}(n_1 + n_2 - 2)\frac{\partial_\alpha m\partial_\mu k}{km}, \\
\widehat{R}_{\mu\nu} &= R_{\mu\nu} + \frac{1}{2}h_{\mu\nu}\frac{\nabla^2 m}{k} + \frac{1}{4}(n_2 - 2)h_{\mu\nu}\frac{(\partial m)^2}{km} + \frac{1}{2}n_1\frac{\nabla_\mu\nabla_\nu k}{k} - \frac{1}{4}n_1\frac{\partial_\mu k\partial_\nu k}{k^2},
\end{aligned} \tag{33.8}$$

где $R_{\alpha\beta}$ и $R_{\mu\nu}$ – тензоры Риччи многообразий $\mathbb{M}_1$ и $\mathbb{M}_2$, построенные, соответственно, по метрикам $g_{\alpha\beta}$ и $h_{\mu\nu}$.

Наконец, получаем выражение для скалярной кривизны сплетенного произведения двух многообразий

$$\widehat{R} = \frac{1}{k}R^g + n_2\frac{\nabla^2 m}{km} + \frac{1}{4}n_2(n_2 - 3)\frac{(\partial m)^2}{km^2} + \frac{1}{m}R^h + n_1\frac{\nabla^2 k}{km} + \frac{1}{4}n_1(n_1 - 3)\frac{(\partial k)^2}{k^2m}, \tag{33.9}$$

где R^g и R^h – скалярные кривизны многообразий $\mathbb{M}_1$ и $\mathbb{M}_2$.

Теперь зафиксируем размерности. В рассматриваемом случае $n_1 = n_2 = 2$. Тогда компоненты тензора Риччи (6.89) принимают вид

$$\begin{aligned}\widehat{R}_{\alpha\beta} &= R_{\alpha\beta} + \frac{\nabla_\alpha \nabla_\beta m}{m} - \frac{\nabla_\alpha m \nabla_\beta m}{2m^2} + \frac{g_{\alpha\beta} \nabla^2 k}{2m} \\ \widehat{R}_{\alpha\mu} &= \widehat{R}_{\mu\alpha} = -\frac{\nabla_\alpha m \nabla_\mu k}{2mk} \\ \widehat{R}_{\mu\nu} &= R_{\mu\nu} + \frac{\nabla_\mu \nabla_\nu k}{k} - \frac{\nabla_\mu k \nabla_\nu k}{2k^2} + \frac{h_{\mu\nu} \nabla^2 m}{2k},\end{aligned}\quad (33.10)$$

где, для краткости, введены обозначения

$$\nabla^2 m := g^{\alpha\beta} \nabla_\alpha \nabla_\beta m, \quad \nabla^2 k := h^{\mu\nu} \nabla_\mu \nabla_\nu k. \quad (33.11)$$

Здесь и далее в этой главе символ ∇ обозначает ковариантную производную с соответствующими символами Кристоффеля. Четырехмерная скалярная кривизна равна

$$\widehat{R} = \frac{1}{k} R^g + 2 \frac{\nabla^2 m}{km} - \frac{(\nabla m)^2}{2km^2} + \frac{1}{m} R^h + 2 \frac{\nabla^2 k}{km} - \frac{(\nabla k)^2}{2k^2 m}, \quad (33.12)$$

где введены обозначения

$$(\nabla m)^2 := g^{\alpha\beta} \partial_\alpha m \partial_\beta m, \quad (\nabla k)^2 := h^{\mu\nu} \partial_\mu k \partial_\nu k. \quad (33.13)$$

Скалярные кривизны поверхностей $\mathbb{U}$ и $\mathbb{V}$ обозначены через R^g и R^h , соответственно.

Таким образом, уравнений Эйнштейна (33.3) для метрики (33.2) принимают вид

$$R_{\alpha\beta} + \frac{\nabla_\alpha \nabla_\beta m}{m} - \frac{\nabla_\alpha m \nabla_\beta m}{2m^2} + \frac{1}{2} g_{\alpha\beta} \left(\frac{\nabla^2 k}{m} - 2k\Lambda \right) = 0, \quad (33.14)$$

$$R_{\mu\nu} + \frac{\nabla_\mu \nabla_\nu k}{k} - \frac{\nabla_\mu k \nabla_\nu k}{2k^2} + \frac{1}{2} h_{\mu\nu} \left(\frac{\nabla^2 m}{k} - 2m\Lambda \right) = 0, \quad (33.15)$$

$$\frac{\nabla_\alpha m \nabla_\mu k}{mk} = 0. \quad (33.16)$$

Перепишем уравнения (33.14) и (33.15) в более удобном виде, выделив из них след, который определяет скалярные кривизны поверхностей:

$$R^g + \frac{\nabla^2 m}{m} - \frac{(\nabla m)^2}{2m^2} + \frac{\nabla^2 k}{m} - 2k\Lambda = 0, \quad (33.17)$$

$$R^h + \frac{\nabla^2 k}{k} - \frac{(\nabla k)^2}{2k^2} + \frac{\nabla^2 m}{k} - 2m\Lambda = 0. \quad (33.18)$$

Бесследовые части уравнений (33.14) и (33.15), умноженные на m и k , принимают простой вид

$$\nabla_\alpha \nabla_\beta m - \frac{\nabla_\alpha m \nabla_\beta m}{2m} - \frac{1}{2} g_{\alpha\beta} \left[\nabla^2 m - \frac{(\nabla m)^2}{2m} \right] = 0, \quad (33.19)$$

$$\nabla_\mu \nabla_\nu k - \frac{\nabla_\mu k \nabla_\nu k}{2k} - \frac{1}{2} h_{\mu\nu} \left[\nabla^2 k - \frac{(\nabla k)^2}{2k} \right] = 0. \quad (33.20)$$

Они не содержат слагаемые с кривизной вовсе, потому что в двух измерениях тензор Риччи полностью определяется скалярной кривизной (6.89) и не имеет бесследовой части.

Отметим, что наличие сингулярности у двумерной скалярной кривизны на поверхности означает ее наличие в полном тензоре кривизны в соответствии с формулой (33.12).

Таким образом, четырехмерные уравнения Эйнштейна (33.3) для метрики вида (33.2) эквивалентны системе уравнений (33.16)–(33.20). Уравнения (33.19) и (33.20) содержат функции, зависящие только от координат x и y , соответственно. В то же время координаты различных поверхностей в уравнениях (33.16), (33.17) и (33.18) перемешаны.

Уравнение (33.16) накладывает жесткие ограничения. Как следствие имеем, что либо поле дилатона k , либо поле дилатона m , либо k и m одновременно должны быть постоянны. Соответственно, возможны три случая:

$$\begin{aligned} \text{A :} & \quad k = \text{const} \neq 0, & m = \text{const} \neq 0, \\ \text{B :} & \quad k = \text{const} \neq 0, & \nabla_\alpha m \neq 0, \\ \text{C :} & \quad \nabla_\mu k \neq 0, & m = \text{const} \neq 0. \end{aligned} \quad (33.21)$$

Рассмотрим эти случаи подробно.

33.3 Произведение поверхностей постоянной кривизны

Наиболее симметричные решения вакуумных уравнений Эйнштейна (33.3) в виде топологического произведения двух поверхностей постоянной кривизны возникают, когда оба дилатонных поля k и m постоянны (случай А в (33.21)). Если k и m постоянны, то уравнения (33.16), (33.19), и (33.20) удовлетворяются, и скалярные кривизны обеих поверхностей $\mathbb{U}$ и $\mathbb{V}$ должны быть постоянны как следствие уравнений (33.17), (33.18), которые принимают следующий вид

$$R^g = 2k\Lambda, \quad R^h = 2m\Lambda. \quad (33.22)$$

Если $\Lambda = 0$, то обе поверхности $\mathbb{U}$ и $\mathbb{V}$ имеют нулевую кривизну, и все пространство-время $\mathbb{M}$ представляет пространство Минковского или его фактор пространство (цилиндр или тор) по группе преобразований, действующей свободно и собственноразрывно, с метрикой Лоренца

$$\hat{g}_{ij} = \text{diag}(+ - - -) \quad \text{или} \quad \hat{g}_{ij} = \text{diag}(- + + +). \quad (33.23)$$

При ненулевой космологической постоянной $\Lambda \neq 0$ обе поверхности $\mathbb{U}$ и $\mathbb{V}$ имеют постоянную ненулевую кривизну. Поверхности постоянной кривизны были рассмотрены в главе 29, и мы приведем лишь выражения для соответствующих метрик. Если $\mathbb{U}$ является полной псевдоримановой поверхностью ненулевой кривизны $R^g = -2K = \text{const}$, то она представляет собой однополостный гиперболоид $\mathbb{L}^2$, вложенный в трехмерное пространство Минковского, с индуцированной метрикой или его универсальной накрывающей (раздел 29.3). Его группой симметрии является группа Лоренца $\mathbb{SO}(1, 2)$. В стереографических координатах метрика однополостного гиперболоида $\mathbb{L}^2$ имеет хорошо известный вид (29.36)

$$ds_{\mathbb{L}}^2 = g_{\alpha\beta} dx^\alpha dx^\beta = \frac{dt^2 - dx^2}{\left[1 + \frac{K}{4}(t^2 - x^2)\right]^2}, \quad (33.24)$$

где введены обозначения $t := x^0$ и $x := x^1$. В отличие от риманова случая псевдориманова поверхность постоянной кривизны одна и та же как для положительной, так и для отрицательной гауссовой кривизны K , при этом меняется только общий знак метрики (33.24), что соответствует перестановке координат $t \leftrightarrow x$.

При $K = 0$ метрика (33.24) совпадает с обычной двумерной метрикой Минковского, и соответствующая поверхность представляет собой плоскость Минковского $\mathbb{R}^{1,1}$ с группой Пуанкаре $\mathbb{IO}(1, 1)$ в качестве группы симметрии.

Положительно определенная метрика на двумерной римановой поверхности постоянной кривизны $R^h = -2K \neq 0$ в стереографических координатах имеет вид (29.24)

$$ds_{\mathbb{H}}^2 = h_{\mu\nu} dy^\mu dy^\nu = \frac{dy^2 + dz^2}{\left[1 + \frac{K}{4}(y^2 + z^2)\right]^2}, \quad (33.25)$$

где $y := y^1$ и $z := y^2$. Эта метрика отличается от (33.24) только знаками.

Для положительных $K > 0$ она соответствует сфере $\mathbb{S}^2$, рассмотренной в разделе 29.1. При $K = 0$ метрика (33.25) соответствует евклидовой плоскости $\mathbb{R}^2$, или цилиндру, или тору. При отрицательных $K < 0$ мы имеем плоскость Лобачевского (гиперболическую плоскость) $\mathbb{H}^2$, рассмотренную в разделе 29.2, или компактную риманову поверхность рода два или выше. Группами симметрии сферы $\mathbb{S}^2$, евклидовой плоскости $\mathbb{R}^2$, и плоскости Лобачевского $\mathbb{H}^2$ являются, соответственно, группы $\mathbb{O}(3)$, $\mathbb{IO}(2)$ и $\mathbb{O}(1, 2)$.

При ненулевой постоянной кривизне всегда можно произвести растяжку координат таким образом, чтобы $K = \pm 1$.

Если скалярные кривизны постоянны (33.22), то решение для ненулевой космологической постоянной $\Lambda \neq 0$ является топологическим произведением двух поверхностей постоянной кривизны с метрикой

$$ds^2 = k \frac{dt^2 - dx^2}{\left[1 - \frac{k\Lambda}{4}(t^2 - x^2)\right]^2} + m \frac{dy^2 + dz^2}{\left[1 - \frac{m\Lambda}{4}(y^2 + z^2)\right]^2}. \quad (33.26)$$

В данном случае можно не говорить о сплетенном произведении поверхностей, т.к. дилатонные поля постоянны. Растягивая координаты, всегда можно добиться выполнения равенств $k = \pm 1$, $m = \pm 1$. Выберем $k = 1$ и $m = -1$ с тем, чтобы метрика имела сигнатуру $(+ - - -)$. Тогда возможны три качественно отличных случая, соответствующих положительной, нулевой и отрицательной космологической постоянной:

$$\begin{aligned} \Lambda < 0: & \quad K^g = +|\Lambda|, \quad K^h = -|\Lambda|, \quad \mathbb{M} = \mathbb{L}^2 \times \mathbb{H}^2, \\ \Lambda = 0: & \quad K^g = 0, \quad K^h = 0, \quad \mathbb{M} = \mathbb{R}^{1,1} \times \mathbb{R}^2 = \mathbb{R}^{1,3}, \\ \Lambda > 0: & \quad K^g = -|\Lambda|, \quad K^h = +|\Lambda|, \quad \mathbb{M} = \mathbb{L}^2 \times \mathbb{S}^2, \end{aligned} \quad (33.27)$$

где K^g и K^h – гауссовы кривизны, соответственно, поверхностей $\mathbb{U}$ и $\mathbb{V}$.

Напомним, что вакуумные уравнения Эйнштейна (33.3) допускают решение в виде пространства-времени постоянной кривизны, которое называется пространством-временем (анти-) де Ситтера. Это пространство-время имеет максимальное число – десять – векторов Киллинга. Хотя полная (четырёхмерная) скалярная кривизна для решения в виде произведения двух поверхностей постоянной кривизны с метрикой (33.26), впрочем как и для всех других решений вакуумных уравнений Эйнштейна (33.3), постоянна, $\hat{R} = 4\Lambda$, решения (33.27) при $\Lambda \neq 0$ не совпадают с решением (анти-) де Ситтера. Действительно, каждая из поверхностей $\mathbb{L}^2$, $\mathbb{H}^2$ и $\mathbb{S}^2$ имеет по три вектора Киллинга, и можно показать (см. например [252]), что четырёхмерное

пространство-время имеет всего шесть векторов Киллинга. Поэтому решения в виде произведения двух поверхностей не совпадают с решением (анти-) де Ситтера.

Все решения в виде произведения двух поверхностей постоянной кривизны относятся к типу D по классификации Петрова [122].

33.4 Пространственно симметричные решения

Во втором случае В (33.21) дилатонное поле k постоянно. Не ограничивая общности, положим $k = 1$. Тогда вся система уравнений Эйнштейна (33.16)–(33.20) сводится к следующей системе:

$$\nabla_\alpha \nabla_\beta m - \frac{\nabla_\alpha m \nabla_\beta m}{2m} - \frac{1}{2} g_{\alpha\beta} \left[\nabla^2 m - \frac{(\nabla m)^2}{2m} \right] = 0, \quad (33.28)$$

$$R^h + \nabla^2 m - 2m\Lambda = 0, \quad (33.29)$$

$$R^g + \frac{\nabla^2 m}{m} - \frac{(\nabla m)^2}{2m^2} - 2\Lambda = 0. \quad (33.30)$$

Уравнение (33.29) представляет собой сумму двух слагаемых, зависящих от координат на разных поверхностях, $x \in \mathbb{U}$ и $y \in \mathbb{V}$, которая должна быть равна нулю. Это значит, что каждое слагаемое равно некоторой постоянной. Зафиксируем эту постоянную следующим уравнением $R^h = -2K = \text{const}$. Таким образом, в случае В поверхность $\mathbb{V}$ – это поверхность постоянной кривизны. При этом возможны три случая, когда гауссова кривизна римановой поверхности $\mathbb{V}$ положительна, $\mathbb{V} = \mathbb{S}^2$, равна нулю, $\mathbb{V} = \mathbb{R}^2$, или отрицательна, $\mathbb{V} = \mathbb{H}^2$. Тогда соответствующие решения вакуумных уравнений Эйнштейна инвариантны относительно групп преобразований $\mathbb{O}(3)$, $\mathbb{IO}(2)$ или $\mathbb{O}(1, 2)$, которые являются группами изометрий поверхностей $\mathbb{S}^2$, $\mathbb{R}^2$ и $\mathbb{H}^2$. Соответствующее четырехмерное пространство-время представляет собой сплетенное произведение поверхности $\mathbb{U}$ с одной из поверхностей $\mathbb{S}^2$, $\mathbb{R}^2$ или $\mathbb{H}^2$, где $\mathbb{U}$ представляется диаграммой Картера–Пенроуза. В частности, при $K = 1$ возникают сферически симметричные решения. Таким образом, в рассматриваемом случае группа симметрии пространства-времени возникает как следствие уравнений движения.

При $K = \text{const}$ уравнение (33.29) принимает вид

$$\nabla^2 m - 2(m\Lambda + K) = 0. \quad (33.31)$$

Исключая случай А, рассмотренный в предыдущем разделе, двинемся дальше, считая, что $\nabla_\alpha m \neq 0$.

Предложение 33.4.1. Уравнение (33.31) является первым интегралом уравнений (33.28) и (33.30).

Доказательство. Продифференцируем уравнение (33.31), используем тождество

$$[\nabla_\alpha, \nabla_\beta] A_\gamma = -R_{\alpha\beta\gamma}^{\delta} A_\delta,$$

где A_α – компоненты произвольного ковекторного поля, для изменения порядка ковариантных производных и используем уравнение (33.28) три раза для исключения вторых производных от m . После небольших алгебраических выкладок мы получим уравнение (33.30). $\square$

Из доказательства предложения следует, что достаточно решить только уравнения (33.28) и (33.31), при этом уравнение (33.30) будет удовлетворено автоматически.

Замечание. Исходное действие Гильберта–Эйнштейна инвариантно относительно общих преобразований координат, и, согласно второй теореме Нетер, между уравнениями движения существует линейная зависимость. Поэтому зависимость уравнений (33.28)–(33.30) не является чем то удивительным и связана с инвариантностью исходного действия. $\square$

Для явного решения уравнений движения (33.28) и (33.31) зафиксируем конформную калибровку для метрики $g_{\alpha\beta}$ на лоренцевой поверхности $\mathbb{U}$:

$$g_{\alpha\beta} dx^\alpha dx^\beta = \Phi d\xi d\eta, \quad (33.32)$$

где $\Phi(\xi, \eta)$ – конформный множитель, который зависит от координат светового конуса ξ, η на $\mathbb{U}$. Соответствующая четырехмерная метрика примет вид

$$ds^2 = \Phi d\xi d\eta + m d\Omega, \quad (33.33)$$

где $d\Omega$ – метрика на римановой поверхности постоянной кривизны $\mathbb{V} = \mathbb{S}^2, \mathbb{R}^2$ или $\mathbb{H}^2$. Знак конформного множителя Φ пока не фиксируем.

При $\Phi > 0$ и $m < 0$ сигнатура метрики (33.33) равна $(+ - - -)$, если пространственная и временная координаты определены так же, как и раньше (30.23). Если изменить знак $m > 0$, то сигнатура метрики станет $(+ - ++)$. Такое же преобразование сигнатуры можно получить, изменив общий знак метрики, $\hat{g}_{ij} \mapsto -\hat{g}_{ij}$, и переставив первые две координаты. Вакуумные уравнения Эйнштейна (33.3) инвариантны относительно одновременного изменения знаков метрики и космологической постоянной. Поскольку в дальнейшем мы построим глобальные решения для всех возможных значений космологической постоянной, $\Lambda \in \mathbb{R}$, то, не ограничивая общности, достаточно рассмотреть случай $m < 0$. При отрицательных m удобно ввести параметризацию

$$m = -q^2, \quad q(\xi, \eta) > 0. \quad (33.34)$$

Символы Кристоффеля для метрики (33.32) в конформной калибровке имеют только две ненулевые компоненты:

$$\Gamma_{\xi\xi}^\xi = \frac{\partial_\xi \Phi}{\Phi}, \quad \Gamma_{\eta\eta}^\eta = \frac{\partial_\eta \Phi}{\Phi}, \quad (33.35)$$

и уравнения (33.28), (33.31) принимают простой вид

$$-\partial_{\xi\xi}^2 q + \frac{\partial_\xi \Phi \partial_\xi q}{\Phi} = 0, \quad (33.36)$$

$$-\partial_{\eta\eta}^2 q + \frac{\partial_\eta \Phi \partial_\eta q}{\Phi} = 0, \quad (33.37)$$

$$-2 \frac{\partial_{\xi\eta}^2 q^2}{\Phi} - (K - \Lambda q^2) = 0. \quad (33.38)$$

Таким образом, полная система уравнений (33.28)–(33.30) в конформной калибровке (33.32) сводится к трем уравнениям на две неизвестные функции q и Φ . Первые два уравнения являются обыкновенными дифференциальными уравнениями, и определяют функции q и Φ с точностью до умножения на произвольную постоянную. Система уравнений (33.36)–(33.38) переопределена и может быть проинтегрирована явно.

Предложение 33.4.2. Условия $\partial_\xi q = 0$ и $\partial_\eta q = 0$ эквивалентны.

Доказательство. Если $\partial_\xi q = 0$, то из уравнения (33.38) следует $q^2 = K/\Lambda = \text{const}$ и, следовательно, $\partial_\eta q = 0$. Обратное утверждение верно по той же причине. $\square$

Поскольку $q = \text{const}$ соответствует уже рассмотренному случаю А, то предположим, что $\partial_\xi q \neq 0$ и $\partial_\eta q \neq 0$. Тогда, разделив уравнения (33.36) и (33.37), соответственно, на $\partial_\xi q$ и $\partial_\eta q$, они легко интегрируются:

$$-\ln|\partial_\xi q| + \ln|\Phi| = \tilde{G}(\eta), \quad (33.39)$$

$$-\ln|\partial_\eta q| + \ln|\Phi| = \tilde{F}(\xi). \quad (33.40)$$

При этом возникают две произвольные функции $\tilde{F}(\xi)$ и $\tilde{G}(\eta)$. Введем монотонные функции $F(\xi)$ и $G(\eta)$ при помощи дифференциальных уравнений

$$F' := \frac{dF}{d\xi} = C e^{\tilde{F}} > 0, \quad G' := \frac{dG}{d\eta} = C e^{\tilde{G}} > 0,$$

где $C > 0$ – некоторая положительная постоянная, которую мы зафиксируем чуть позже. Тогда разность уравнений (33.39) и (33.40) примет вид

$$\frac{|\partial_\xi q|}{F'} = \frac{|\partial_\eta q|}{G'}. \quad (33.41)$$

Конформная калибровка для двумерной метрики (33.32) определена с точностью до конформных преобразований. Воспользуемся этой свободой и перейдем к новым координатам $\xi, \eta \mapsto F, G$. Это всегда можно сделать, т.к. якобиан преобразования координат отличен от нуля, $F'G' \neq 0$. При конформном преобразовании координат конформный множитель преобразуется по правилу $\Phi \mapsto \Phi/(F'G')$, что следует из вида метрики в конформной калибровке (33.32).

Предложение 33.4.3. Уравнения (33.36)–(33.38) ковариантны относительно конформных преобразований

$$\xi, \eta \mapsto F, G, \quad \Phi \mapsto \tilde{\Phi} = \frac{\Phi}{F'G'}. \quad (33.42)$$

Доказательство. Прямая проверка. $\square$

Таким образом, произвольные функции $\tilde{F}(\xi)$ и $\tilde{G}(\eta)$, возникшие в первых интегралах (33.39), (33.40), соответствуют конформным преобразованиям.

Перейдем к новым координатам $\xi, \eta \mapsto F, G$.

Предложение 33.4.4. Если $\partial_\xi q \partial_\eta q > 0$, то в новых координатах функция $q(\tau)$ зависит только от временной координаты $\tau := \frac{1}{2}(F + G)$. Если $\partial_\xi q \partial_\eta q < 0$, то функция $q(\sigma)$ зависит только от пространственной координаты $\sigma := \frac{1}{2}(F - G)$.

Доказательство. Из-за знаков модулей в уравнении (33.41), возможны два случая.

Если $\partial_\xi q \partial_\eta q > 0$, то справедливо равенство

$$\frac{\partial q}{\partial(F - G)} = \partial_\xi q \frac{\partial \xi}{\partial(F - G)} + \partial_\eta q \frac{\partial \eta}{\partial(F - G)} = \frac{\partial_\xi q}{F'} - \frac{\partial_\eta q}{G'} = 0. \quad (33.43)$$

Последнее равенство вытекает из уравнения (33.41). Поэтому, переходя к координатам $F, G \mapsto \tau, \sigma$, получаем сделанное утверждение.

Аналогично, если $\partial_\xi q \partial_\eta q < 0$, то выполнено равенство

$$\frac{\partial q}{\partial(F + G)} = \partial_\xi q \frac{\partial \xi}{\partial(F + G)} + \partial_\eta q \frac{\partial \eta}{\partial(F + G)} = \frac{\partial_\xi q}{F'} + \frac{\partial_\eta q}{G'} = 0. \quad \square \quad (33.44)$$

Теперь из каждого из двух уравнений (33.39) или (33.40) следует одно и то же равенство

$$|\Phi| = \frac{1}{2C} F' G' |q'|,$$

где q' обозначает производную функции q либо по $\tau := \frac{1}{2}(F + G)$, либо по $\sigma := \frac{1}{2}(F - G)$. Постоянная C соответствует растяжке новых координат F, G , и, для упрощения последующих формул, положим $C = 1/2$. При конформном преобразовании конформный множитель преобразуется по правилу (33.42). Поэтому после конформного преобразования (33.42) будет выполнено равенство

$$|\tilde{\Phi}| = |q'|. \quad (33.45)$$

В дальнейшем знак тильды у конформного множителя мы, для краткости, опустим.

Таким образом, координаты всегда можно выбрать таким образом, чтобы функции q и Φ зависели одновременно только от времениподобной или пространственно-подобной координаты

$$\zeta = \frac{1}{2}(F \pm G) := \begin{cases} \tau, & \partial_\xi q \partial_\eta q > 0, \\ \sigma, & \partial_\xi q \partial_\eta q < 0. \end{cases} \quad (33.46)$$

Это значит, что двумерная метрика (33.32) обладает вектором Киллинга, ∂_σ или ∂_τ , как следствие уравнений (33.36) и (33.37). Назовем эти решения, соответственно, однородными и статическими, хотя это и относится только к определенной системе координат. Существование вектора Киллинга является обобщением *теоремы Бирхгоффа* [253], утверждающей, что произвольное сферически симметричное решение вакуумных уравнений Эйнштейна должно быть статическим. (Это утверждение было опубликовано ранее в статье [254].) Обобщение заключается в том, что наличие вектора Киллинга доказано не только для сферически симметричных решений ($K = 1$), но и для решений, инвариантных относительно групп преобразований $\mathbb{IO}(2)$ ($K = 0$) и $\mathbb{SO}(1, 2)$ ($K = -1$).

Окончательно, решением уравнений (33.36), (33.37) в фиксированной системе координат является равенство (33.45) и утверждение о том, что функции q и Φ зависят только от одной переменной ζ (33.46). Осталось решить только одно уравнение (33.38).

В статическом, $q = q(\sigma)$, и однородном, $q = q(\tau)$, случаях уравнение (33.38) принимает вид

$$(q^2)'' = 2(K - \Lambda q^2)\Phi, \quad q = q(\sigma), \quad (33.47)$$

$$(q^2)'' = -2(K - \Lambda q^2)\Phi, \quad q = q(\tau). \quad (33.48)$$

Чтобы проинтегрировать полученные уравнения, необходимо выразить Φ через q с помощью уравнения (33.45), а для этого необходимо раскрыть знаки модулей.

Рассмотрим подробно статический случай $q = q(\sigma)$, $\Phi > 0$ и $q' > 0$. Тогда уравнение (33.47) с учетом (33.45) примет вид

$$(q^2)'' = 2(K - \Lambda q^2)q'.$$

Его легко проинтегрировать

$$(q^2)' = 2 \left(Kq - \frac{\Lambda q^3}{3} - 2M \right),$$

где $M = \text{const}$ – постоянная интегрирования. В дальнейшем мы увидим, что она совпадает с массой в решении Шварцшильда. Выполнив дифференцирование в левой части и поделив на $2q > 0$, получим уравнение

$$q' = K - \frac{2M}{q} - \frac{\Lambda q^2}{3}.$$

Поскольку в рассматриваемом случае $q' = \Phi$, то отсюда следует выражение для конформного множителя через переменную q :

$$\Phi(q) = K - \frac{2M}{q} - \frac{\Lambda q^2}{3}. \quad (33.49)$$

Если $q = q(\sigma)$, $\Phi > 0$ и $q' < 0$, то аналогичное интегрирование приводит к уравнению

$$q' = -\Phi(q),$$

где в правой части стоит тот же самый конформный множитель (33.49). Этот случай можно объединить с предыдущим, записав уравнение для q в виде

$$|q'| = \Phi(q), \quad q = q(\sigma), \quad \Phi > 0. \quad (33.50)$$

Аналогично интегрируется статический случай при $\Phi < 0$:

$$|q'| = -\Phi(q), \quad q = q(\sigma), \quad \Phi < 0. \quad (33.51)$$

Если решение однородно, $q = q(\tau)$ и $\Phi > 0$, $q' > 0$, то интегрирование уравнения (33.48) приводит к равенству

$$q' = -\left(K - \frac{2M}{q} - \frac{\Lambda q^2}{3}\right).$$

То есть в этом случае конформный множитель надо отождествить с правой частью

$$\hat{\Phi} = -\left(K - \frac{2M}{q} - \frac{\Lambda q^2}{3}\right). \quad (33.52)$$

Поскольку выражение конформного множителя в однородном случае через q отличается знаком, то мы пометили его шляпкой. Таким образом, однородные решения уравнений Эйнштейна можно записать в виде

$$|q'| = \hat{\Phi}(q), \quad q = q(\tau), \quad \hat{\Phi} > 0. \quad (33.53)$$

$$|q'| = -\hat{\Phi}(q), \quad q = q(\tau), \quad \hat{\Phi} < 0. \quad (33.54)$$

Если конформный множитель отрицателен, то сигнатура метрики равна $(-+-)$. В этом случае, сделав замену $\tau \leftrightarrow \sigma$, мы вернемся к прежней сигнатуре метрики $(+---)$. Это преобразование позволяет объединить стационарные и однородные решения, написав знак модуля у конформного множителя в выражении для метрики (33.33). Тогда общее решение вакуумных уравнений Эйнштейна (33.3) в соответствующей системе координат примет вид

$$ds^2 = |\Phi|(d\tau^2 - d\sigma^2) - q^2 d\Omega, \quad (33.55)$$

где конформный множитель Φ имеет вид (33.49). При этом переменная q зависит либо от σ (статическое локальное решение), либо от τ (однородное локальное решение) через дифференциальное уравнение

$$\left| \frac{dq}{d\zeta} \right| = \pm \Phi(q), \quad (33.56)$$

где выполнено правило знаков:

$$\begin{aligned} \Phi > 0 : \quad \zeta = \sigma, \quad \text{знак} + \text{ (статическое локальное решение),} \\ \Phi < 0 : \quad \zeta = \tau, \quad \text{знак} - \text{ (однородное локальное решение).} \end{aligned} \quad (33.57)$$

Таким образом, из четырехмерных уравнений Эйнштейна вытекает, что на поверхности $\mathbb{U}$ возникает метрика с одним вектором Киллинга, которая была подробно рассмотрена в главе 30. Теперь с помощью метода конформных блоков можно построить глобальные (максимально продолженные вдоль экстремалей) решения вакуумных уравнений Эйнштейна. Число особенностей и нулей конформного множителя (33.49) зависит от соотношения между постоянными K , M и Λ . Поэтому возможно существование многих существенно различных глобальных решений, которые мы рассмотрим в следующих разделах.

Конформный множитель (33.49) имеет одну сингулярность: простой полюс при $q = 0$. Поэтому, согласно правилам построения глобальных решений из раздела 30.4 каждое глобальное решение соответствует одному из интервалов $(-\infty, 0)$ или $(0, \infty)$. Из вида конформного множителя (33.62) следует, что эти глобальные решения получаются друг из друга преобразованием $M \mapsto -M$. Поэтому, не ограничивая общности, мы опишем только глобальные решения, соответствующие положительному интервалу $(q_-, q_+) = (0, \infty)$. Впрочем, это предположение уже было сделано при параметризации (33.34).

Поскольку конформный множитель $\Phi(q)$ является гладкой функцией при $q > 0$, то все возникающие лоренцевы поверхности $\mathbb{U}$ и метрика на них, являются гладкими.

Используя уравнения (33.30), (33.34) и (33.47), нетрудно вычислить скалярную кривизну поверхности $\mathbb{U}$:

$$R^g = \frac{2}{3}\Lambda + \frac{4M}{q^3}. \quad (33.58)$$

Она не зависит от гауссовой кривизны K римановой поверхности $\mathbb{V}$ и сингулярна при $q = 0$, если $M \neq 0$. Отметим, что сингулярная часть двумерной скалярной кривизны (33.58) пропорциональна собственному значению четырехмерного тензора Вейля (см., например, [169]):

$$\frac{1}{48} \hat{C}_{ijkl} \hat{C}^{ijkl} = \left(-\frac{M}{q^3} \right)^2. \quad (33.59)$$

Теперь перейдем к описанию всех пространственно симметричных глобальных решений вакуумных уравнений Эйнштейна.

33.4.1 Сферически симметричные решения $K = 1$

При $K = 1$ риманова поверхность $\mathbb{V}$ представляет собой сферу $\mathbb{S}^2$, и все решения сферически симметричны. Для сферы единичного радиуса метрику (33.25) запишем в сферических координатах

$$d\Omega = d\theta^2 + \sin^2\theta d\varphi^2. \quad (33.60)$$

Сферически симметричную метрику пространства-времени, которая удовлетворяет уравнениям Эйнштейна, можно записать в виде

$$ds^2 = |\Phi(q)|(d\tau^2 - d\sigma^2) - q^2(d\theta^2 + \sin^2\theta d\varphi^2), \quad (33.61)$$

где

$$\Phi(q) = 1 - \frac{2M}{q} - \frac{\Lambda q^2}{3}. \quad (33.62)$$

Переменная q связана с σ или τ дифференциальным уравнением (33.56), где выполнено правило знаков (33.57).

Определение. Координаты τ, σ , в которых записана сферически симметричная метрика (33.61), называются *черепашиными*. Это название, по-видимому, произошло потому что диаграммы Картера–Пенроуза чем то напоминают рисунок панциря черепахи. $\square$

Обобщение решения Шварцшильда на случай ненулевой космологической постоянной (33.61) было получено Коттлером [255].

В рассматриваемом случае все решения параметризуются двумя постоянными: космологической постоянной Λ и массой M . Вторую постоянную мы будем называть массой, хотя она и не имеет физического смысла массы для большинства решений, отличных от решения Шварцшильда.

Как было отмечено в предыдущем разделе, мы рассматриваем глобальные решения вакуумных уравнений Эйнштейна, соответствующих положительному интервалу $(q_-, q_+) = (0, \infty)$. Структура глобальных решений определяется количеством и типом положительных корней уравнения $\Phi(q) = 0$ или эквивалентного кубического уравнения

$$\frac{\Lambda}{3}q^3 - q + 2M = 0. \quad (33.63)$$

При ненулевой космологической постоянной это уравнение может иметь до трех нулей. Элементарный анализ показывает, что по крайней мере один из корней отрицателен. Это значит, что при положительных q возможно существование не более двух горизонтов.

При положительной космологической постоянной $\Lambda > 0$ существуют следующие возможности в зависимости от величины массы M (см. рис.33.1 слева). Если $M > \frac{1}{3\sqrt{\Lambda}}$, то уравнение (33.63) не имеет положительных корней. При $M = \frac{1}{3\sqrt{\Lambda}}$ возникает один положительный корень второго порядка. В интервале значений $0 < M < \frac{1}{3\sqrt{\Lambda}}$ существуют два положительных корня. При неположительных значениях $M \leq 0$ имеется один простой положительный нуль.

При отрицательных значениях космологической постоянной $\Lambda < 0$ мы имеем один простой положительный нуль при $M > 0$ и ни одного нуля при $M \leq 0$ (см. рис.33.1 справа).

Перейдем к классификации решений.

Пространство-время Минковского $\Lambda = 0, M = 0$

Наиболее простое сферически симметричное решение получается при $\Lambda = 0$ и $M = 0$ (см. рис.33.2). В этом случае $\Phi = 1$, и метрика принимает вид

$$ds^2 = d\tau^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2), \quad r \in (0, \infty), \quad (33.64)$$

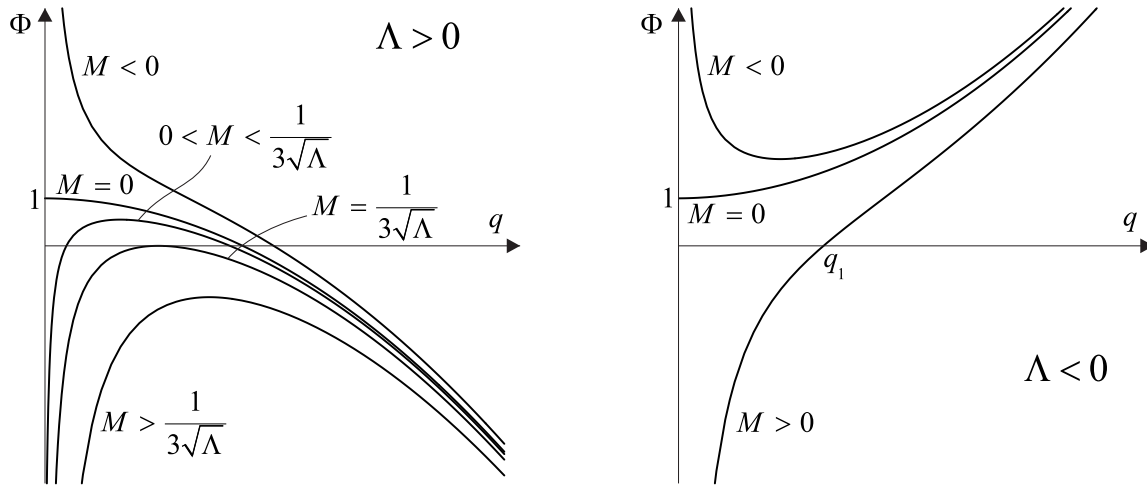


Рис. 33.1: Поведение конформного множителя $\Phi = 1 - \frac{2M}{q} - \frac{\Lambda q^2}{3}$ в зависимости от величины постоянной M (массы) при положительной (слева) и отрицательной (справа) космологической постоянной Λ .

где мы переобозначили $\sigma \mapsto r$. Точка $r = 0$ является координатной сингулярностью. Переходя к декартовым координатам в четырехмерном пространстве-времени и добавляя мировую линию начала сферической системы координат $r = 0$, мы получим пространство-время Минковского $\mathbb{R}^{1,3}$. При этом пространственная координата r естественным образом отождествляется с радиусом сферической системы координат. В этом случае пространство-время нельзя представить в виде топологического произведения $\mathbb{U} \times \mathbb{S}^2$, и поэтому диаграмма Картера–Пенроуза на рис.33.2 не продолжается.

Черная дыра Шварцшильда $\Lambda = 0$, $M > 0$

Решение Шварцшильда соответствует нулевой космологической постоянной $\Lambda = 0$ и положительной массе $M > 0$. Оно уже обсуждалось в разделе 30.5.1. В этом случае конформный множитель (33.62) имеет один простой нуль в точке $q_1 = 2M$ и, следовательно, один горизонт. Соответствующая диаграмма Картера–Пенроуза также изображена на рис.33.2.

Голая сингулярность $\Lambda = 0$, $M < 0$

При нулевой космологической постоянной $\Lambda = 0$ и отрицательной массе $M < 0$ уравнение (33.63) не имеет положительных корней. Этот случай также был рассмотрен в разделе 30.5.1. При указанных значениях постоянных существуют две диаграммы Картера–Пенроуза, каждая из которых состоит из одного треугольного конформного блока, изображенного на рис.33.2, и его пространственного отражения. Каждый из конформных блоков представляет собой максимально продолженную поверхность $\mathbb{U}$. Сингулярная граница времениподобна и называется *голой сингулярностью*, т.к. не окружена горизонтом.

Решение де Ситтера $\Lambda > 0$, $M = 0$

Решение де Ситтера (см. раздел 35.4.2) соответствует положительной космологической постоянной и нулевой массе. При этом пространство-время представляет собой

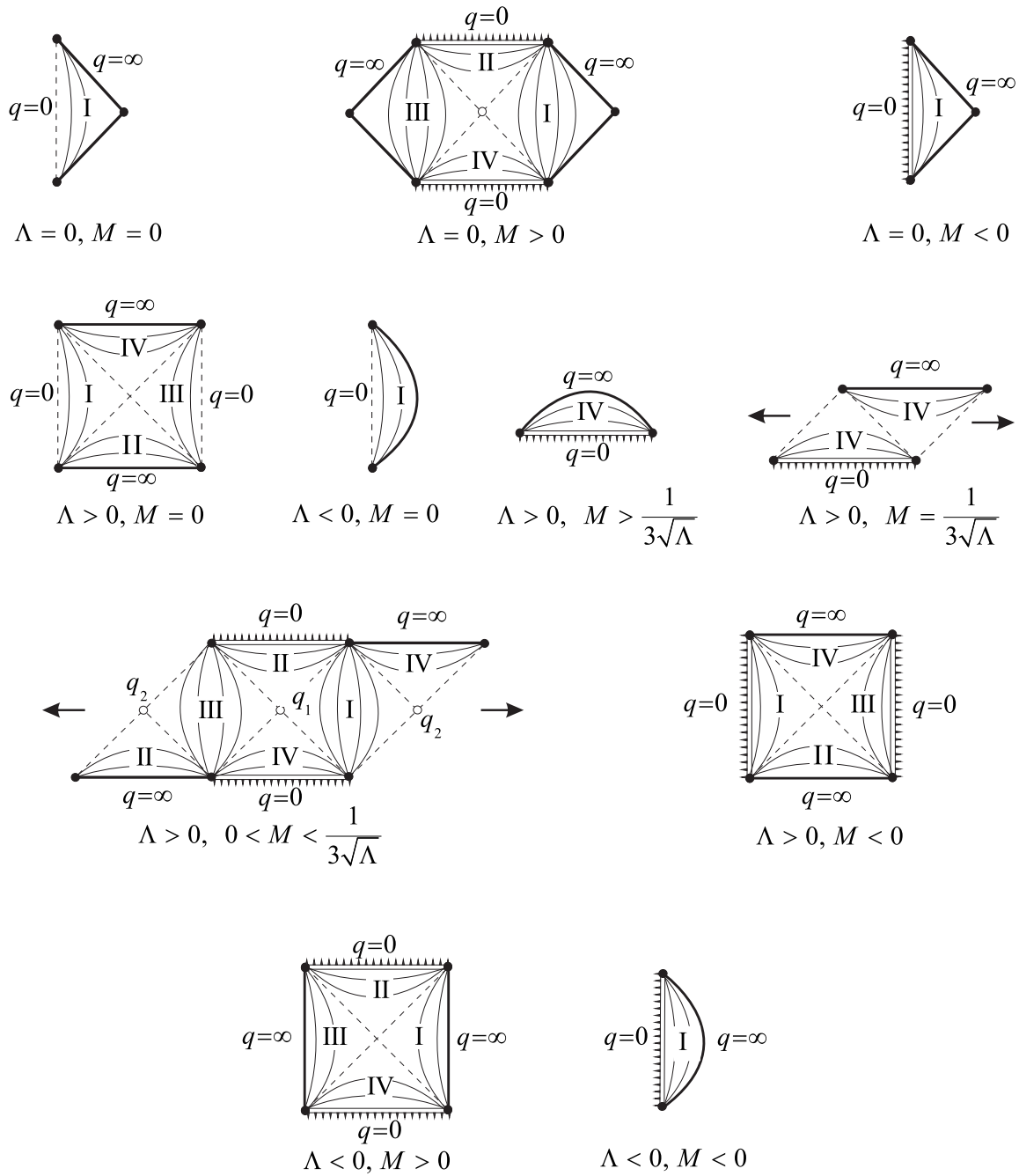


Рис. 33.2: Сферически симметричные решения для конформного множителя $\Phi = 1 - \frac{2M}{q} - \frac{\Lambda q^2}{3}$. Условные обозначения те же, что и на рис.30.3. Стрелки показывают возможное периодическое продолжение решений.

многообразие постоянной (в наших обозначениях положительной) скалярной кривизны и может быть представлено, как четырехмерный гиперboloид, вложенный в пятимерное пространство Минковского $\mathbb{R}^{1,4}$, с индуцированной метрикой. Его группой симметрии является группа Лоренца $\mathbb{O}(1, 4)$, а метрика имеет максимальное число – десять – векторов Киллинга, что совпадает с размерностью группы симметрии. Конформный множитель (33.62) имеет один простой положительный нуль, соответствующий горизонту.

Статические и однородные решения в координатах Шварцшильда имеют вид

$$ds^2 = \left(1 - \frac{\Lambda}{3}r^2\right) d\tau^2 - \frac{dr^2}{1 - \frac{\Lambda}{3}r^2} - r^2 d\Omega^2, \quad (33.65)$$

$$\begin{aligned} 0 < r < \sqrt{\frac{3}{\Lambda}}, \quad -\infty < \tau < \infty, \\ ds^2 = -\frac{dt^2}{1 - \frac{\Lambda}{3}t^2} + \left(1 - \frac{\Lambda}{3}t^2\right) d\sigma^2 - t^2 d\Omega^2, \\ \sqrt{\frac{3}{\Lambda}} < t < \infty, \quad -\infty < \sigma < \infty. \end{aligned} \quad (33.66)$$

Поскольку уравнение (33.56) в этом случае интегрируется явно, то двумерную часть метрики можно записать также в конформно плоском виде

$$ds^2 = \frac{1}{\operatorname{ch}^2\left(\sqrt{\frac{\Lambda}{3}}\sigma\right)} (d\tau^2 - d\sigma^2) - \frac{3}{\Lambda} \operatorname{th}^2\left(\sqrt{\frac{\Lambda}{3}}\sigma\right) d\Omega^2, \quad (33.67)$$

$$ds^2 = \frac{1}{\operatorname{sh}^2\left(\sqrt{\frac{\Lambda}{3}}\tau\right)} (d\tau^2 - d\sigma^2) - \frac{3}{\Lambda} \operatorname{cth}^2\left(\sqrt{\frac{\Lambda}{3}}\tau\right) d\Omega^2, \quad (33.68)$$

соответственно, для статического и однородного случая. Область определения $r \in (0, \sqrt{3/\Lambda})$ переходит в $\sigma \in (0, \infty)$, а $t \in (\sqrt{3/\Lambda}, \infty)$ – в $\tau \in (0, \infty)$.

Поведение конформного множителя показано на рис.33.1 слева. Он имеет один простой нуль в некоторой точке, которую мы обозначим q_1 . В интервале $(0, q_1)$ конформный множитель положителен и ему ставится в соответствие два треугольных статических конформных блока I и III с метрикой (33.67). В интервале (q_1, ∞) конформный множитель отрицателен, и ему соответствуют два треугольных однородных конформных блока II и IV с метрикой (33.68). Эти четыре конформных блока склеиваются единственным образом по правилам, сформулированным в разделе 30.4. Соответствующая диаграмма Картера–Пенроуза показана на рис.33.2. Решение де Ситтера получится после добавления времениподобной границы $r = 0$ к сплетенному произведению $\mathbb{U} \times \mathbb{S}^2$.

Заметим, что конформный множитель для решения де Ситтера является квадратичным полиномом, $\Phi = 1 - \frac{\Lambda q^2}{3}$. Как показано в разделе 30.5.5, при максимальном продолжении поверхности с такой метрикой возникает однополостный гиперboloид. В рассматриваемом случае мы этого не делаем не потому что возникают какие то проблемы с поверхностью $\mathbb{U}$, а потому что при $q = 0$ коэффициент при угловой части метрики (33.61) становится равным нулю и метрика четырехмерного пространства-времени вырождается. Здесь возникает такая же ситуация, как в пространстве Минковского $\mathbb{R}^{1,3}$ в сферической системе координат.

Решение анти-де Ситтера $\Lambda < 0$, $M = 0$

Изменение знака космологической постоянной в уравнениях Эйнштейна приводит к качественному изменению решений. При нулевой массе мы получаем решение анти-де Ситтера (см. раздел 35.4.3). Соответствующее пространство-время представляет собой многообразие постоянной отрицательной кривизны, которое можно представить в виде четырехмерного гиперboloида, вложенного в плоское пространство $\mathbb{R}^{2,3}$ с метрикой $\eta_{\alpha\beta} = \text{diag}(1, 1, -1, -1, -1)$. Метрика анти де Ситтера симметрична относительно действия группы вращений $\mathbb{O}(2, 3)$ и имеет также максимальное число – десять – векторов Киллинга. Конформный множитель (33.62) не имеет нулей и всегда положителен. Поэтому решение статично и не имеет горизонтов. В координатах Шварцшильда метрика имеет тот же вид (33.65), что и для решения де Ситтера, однако из-за отрицательного знака Λ область изменения r совпадает со всем положительным интервалом $(0, \infty)$. Если двумерная часть метрики является конформно плоской, то метрика анти-де Ситтера принимает вид

$$ds^2 = \frac{1}{\cos^2 \left(\sqrt{\frac{|\Lambda|}{3}} \sigma \right)} (d\tau^2 - d\sigma^2) - \frac{3}{|\Lambda|} \text{th}^2 \left(\sqrt{\frac{|\Lambda|}{3}} \sigma \right) d\Omega. \quad (33.69)$$

При этом координата σ меняется в конечном интервале $\sigma \in \left(0, \frac{\pi}{2} \sqrt{\frac{3}{|\Lambda|}}\right)$. Соответствующая диаграмма Картера–Пенроуза имеет вид линзы и изображена на рис.33.2. Это решение неполно на границе $r = 0$ и не может быть продолжено, что является следствием выбора сферической системы координат. Для получения полного решения анти-де Ситтера мировую линию $r = 0$ необходимо добавить к многообразию, что можно сделать путем перехода к другой системе координат.

Однородная пространственная сингулярность $\Lambda > 0$, $M > \frac{1}{3\sqrt{\Lambda}}$

При $\Lambda > 0$, $M > \frac{1}{3\sqrt{\Lambda}}$ конформный множитель Φ для положительных $q > 0$ не имеет нулей и всегда отрицателен. Решение однородно и не имеет горизонтов. Поверхность $\mathbb{U}$ в этом случае представляется диаграммой Картера–Пенроуза в виде линзы, изображенной на рис.33.2, и ее отражения во времени. При конечном значении времени в прошлом имеется истинная сингулярность и у четырехмерной кривизны, и у тензора кривизны поверхности $\mathbb{U}$. Многообразие не может быть продолжено через эту сингулярность. Космологическая интерпретация этого решения простая. Вселенная рождается в конечном прошлом из сингулярности и развивается вечно. В пределе $t \rightarrow \infty$ метрика на поверхности $\mathbb{U}$ стремится к метрике де Ситтера. Это же верно для трех следующих глобальных решений.

Двойной горизонт $\Lambda > 0$, $M = \frac{1}{3\sqrt{\Lambda}}$

Если $\Lambda > 0$, $M = \frac{1}{3\sqrt{\Lambda}}$ то конформный множитель Φ имеет один нуль второго порядка, соответствующий горизонту $q_1 = 1/\sqrt{\Lambda}$. При $q > 0$ функция Φ неотрицательна и, следовательно, все конформные блоки однородны. Два однородных конформных блока соответствуют интервалу $(0, q_1)$ и два – интервалу (q_1, ∞) . Они склеиваются в фундаментальную область, показанную на рис.33.2. Эту фундаментальную область можно либо продолжить влево и вправо, что приводит к универсальному накрывающему пространству, либо отождествить противоположные неполные края. В последнем случае возникающая поверхность диффеоморфна цилиндру. В случае одного

двойного горизонта имеются также поверхности, возникающие из построенных выше отражением во времени.

Два горизонта $\Lambda > 0$, $0 < M < \frac{1}{3\sqrt{\Lambda}}$

При $\Lambda > 0$, $0 < M < \frac{1}{3\sqrt{\Lambda}}$ конформный множитель Φ имеет два нуля и, значит, два горизонта в точках q_1 и q_2 . Двум интервалам $(0, q_1)$ и $(q_2, 0)$, в которых конформный множитель отрицателен, ставится в соответствие по два однородных треугольных конформных блока. В интервале (q_1, q_2) конформный множитель положителен, и ему ставится в соответствие два статических квадратных конформных блока. Фундаментальная область возникает после склейки всех шести конформных блоков по правилам, описанным в разделе 30.4. Как и в предыдущем примере, фундаментальную область можно либо продолжить влево и вправо, и получить универсальное накрывающее пространство, как показано стрелками на рис.33.2, либо отождествить неполные края.

Две статические сингулярности $\Lambda > 0$, $M < 0$

При положительной космологической постоянной $\Lambda > 0$ и отрицательной массе $M < 0$ конформный множитель Φ имеет один простой нуль в точке q_1 . Интервалу $(0, q_1)$, в котором конформный множитель положителен, ставится в соответствие два треугольных статических конформных блока. Интервалу (q_1, ∞) , где конформный множитель отрицателен, ставится в соответствие два треугольных однородных конформных блока. После склейки этих блоков возникает диаграмма Картера–Пенроуза, показанная на рис.33.2. Мы имеем две сингулярности в статических областях I, III, разделенных горизонтами. Из области I вдоль времениподобной кривой можно достичь только правую сингулярность. Это же верно для области III и левой сингулярности. Из области II имеется возможность попасть в обе сингулярности. Из области IV никакой сингулярности достичь невозможно, и жизнь там продолжается вечно, если вы родились в этой области.

Черная дыра анти-де Ситтера $\Lambda < 0$, $M > 0$

Если $\Lambda < 0$ и $M > 0$, то конформный множитель Φ имеет один нуль и, соответственно, один горизонт в точке q_1 , как показано на рис.33.1 справа. Интервалу $(0, q_1)$, где конформный множитель отрицателен, ставится в соответствие два треугольных однородных конформных блока. В интервале (q_1, ∞) конформный множитель положителен, и ему соответствует два треугольных статических конформных блока. Эти четыре конформных блока склеиваются в диаграмму Картера–Пенроуза, показанную на рис.33.2. Она аналогична диаграмме для шварцшильдовской черной дыры. Отличие заключается в том, что полные левая и правая границы в рассматриваемом случае времениподобны. Сингулярности кривизны расположены в конечном прошлом (белая дыра) и в конечном будущем (черная дыра). Они имеют тот же характер, что и в решении Шварцшильда. Метрика пространства-времени не является асимптотически плоской и стремится к метрике анти-де Ситтера при $q = r \rightarrow \infty$.

Голая сингулярность $\Lambda < 0$, $M < 0$

При $\Lambda < 0$ и $M < 0$ имеется голая сингулярность без горизонтов. Соответствующая диаграмма Картера–Пенроуза совпадает с диаграммой Картера–Пенроуза для $\Lambda > 0$,

$M > \frac{1}{3\sqrt{\Lambda}}$ изображенной на рис.33.2, но повернутой на угол $\pi/2$, т.к. конформный множитель положителен $\Phi > 0$, и, следовательно, метрика статична. Метрика при $q = r \rightarrow \infty$ также стремится к метрике анти-де Ситтера.

33.4.2 Планарные решения $K = 0$

В случае $K = 0$ метрика на римановой поверхности $\mathbb{V}$ (33.25) становится евклидовой

$$d\Omega_p := dy^2 + dz^2. \quad (33.70)$$

Это значит, что соответствующая максимально продолженная поверхность $\mathbb{V}$ является либо евклидовой плоскостью $\mathbb{R}^2$ с группой симметрии Пуанкаре $\mathbb{IO}(2)$, либо ее компактификацией (цилиндр, тор). Будем называть соответствующие четырехмерные глобальные решения вакуумных уравнений Эйнштейна *планарными*. Для решений этого типа метрика в координатах Шварцшильда имеет вид

$$ds^2 = \Phi(q)d\zeta^2 - \frac{dq^2}{\Phi(q)} - q^2\Omega_p, \quad (33.71)$$

где

$$\Phi(q) = -\frac{2M}{q} - \frac{\Lambda q^2}{3}. \quad (33.72)$$

Координаты q и ζ определены в (33.56). Множество планарных решений так же как и в сферически симметричном случае параметризуется двумя постоянными: космологической постоянной Λ и массой M .

Замечание. Для планарных решений координата q не может быть интерпретирована, как радиус пространства. Поэтому мы с самого начала отказались от ее обозначения через r , т.к. решение Шварцшильда является лишь частным случаем описываемого общего подхода. $\square$

Физическая интерпретация планарных решения весьма интересна. Например, при $\mathbb{V} = \mathbb{T}^2$ трехмерное пространство представляет собой прямое произведение тора $\mathbb{T}^2$ и прямой $\mathbb{R}$. Это пространство содержит нестягиваемые замкнутые пространственноподобные кривые, т.е. имеет нетривиальную фундаментальную группу. С физической точки зрения такие пространства описывают кротовые норы. В этом случае все горизонты также представляют собой торы.

Для построения глобальных планарных решений вакуумных уравнений Эйнштейна необходимо знать поведение конформного множителя (33.72) при различных значениях Λ и M . При $\Lambda \neq 0$ оно ясно из рис.33.1 и получается сдвигом оси абсцисс на единицу вверх, как показано на рис.33.3.

При $M = 0$ конформный множитель является квадратичным полиномом. Этот случай был рассмотрен в разделе 30.5.5 и соответствует поверхностям $\mathbb{U}$ постоянной кривизны.

Если космологическая постоянная положительна, $\Lambda > 0$, то конформный множитель при $M > 0$ не имеет нулей. При $M < 0$ существует один положительный нуль.

Для отрицательной космологической постоянной, $\Lambda < 0$, ситуация противоположна. При $M > 0$ есть один простой нуль, а при $M < 0$ нули отсутствуют.

Перечислим все возможные максимально продолженные поверхности $\mathbb{U}$. Как и ранее, достаточно рассмотреть глобальные решения, соответствующие интервалу $q \in (0, \infty)$.

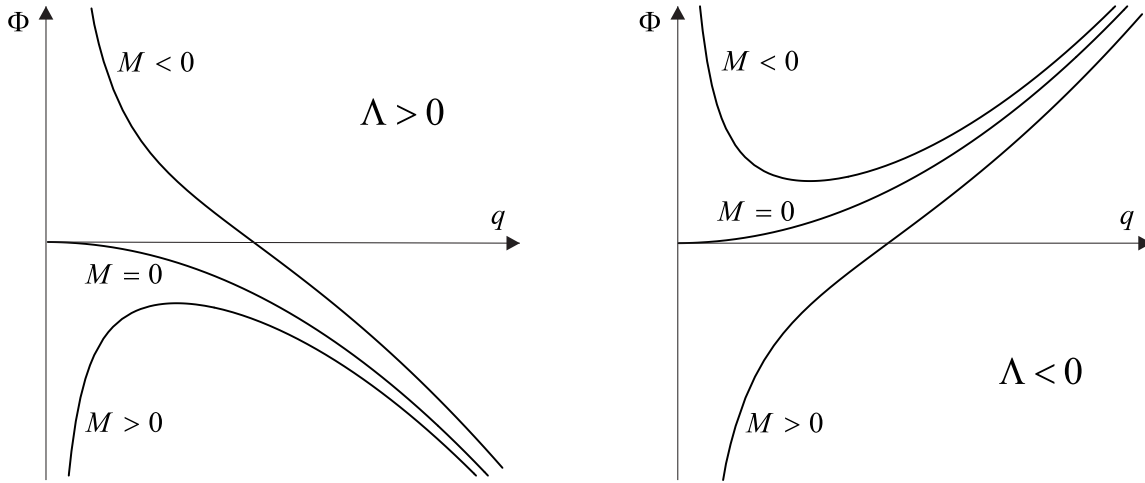


Рис. 33.3: Поведение конформного множителя $\Phi = -\frac{2M}{q} - \frac{\Lambda q^2}{3}$ в зависимости от величины массы M при положительной (слева) и отрицательной (справа) космологической постоянной Λ .

Однородные и голые сингулярности $\Lambda = 0$, $M \neq 0$

При нулевой космологической постоянной и положительной массе $M > 0$ конформный множитель при $q > 0$ отрицателен и не имеет нулей (горизонтов). Поэтому глобальное решение состоит из одного треугольного конформного блока типа IV, однородно и имеет пространственноподобную сингулярность при $q = 0$. В шварцшильдовских и черепаших координатах метрика имеет вид

$$ds^2 = \frac{t}{2M} dt^2 - \frac{2M}{t} d\sigma^2 - t^2 d\Omega_P, \quad t \in (0, \infty), \quad (33.73)$$

$$ds^2 = \sqrt{\frac{M}{|\tau|}} (d\tau^2 - d\sigma^2) - 4M|\tau| d\Omega_P, \quad \tau < 0 \quad \text{или} \quad \tau > 0. \quad (33.74)$$

Соответствующая диаграмма Картера–Пенроуза изображена на рис.33.4.

Как и ранее, существует также глобальное решение типа II, полученное из показанного на рисунке обращением времени.

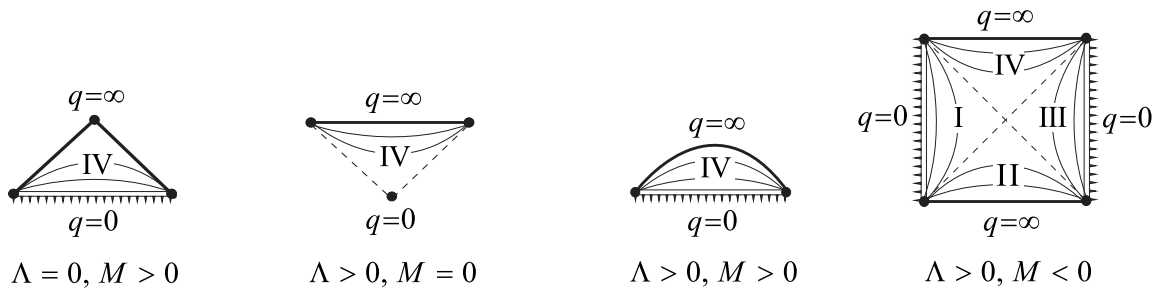


Рис. 33.4: Планы решения $K = 0$. Условные обозначения те же, что и на рис.30.3.

Если масса отрицательна, $M < 0$, то оба глобальных решения имеют такой же вид как и при $M > 0$, но являются статическими. Поэтому соответствующие диаграммы Картера–Пенроуза необходимо просто повернуть на угол $\pi/2$. В таком случае возникают голые сингулярности.

Решение (анти-)де Ситтера $\Lambda \neq 0$, $M = 0$

При положительной космологической постоянной $\Lambda > 0$ и нулевой массе $M = 0$ конформный множитель

$$\Phi = -\frac{\Lambda t^2}{3}, \quad q := t \in (0, \infty),$$

всюду отрицателен и не имеет нулей. Соответствующее глобальное решение однородно и не имеет горизонтов. Оно представляется треугольным конформным блоком типа IV, изображенном на рис.33.4. Есть также глобальное решение типа II, полученное из решения типа IV обращением времени.

В шварцшильдовских и черепаших координатах метрика имеет вид

$$ds^2 = \frac{3}{\Lambda t^2} dt^2 - \frac{\Lambda t^2}{3} d\sigma^2 - t^2 d\Omega_p, \quad t \in (0, \infty), \quad (33.75)$$

$$ds^2 = \frac{3}{\Lambda \tau^2} (d\tau^2 - d\sigma^2) - \frac{9}{\Lambda^2 \tau^2} d\Omega_p, \quad \tau \in (0, \infty). \quad (33.76)$$

Вдоль границы $q = 0$ (т.е. $t = 0$) диаграммы Картера–Пенроуза 4-х мерная метрика вырождается и ее нельзя продолжать. Эти светоподобные края показаны на рис.33.4 пунктиром.

Можно проверить, что метрики (33.75) и (33.76) описывают пространства постоянной четырехмерной кривизны. Это означает, что эти метрики представляют собой решение де Ситтера.

В случае отрицательной космологической постоянной $\Lambda < 0$ зависимость компонент метрики от времени в выражениях (33.75) и (33.76) необходимо заменить на зависимость от пространственной координаты, и соответствующие диаграммы Картера–Пенроуза повернуть на угол $\pi/2$. В этом случае пространство-время снова является пространством постоянной кривизны, и возникает новое представление решения анти-де Ситтера.

Однородные и голые сингулярности $\Lambda M > 0$

При положительной космологической постоянной, $\Lambda > 0$, и положительной массе, $M > 0$, конформный множитель

$$\Phi = -\frac{\Lambda t^3 + 6M}{3t} < 0, \quad q := t \in (0, \infty), \quad (33.77)$$

отрицателен и не имеет нулей (горизонтов). В шварцшильдовских и черепаших координатах метрика имеет вид

$$ds^2 = \frac{3t}{\Lambda t^3 + 6M} dt^2 - \frac{\Lambda t^3 + 6M}{3t} d\sigma^2 - t^2 d\Omega_p, \quad t \in (0, \infty), \quad (33.78)$$

$$ds^2 = |\Phi| (d\tau^2 - d\sigma^2) - t^2 d\Omega_p, \quad -\frac{\pi}{12\sqrt{3}M} < \tau < \frac{\pi}{4\sqrt{3}M}, \quad (33.79)$$

где функция $t = t(\tau)$ определена уравнением (33.56), которое интегрируется и дает неявную связь $t = t(\tau)$ (при $q' > 0$):

$$\frac{\Lambda}{3}(\tau + \text{const}) = \frac{1}{6a} \ln \frac{a^2 - at + t^2}{(a + t)^2} + \frac{1}{a\sqrt{3}} \operatorname{arctanh} \left(\frac{2t - a}{a\sqrt{3}} \right), \quad a := 6M/\Lambda. \quad (33.80)$$

Соответствующее максимально продолженное решение представляется диаграммой Картера–Пенроуза в виде линзы типа IV, которая изображена на рис.33.4, и ее отражением во времени (диаграмма типа II). Оно описывает однородную пространственную сингулярность при $t = 0$ и является асимптотически пространством-временем де Ситтера при $t \rightarrow \infty$.

Глобальные решения, возникающие при $\Lambda < 0$, $M < 0$, получаются из рассмотренного выше решения преобразованием $\tau \leftrightarrow \sigma$. Соответствующие решения статичны. Диаграммы Картера–Пенроуза также имеют вид линзы и получаются из предыдущего случая поворотом на угол $\pi/2$.

Черное кольцо $\Lambda M < 0$

Если $\Lambda > 0$ и $M < 0$, то конформный множитель (33.77) имеет один простой нуль, и пространство-время содержит один горизонт. В этом случае имеется два статических и два однородных конформных блока, которые вместе склеиваются вдоль горизонта в диаграмму Картера–Пенроуза, изображенную на рис.33.4 справа. Она имеет тот же вид, что и в сферически симметричном случае $K = 1$, $\Lambda > 0$ и $M < 0$ (рис.33.2). Метрика для однородного решения имеет вид (33.78), но из-за условия (33.80) области определения координат t и τ меняются. Для статического решения метрика задается уравнением (33.71) с соответствующей областью определения координат r и σ , связанных тем же уравнением, что и t и τ .

Решения, возникающие при $\Lambda < 0$, $M > 0$, получаются из решений, соответствующих $\Lambda > 0$, $M < 0$, путем перестановки пространственной и временной координаты $\tau \leftrightarrow \sigma$. Соответствующую диаграмму Картера–Пенроуза необходимо повернуть на угол $\pi/2$. По виду она совпадает с черной дырой анти де Ситтера. В этом случае после компактификации $\mathbb{V} = \mathbb{T}^2$ горизонт черной дыры имеет вид тора $\mathbb{T}^2$, а пространство-время является асимптотически пространством анти-де Ситтера. Следовательно, в пространстве возникает кольцевая черная дыра (черное кольцо). Напомним, что в этом случае все пространство имеет вид топологического произведения прямой $\mathbb{R}$ и тора $\mathbb{T}^2$.

33.4.3 Гиперболические глобальные решения $K = -1$

При $K = -1$ поверхность $\mathbb{V}$ представляет собой двуполостный гиперboloид $\mathbb{H}^2$ (плоскость Лобачевского). Точнее, верхнюю полу двуполостного гиперboloида (см. раздел 29.2). После компактификации $\mathbb{H}^2$ в качестве поверхности $\mathbb{V}$ получится компактная риманова поверхность рода два и выше. Отметим, что группой изометрий однополостного гиперboloида $\mathbb{H}^2$ является группа преобразований Лоренца $\mathbb{O}(1, 2)$. То есть возникают решения вакуумных уравнений Эйнштейна симметричные относительно действия группы Лоренца $\mathbb{SO}(1, 2)$ не в пространстве-времени, а на пространственных сечениях $t = \text{const}$.

Метрика двуполостного гиперboloида (33.25) единичного радиуса в гиперболической системе координат имеет вид (29.16)

$$d\Omega_{\text{H}} = d\theta^2 + \text{sh}^2 \theta d\varphi^2. \quad (33.81)$$

Соответствующее вакуумное решение уравнений Эйнштейна можно записать в координатах Шварцшильда

$$ds^2 = \Phi(q)d\zeta^2 - \frac{dq^2}{\Phi(q)} - q^2 d\Omega_{\text{H}}, \quad (33.82)$$

где конформный множитель имеет вид

$$\Phi(q) = -1 - \frac{2M}{q} - \frac{\Lambda q^2}{3}. \quad (33.83)$$

Чтобы описать глобальные гиперболические решения, заметим, что конформный множитель Φ общего вида (33.49), меняет знак на противоположный при преобразовании всех постоянных:

$$K \mapsto -K, \quad \Lambda \mapsto -\Lambda, \quad M \mapsto -M.$$

Это означает, что все глобальные решения в случае $K = -1$, можно получить из сферически симметричных решений, если поменять пространственную и временную координату $\tau \leftrightarrow \sigma$ на поверхности $\mathbb{U}$, а также изменить знак космологической постоянной и массы.

Пример 33.4.1. Аналогом решения Шварцшильда является решение с нулевой космологической постоянной $\Lambda = 0$ и отрицательной массой $M < 0$. В этом случае метрика для однородного и статичного конформного блока принимает вид

$$ds^2 = \frac{dt^2}{1 + \frac{2M}{t}} - \left(1 + \frac{2M}{t}\right) d\sigma^2 - t^2 d\Omega_n, \quad -2M < t < \infty, \quad (33.84)$$

$$ds^2 = \left(1 + \frac{2M}{r}\right) d\tau^2 - \frac{dr^2}{1 + \frac{2M}{r}} - r^2 d\Omega_n, \quad 0 < r < -2M. \quad (33.85)$$

При этом диаграмму Картера–Пенроуза для решения Шварцшильда необходимо повернуть на угол $\pi/2$, рис.33.5. Соответствующее глобальное решение имеет левую и правую статические сингулярности при $q = r = 0$. Свойства этого глобального решения похожи на свойства двух статических сферически симметричных сингулярностей, рассмотренных в разделе 33.4.1, но с другим асимптотическим поведением. Пространство-время является асимптотически плоским в бесконечно удаленном прошлом и будущем (оба случая соответствуют пределу $q = t \rightarrow \infty$). $\square$

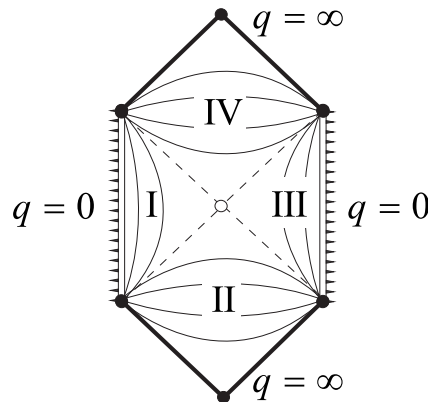


Рис. 33.5: Статические сингулярности в гиперболическом случае.

Остальные глобальные решения для $K = -1$ анализируются аналогично, и мы оставляем это любознательному читателю.

Не так легко представить себе пространство (трехмерное сечение четырехмерного пространства-времени, задаваемое уравнением $t = \text{const}$) при $K = -1$. В рассматриваемом случае оно представляет собой произведение интервала (конечного или

бесконечного в зависимости от значений Λ и M) и однополостного гиперboloида $\mathbb{H}^2$. После компактификации, когда поверхность $\mathbb{V}$ представляет собой замкнутую риманову поверхность рода два и выше, глобальное решение можно интерпретировать, как множество кротовых нор, число которых совпадает с родом поверхности (числом ручек). Топология горизонта совпадает с топологией поверхности $\mathbb{V}$.

33.5 Лоренц-инвариантные решения

В настоящем разделе мы рассмотрим случай С (33.21), когда второе дилатонное поле в сплетенном произведении (33.2) постоянно, $m = \text{const}$. В этом случае, как будет показано ниже, псевдориманова поверхность $\mathbb{U}$ должна быть поверхностью постоянной кривизны. Следовательно, она представляет собой однополостный гиперboloид, $\mathbb{U} = \mathbb{L}^2$ или его универсальную накрывающую, которые были подробно описаны в разделе 29.3. В этом случае глобальные решения вакуумных уравнений Эйнштейна имеют вид топологического произведения $\mathbb{M} = \mathbb{L}^2 \times \mathbb{V}$. Второй сомножитель $\mathbb{V}$ представляет собой максимально продолженную риманову поверхность с одним вектором Киллинга, которые были рассмотрены в разделе 31.3. Как было показано, поверхность $\mathbb{V}$ может иметь конические сингулярности или сингулярности кривизны вдоль края поверхности $\mathbb{V}$. С физической точки зрения этим сингулярностям соответствуют космические струны или сингулярные доменные стенки, которые эволюционируют во времени.

Случай С похож на пространственно симметричные решения, рассмотренные в случае В, однако имеет также несколько существенно новых черт. Во-первых, мы не можем ограничить себя только положительно определенными метриками $h_{\mu\nu}$ на $\mathbb{V}$, потому что уравнения Эйнштейна (33.18) неинвариантны относительно преобразования $h_{\mu\nu} \rightarrow -h_{\mu\nu}$ при заданном m . Отметим, что при $k = \text{const}$ преобразование $g_{\alpha\beta} \rightarrow -g_{\alpha\beta}$ всегда можно дополнить перестановкой пространственной и временной координаты на $\mathbb{U}$, $\tau \leftrightarrow \sigma$, что вместе оставляют уравнение (33.17) инвариантным. В случае евклидовой метрики на $\mathbb{V}$ это невозможно. Поэтому, не ограничивая общности, мы зафиксируем $m = 1$, но допустим, что метрика $h_{\mu\nu}$ может быть как положительно, так и отрицательно определена. В обоих случаях сигнатура четырехмерной метрики будет лоренцевой: либо $(+ - - -)$, либо $(+ - + +)$.

Решение уравнений (33.17)–(33.20) проводится так же, как и для метрики лоренцевой сигнатуры, при этом необходимо функцию m заменить на k и метрику $g_{\alpha\beta}$ на $h_{\mu\nu}$. Поэтому мы только кратко обозначим основные этапы вычислений, подчеркнув те моменты, которые специфичны для евклидовой сигнатуры.

При $m = 1$ полная система вакуумных уравнений Эйнштейна (33.17)–(33.20) принимает вид

$$\nabla_\mu \nabla_\nu k - \frac{\nabla_\mu k \nabla_\nu k}{2k} - \frac{1}{2} h_{\mu\nu} \left[\nabla^2 k - \frac{(\nabla k)^2}{2k} \right] = 0, \quad (33.86)$$

$$R^g + \nabla^2 k - 2k\Lambda = 0, \quad (33.87)$$

$$R^h + \frac{\nabla^2 k}{k} - \frac{(\nabla k)^2}{2k^2} - 2\Lambda = 0. \quad (33.88)$$

Как и в случае В, в уравнение (33.87) входит сумма функций от разных аргументов: $R^g = R^g(x)$ и $k = k(y)$. Поэтому скалярная кривизна поверхности $\mathbb{U}$ должна быть постоянной, $R^g = -2K = \text{const}$. Отсюда вытекает, что поверхность $\mathbb{U}$ является однополостным гиперboloидом $\mathbb{L}^2$ или его универсальной накрывающей.

Это – очень важное следствие вакуумных уравнений Эйнштейна, т.к. в рассматриваемом случае $\mathcal{C}$ все решения должны быть $\mathbb{SO}(1, 2)$ инвариантны, где группа преобразований Лоренца $\mathbb{SO}(1, 2)$ действует на однополостном гиперboloиде. Поэтому глобальные решения класса $\mathcal{C}$ названы лоренц-инвариантными.

Тогда уравнение (33.87) принимает вид

$$\nabla^2 k - 2(k\Lambda + K) = 0. \quad (33.89)$$

Как и в случае В уравнение (33.88) является следствием уравнений (33.86) и (33.89). Поэтому для нахождения решений вакуумных уравнений Эйнштейна достаточно решить уравнения (33.86) и (33.89).

Следующий шаг состоит в фиксировании координат на поверхности $\mathbb{V}$. Конформно евклидова метрика на поверхности $\mathbb{V}$ имеет вид

$$h_{\mu\nu} dy^\mu dy^\nu = \Phi dz d\bar{z} = \Phi(d\sigma^2 + d\rho^2). \quad (33.90)$$

Здесь $\Phi(z, \bar{z})$ является функцией комплексных координат

$$z := \sigma + i\rho, \quad \bar{z} = \sigma - i\rho, \quad (33.91)$$

где $\sigma = y^2$, $\rho = y^3$. При этом метрика всего четырехмерного пространства-времени равна

$$ds^2 = k d\Omega_L + \Phi dz d\bar{z}, \quad (33.92)$$

где $d\Omega_L$ – метрика постоянной кривизны на однополостном гиперboloиде $\mathbb{L}^2$, заданная, например, уравнением (33.24).

Не ограничивая общности, рассмотрим положительные $k > 0$. В противном случае можно просто переставить первые две координаты. Тогда удобно ввести параметризацию

$$k = q^2, \quad q > 0.$$

Для двух неизвестных функций q и Φ вместо уравнений (33.36)–(33.38) возникает следующая система уравнений

$$\partial_{zz}^2 q - \frac{\partial_z \Phi \partial_z q}{\Phi} = 0, \quad (33.93)$$

$$\partial_{\bar{z}\bar{z}}^2 q - \frac{\partial_{\bar{z}} \Phi \partial_{\bar{z}} k}{\Phi} = 0, \quad (33.94)$$

$$2 \frac{\partial_z \partial_{\bar{z}} q^2}{\Phi} - (K + \Lambda q^2) = 0. \quad (33.95)$$

Аналогично случаю В, решением уравнений (33.93) и (33.94) являются функции одного аргумента: $q = q(z \pm \bar{z})$ и $\Phi = \Phi(z \pm \bar{z})$, при этом функция Φ определяется уравнением

$$|\Phi| = |q'|, \quad (33.96)$$

где штрих обозначает дифференцирование по соответствующему аргументу. В полученной формуле нижний и верхний знаки соответствуют положительно и отрицательно определенной римановой метрике на $\mathbb{V}$. Таким образом, функции q и Φ зависят либо от координаты σ , либо от $i\rho$. Поскольку, благодаря вращательной $\mathbb{SO}(2)$ симметрии конформно евклидовой метрики (33.90), оба выбора равнозначны, то для определенности мы предположим, что функции $q(\sigma)$ и $\Phi(\sigma)$ зависят от σ .

После этого уравнение (33.95) упростится:

$$\frac{1}{2}(q^2)'' = (K + \Lambda q^2)\Phi, \quad (33.97)$$

где штрих обозначает дифференцирование по σ . Чтобы его проинтегрировать, в уравнении (33.96) необходимо раскрыть знаки модулей.

Рассмотрим случай $\Phi q' > 0$. Тогда уравнение (33.97) с учетом (33.96) примет вид

$$\frac{1}{2}(q^2)'' = (K + \Lambda q^2)q',$$

и его легко проинтегрировать:

$$q' = K - \frac{2M}{q} + \frac{\Lambda q^2}{3}, \quad (33.98)$$

где M – произвольная постоянная интегрирования. Хотя в рассматриваемом случае ее нельзя интерпретировать как массу, мы будем использовать старые обозначения, чтобы облегчить сравнение. Учитывая уравнение (33.96) получаем выражение для конформного множителя

$$\Phi(q) = K - \frac{2M}{q} + \frac{\Lambda q^2}{3} \quad (33.99)$$

Случай $\Phi q' < 0$ интегрируется аналогично.

Окончательно, общее решение вакуумных уравнений Эйнштейна в случае С имеет вид

$$ds^2 = q^2 d\Omega_L + \Phi(q)(d\sigma^2 + d\rho^2), \quad (33.100)$$

где конформный множитель задан уравнением (33.99) и функция $q = q(\sigma)$ определяется уравнением (33.96). Таким образом, метрика на поверхности $\mathbb{V}$ имеет один вектор Киллинга ∂_ρ и имеет тот же вид, что и в разделе 31.1. Поэтому мы можем использовать развитую там технику для построения глобальных решений вакуумных уравнений Эйнштейна.

Выбирая функцию $q(\sigma)$ в качестве одной из координат, метрику (33.100) можно записать в виде, напоминающем метрику Шварцшильда,

$$ds^2 = q^2 d\Omega_L + \frac{dq^2}{\Phi(q)} + \Phi(q) d\sigma^2. \quad (33.101)$$

Результирующая метрика имеет три вектора Киллинга, соответствующих группе симметрии $\mathbb{SO}(1, 2)$ однополостного гиперболоида постоянной кривизны $\mathbb{L}^2$, и один дополнительный вектор Киллинга ∂_ρ на поверхности $\mathbb{V}$.

Вычисления, аналогичные случаю $k = 1$, приводят к следующему выражению для скалярной кривизны поверхности $\mathbb{V}$

$$R^h = \frac{2}{3}\Lambda + \frac{4M}{q^3}.$$

При этом для инвариантного собственного значения тензора Вейля мы получаем то же выражение (33.59), что и в случае В.

33.5.1 Лоренц-инвариантные решения $K = 1$

Прежде всего отметим, что случаи $K = 1$ и $K = -1$ связаны между собой перестановкой первых двух координат $\tau \leftrightarrow \sigma$. Мы выберем значение $K = 1$, чтобы выражение для конформного множителя Φ имело, с точностью до изменения знака космологической постоянной, тот же вид, что и для сферически симметричного случая. Метрика для однополостного гиперболоида в гиперболической полярной системе координат имеет вид (29.28). Поэтому четырехмерная метрика пространства-времени в координатах Шварцшильда запишется следующим образом

$$ds^2 = q^2(d\theta^2 - \text{ch}^2\theta d\varphi^2) + \frac{dq^2}{\Phi(q)} + \Phi(q)d\rho^2, \quad (33.102)$$

где конформный множитель,

$$\Phi = 1 - \frac{2M}{q} + \frac{\Lambda q^2}{3}, \quad (33.103)$$

имеет тот же вид, что и в решении Коттлера [255], но в рассматриваемом случае этот нетривиальный конформный множитель входит в евклидову часть метрики.

На поверхности $\mathbb{V}$ метрика может быть как отрицательно ($\Phi < 0$), так и положительно ($\Phi > 0$) определена. Для отрицательно определенной метрики сигнатура метрики пространства-времени равна $(+ - - -)$, и роль времени играет координата θ . Поэтому времениподобная координата принимает значения на всей вещественной оси $\theta \in \mathbb{R}$, и трехмерное пространство представляет собой произведение окружности $\varphi \in [0, 2\pi)$ и поверхности $\mathbb{V}$, которая будет построена ниже. Если в качестве $\mathbb{U}$ выбрать универсальную накрывающую однополостного гиперболоида $\mathbb{L}^2$, то трехмерное пространство будет представлять произведение $\mathbb{R} \times \mathbb{V}$. Эволюция этих пространств во времени длится вечно, и если поверхность $\mathbb{V}$ имеет сингулярность, то ей будет соответствовать времениподобная кривая.

Для положительно определенной метрики на $\mathbb{V}$ сигнатура четырехмерной метрики равна $(+ - ++)$, и времениподобной координатой является угол φ . При $\mathbb{U} = \mathbb{L}^2$ он принимает значения на окружности $\varphi \in [0, 2\pi)$, и трехмерное пространство представляет собой произведение прямой $\theta \in \mathbb{R}$ и поверхности $\mathbb{V}$. Соответствующее пространство-время содержит замкнутые времениподобные кривые (включая экстремали), если только в качестве поверхности $\mathbb{U}$ не выбрано универсальное накрывающее пространство для $\mathbb{L}^2$.

Поскольку метрика на поверхности $\mathbb{V}$ имеет тот же локальный вид, что и в разделе 31.1, то мы можем построить максимально продолженные поверхности $\mathbb{V}$. Вид римановой поверхности определяется конформным множителем, показанным на рис.33.1 при $\Lambda \neq 0$. Надо только помнить, что в выражение для конформного множителя космологической постоянной входит с противоположным знаком. Кроме того, при отождествлении $\rho \sim \rho + L$ возможно появление конических сингулярностей, которые соответствуют космическим струнам.

Перейдем к классификации глобальных решений вакуумных уравнений Эйнштейна в рассматриваемом случае $\mathcal{C}$.

Пространство-время Минковского $\Lambda = 0$, $M = 0$

При нулевых значениях постоянных поверхность $\mathbb{V}$ представляет собой полуплоскость $q \in (0, \infty)$, $\rho \in (-\infty, \infty)$. Она неполна при $q = 0$ из-за координатной сингулярности, аналогичной сферически симметричному случаю.

Космическая струна $\Lambda = 0$, $M > 0$

При нулевой космологической постоянной $\Lambda = 0$ и положительных значениях $M > 0$ возникает решение, соответствующее решению Шварцшильда. В этом случае существуют две несвязные между собой максимально продолженные римановы поверхности, изображенные на рис.33.6 слева и в центре с положительно и отрицательно определенной метрикой. Для наглядности мы отождествили точки с координатами ρ и $\rho + L$.

Внешнему решению Шварцшильда соответствует интервал $q \in (2M, \infty)$. В этой области конформный множитель положителен и четырехмерная метрика имеет сигнатуру $(+ - + +)$. Соответствующая поверхность $\mathbb{V}$ является гладким многообразием. При этом возможно появление конической сингулярности на горизонте $q = 2M$. С топологической точки зрения коническая сингулярность представляет собой плоскость (мировую поверхность космической струны), а все пространство-время описывает бесконечную эволюцию бесконечной космической струны. Для этого решения кривизна не имеет больше никаких сингулярностей, и если выполнено уравнение (31.30), то пространство-время вообще не имеет сингулярностей. Специфическим свойством этого решения является то, что периметр цилиндра, изображенного на рис.33.6 слева, стремится к постоянной при $q \rightarrow \infty$. Это свойство соответствует тому, что решение Шварцшильда является асимптотически плоским.

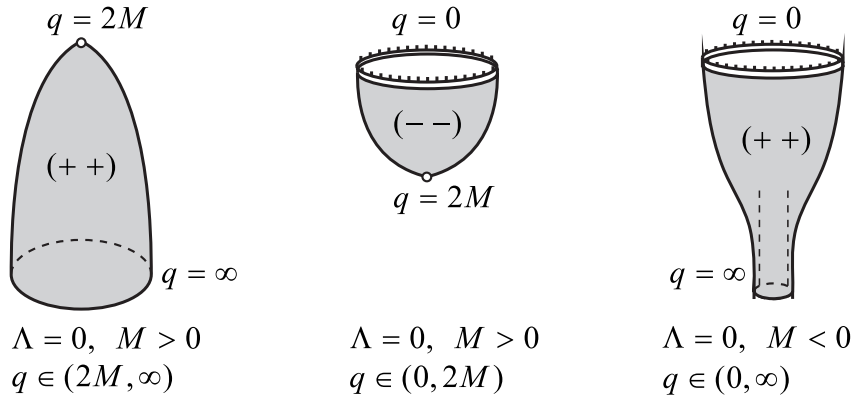


Рис. 33.6: Максимально продолженная риманова поверхность $\mathbb{V}$ при нулевой космологической постоянной $\Lambda = 0$. Условные обозначения те же, что и на рис. 30.3. Незакрашенные окружности обозначают возможные конические сингулярности.

Для внутреннего решения Шварцшильда $q \in (0, 2M)$, конформный множитель отрицателен, и метрика пространства-времени имеет сигнатуру $(+ - - -)$. Максимально продолженная поверхность $\mathbb{V}$ представляет собой в этом случае диск конечного радиуса с сингулярным краем $q = 0$, который изображен на рис.33.6 в центре. Если выполнено уравнение (31.30), то поверхность $\mathbb{V}$ гладкая и имеет особенность только на крае. В противном случае в центре диска $q = 2M$ возникает коническая сингулярность, соответствующая космической струне. Поэтому трехмерное пространство представляет собой космическую струну, окруженную цилиндрической доменной стенкой, на которой кривизна сингулярна. Сингулярный край невозможно адекватно изобразить на рисунке, поскольку он находится на конечном расстоянии, но в то же время имеет бесконечный периметр.

Доменная стенка $\Lambda = 0$, $M < 0$

Если космологическая постоянная равна нулю, а масса отрицательна, то решение Шварцшильда описывает голую сингулярность. В этом случае конформный множитель при $q \in (0, \infty)$ положителен и сигнатура метрики равна $(+ - + +)$. Максимально продолженная риманова поверхность $\mathbb{V}$ представляет собой полуплоскость, $q \in (0, \infty)$, $\rho \in \mathbb{R}$. Сингулярность кривизны расположена вдоль края $q = 0$, что можно интерпретировать, как доменную стенку, на которой кривизна сингулярна. С четырехмерной точки зрения это решение описывает эволюцию доменной стенки, которая расположена на конечном расстоянии.

Благодаря трансляционной симметрии, точки с координатами ρ и $\rho + L$ можно отождествить. Возникающая при этом поверхность изображена на рис.33.6 справа. Соответствующее трехмерное пространство представляет собой цилиндрическую доменную стенку.

Решение де Ситтера $\Lambda > 0$, $M = 0$

При нулевой массе и положительной космологической постоянной конформный множитель (33.99) положителен для $q \in (0, \infty)$. Поверхность $\mathbb{V}$ имеет метрику с сигнатурой $(+ - + +)$ и показана на рис. 33.7 слева. Для наглядности она изображена после отождествления точек с координатами ρ и $\rho + L$. Пространство де Ситтера соответствует универсальной накрывающей поверхности $\mathbb{V}$. Поверхность $\mathbb{V}$ является гладкой и не имеет сингулярностей. Однако она неполна при $q = 0$. Ее продолжение через край $q = 0$ невозможно, т.к. четырехмерная метрика пространства-времени вырождается.

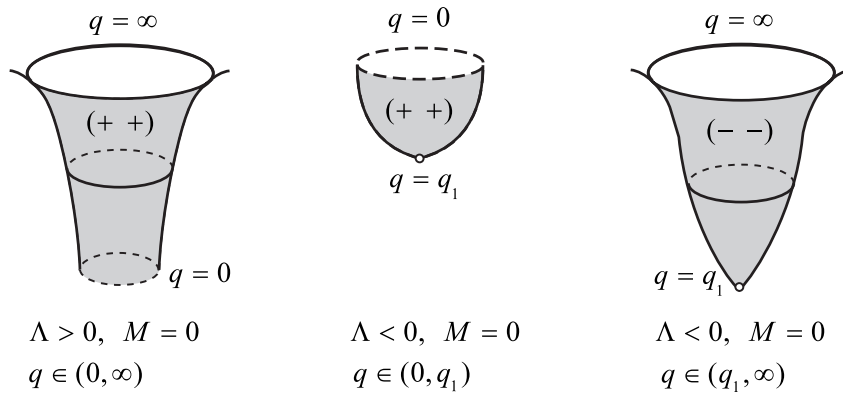


Рис. 33.7: Максимально продолженные римановы поверхности $\mathbb{V}$ при нулевой массе $M = 0$. Условные обозначения те же, что и на рис. 30.3. Незакрашенные окружности обозначают возможные конические сингулярности.

Решение анти-де Ситтера $\Lambda < 0$, $M = 0$

При отрицательной космологической постоянной $\Lambda < 0$ и нулевой константе $M = 0$ возникает решение анти-де Ситтера. Конформный множитель при этом имеет один простой нуль (горизонт) в точке $q_1 = \sqrt{-3/\Lambda}$. Следовательно, существуют две несвязные между собой поверхности $\mathbb{V}$: одна с положительно, а другая с отрицательно определенной метрикой.

При $q \in (0, q_1)$ конформный множитель положителен, метрика пространства-времени имеет сигнатуру $(+ - ++)$, и поверхность $\mathbb{V}$ показана на рис.33.7 в центре. С топологической точки зрения эта поверхность представляет собой диск конечного радиуса с положительно определенной метрикой. В центре диска $q = q_1$ возможно появление конической сингулярности. Ее продолжение через край $q = 0$ невозможно, т.к. четырехмерная метрика вырождается.

Если $q \in (q_1, \infty)$, то конформный множитель отрицателен и метрика пространства-времени имеет сигнатуру $(+ - --)$. Соответствующая поверхность $\mathbb{V}$ показана на рис.33.7 справа. С топологической точки зрения она представляет собой плоскость, в центре которой возможно появление конической сингулярности.

Доменная стенка $\Lambda < 0$, $M > \frac{1}{3\sqrt{-\Lambda}}$

При отрицательной космологической постоянной $\Lambda < 0$ и $M > \frac{1}{3\sqrt{-\Lambda}}$ конформный множитель отрицателен при $q \in (0, \infty)$, не имеет нулей, и сигнатура метрики пространства-времени равна $(+ - --)$. С топологической точки зрения поверхность $\mathbb{V}$ представляет собой полуплоскость $q > 0$, $\rho \in \mathbb{R}$ с отрицательно определенной метрикой и сингулярной кривизной на границе $q = 0$. Глобальное решение описывает бесконечную эволюцию бесконечной плоской доменной стенки сингулярной кривизны.

Компактификация по координате $\rho \sim \rho + L$ дает внутренность цилиндрической доменной стенки, показанной на рис.33.8. Длина направляющей окружности цилиндра стремится к бесконечности при $q \rightarrow 0$ и $q \rightarrow \infty$. Сингулярный край $q = 0$ находится на конечном расстоянии.

Двойной горизонт $\Lambda < 0$, $M = \frac{1}{3\sqrt{-\Lambda}}$

При указанных значениях постоянных конформный множитель отрицателен и имеет двойной нуль в точке q_1 . Метрика пространства-времени в обоих интервалах $q \in (0, q_1)$ и $q \in (q_1, \infty)$ имеет сигнатуру $(+ - --)$. В этом случае имеем две поверхности, изображенные на рис.33.8. Поверхность $\mathbb{V}$, соответствующая интервалу $q \in (q_1, \infty)$, не является односвязной. Ее можно представить себе в виде плоскости, центр которой $q = q_1$ удален в бесконечность, т.к. он лежит на бесконечном расстоянии.

Поверхность для интервала $(0, q_1)$ представляет собой цилиндрическую доменную стенку сингулярностей кривизны. Сингулярный край $q = 0$ лежит на конечном расстоянии.

Два горизонта $\Lambda < 0$, $0 < M < \frac{1}{3\sqrt{-\Lambda}}$

Если космологическая постоянная отрицательна $\Lambda < 0$ и масса лежит в интервале $0 < M < \frac{1}{3\sqrt{-\Lambda}}$, то конформный множитель имеет максимальное число нулей: два простых нуля, расположенных в точках $q_{1,2}$ (см. рис.33.1 слева). В интервалах $(0, q_1)$ и (q_2, ∞) конформный множитель отрицателен, и сигнатура метрики пространства-времени равна $(+ - --)$. В интервале (q_1, q_2) конформный множитель положителен, и сигнатура метрики имеет вид $(+ - ++)$. Максимально продолженная поверхность $\mathbb{V}$ является одной из трех поверхностей, показанных на рис. 33.8 в нижнем ряду.

При $q \in (0, q_1)$ поверхность $\mathbb{V}$ представляет собой диск конечного радиуса, на крае которого кривизна сингулярна. В центре диска, возможно, расположена коническая сингулярность. Это зависит от отождествления $\rho \sim \rho + L$.

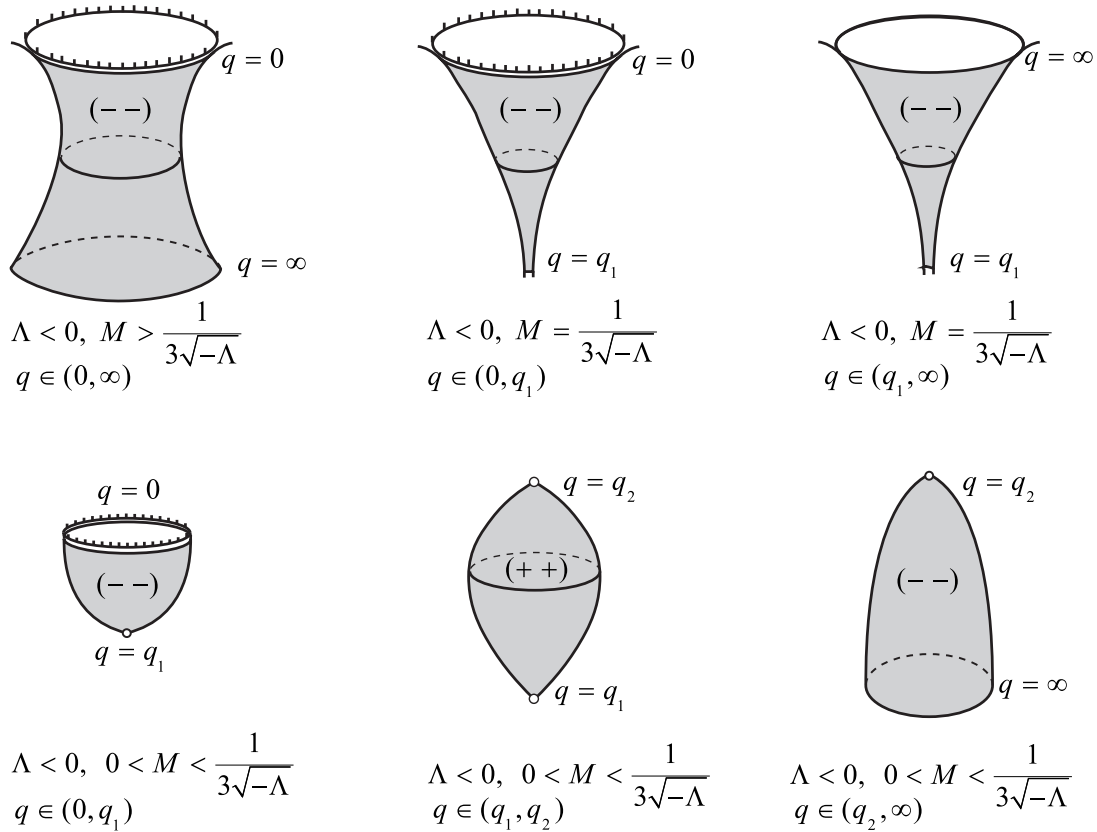


Рис. 33.8: Максимально продолженные римановы поверхности $\mathbb{V}$ при отрицательной космологической постоянной $\Lambda < 0$ и $M \neq 0$. Незакрашенные окружности обозначают возможные конические сингулярности.

Если $q \in (q_1, q_2)$, то поверхность $\mathbb{V}$ топологически представляет собой сферу, в общем случае имеющую две конические сингулярности. Алгебраический анализ системы уравнений $\Phi(q_1) = \Phi(q_2) = 0$ и $\Phi'(q_1) = -\Phi'(q_2)$ при $q_1 \neq q_2$ показывает, что таких решений не существует. Поэтому, подгоняя период компактификации L , можно устранить одну из конических сингулярностей, но не обе одновременно. Это значит, что в рассматриваемом случае всегда должна существовать космическая струна. С топологической точки зрения трехмерное пространство представляет собой произведение сферы, содержащей одну или две конические сингулярности, и прямой.

При $q \in (q_2, \infty)$ поверхность $\mathbb{V}$ – это плоскость, в центре которой, возможно, расположена коническая сингулярность. Экстремали при $q \rightarrow \infty$ полны, а кривизна стремится к постоянной.

Космическая струна $\Lambda < 0, M < 0$

Если и космологическая постоянная $\Lambda < 0$, и постоянная $M < 0$, то конформный множитель Φ имеет один простой нуль в точке q_1 . В этом случае возможно существование двух поверхностей $\mathbb{V}$.

При $q \in (0, q_1)$ конформный множитель положителен, и сигнатура метрики пространства-времени равна $(+ - + +)$. Топологически поверхность $\mathbb{V}$ имеет тот же вид, что и рассмотренная в предыдущем случае поверхность для $\Lambda < 0, 0 < M < \frac{1}{3\sqrt{-\Lambda}}$ и $q \in (0, q_1)$, но с отрицательно определенной метрикой.

Вторая поверхность $\mathbb{V}$ соответствует интервалу $q \in (q_1, \infty)$. Конформный множи-

тель отрицателен, и сигнатура метрики пространства-времени равна $(+ - - -)$. Поверхность $\mathbb{V}$ имеет тот же вид, что и в случае $\Lambda < 0$, $M = 0$ и $q \in (q_1, \infty)$ на рис.33.7. Глобальное решение вакуумных уравнений Эйнштейна в общем случае описывает бесконечную космическую струну, причем других сингулярностей у кривизны нет.

Космическая струна $\Lambda > 0$, $M > 0$

При положительной космологической постоянной $\Lambda > 0$ и массе $M > 0$. Конформный множитель имеет один простой нуль в точке q_1 , как видно из рис.33.1, справа. Он отрицателен в интервале $q \in (0, q_1)$ и положителен при $q \in (q_1, \infty)$. Этот случай аналогичен предыдущему. Мы имеем две максимально продолженные поверхности $\mathbb{V}$, на которых необходимо просто изменить сигнатуру метрики.

Доменная стенка $\Lambda > 0$, $M < 0$

Если космологическая постоянная положительна, $\Lambda > 0$, и постоянная M отрицательна, то конформный множитель положителен при $q \in (0, \infty)$. Сигнатура метрики пространства-времени равна $(+ - ++)$. В этом случае имеется одна поверхность $\mathbb{V}$, т.к. нули у конформного множителя отсутствуют. Топологически поверхность $\mathbb{V}$ такая же, как и в случае $\Lambda < 0$, $M > 1/3\sqrt{-\Lambda}$, но с положительно определенной метрикой.

33.5.2 Решения с плоскостью Минковского $K = 0$

При $K = 0$ поверхность $\mathbb{U}$ представляет собой или плоскость Минковского $\mathbb{R}^{1,1}$, или цилиндр, или тор. При этом возникают новые решения, интересные с топологической точки зрения. Соответствующая метрика в координатах Шварцшильда имеет вид

$$ds^2 = q^2(dt^2 - dx^2) + \frac{dq^2}{\Phi(q)} + \Phi(q)d\rho^2, \quad (33.104)$$

где

$$\Phi(q) = -\frac{2M}{q} + \frac{\Lambda q^2}{3}.$$

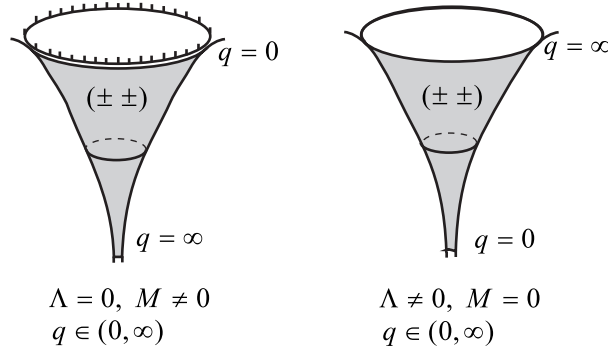
В зависимости от значения постоянных Λ и M , входящих в конформный множитель, возможны четыре качественно отличных случая.

Доменная стенка $\Lambda = 0$, $M \neq 0$

При нулевой космологической постоянной $\Lambda = 0$, но отличной от нуля постоянной $M \neq 0$ максимально продолженная поверхность $\mathbb{V}$ изображена на рис. 33.9. Метрика на этой поверхности положительно и отрицательно определена, соответственно, при $M < 0$ и $M > 0$.

Решение (анти-)де Ситтера $\Lambda \neq 0$, $M = 0$

Если масса равна нулю, то возникает еще одно представление решения (анти-)де Ситтера для положительной и отрицательной космологической постоянной. Соответствующая поверхность $\mathbb{V}$ показана на рис. 33.9.

Рис. 33.9: Глобальные решения с плоскостью Минковского, $K = 0$.

Пусть $\Phi q' > 0$ или, что эквивалентно, $\Lambda q' > 0$. Тогда уравнение (31.2) принимает вид

$$q' = \frac{\Lambda q^2}{3}$$

и легко интегрируется:

$$q = -\frac{3}{\Lambda \sigma},$$

где мы опустили несущественную постоянную интегрирования, соответствующую сдвигу σ . Тогда метрика в черепаших координатах равна

$$ds^2 = \frac{9}{\Lambda^2 \sigma^2} (dt^2 - dx^2) + \frac{3}{\Lambda \sigma^2} (d\sigma^2 + d\rho^2). \quad (33.105)$$

Растяжением координат t и x она приводится к конформно плоскому виду.

Космическая струна $\Lambda > 0$, $M > 0$

Конформный множитель Φ при $\Lambda > 0$ и $M > 0$ имеет один простой нуль в точке q_1 . При $q \in (0, q_1)$ и $q \in (q_1, \infty)$, конформный множитель, соответственно, отрицателен и положителен. Поэтому метрики пространства-времени имеют сигнатуры $(+ - - -)$ и $(+ - ++)$. Максимально продолженные поверхности $\mathbb{V}$ топологически такие же как и для конформного множителя $\Phi = 1 - \frac{2M}{q} + \frac{\Lambda q^2}{3}$, который был рассмотрен ранее. Меняется только величина q_1 . Глобальные решения описывают космическую струну и космическую струну, окруженную доменной стенкой сингулярностей кривизны.

Изменение знака обеих постоянных $\Lambda < 0$, $M > 0$ приводит к изменению сигнатуры метрики. Качественные свойства решений остаются прежними.

Доменная стенка $\Lambda > 0$, $M < 0$

В рассматриваемом случае конформный множитель Φ положителен и не имеет нулей. Сигнатура метрики пространства-времени равна $(+ - ++)$. Существует всего одна максимально продолженная поверхность $\mathbb{V}$. Топологически она такая же, как и для конформного множителя $\Phi = 1 - \frac{2M}{q} + \frac{\Lambda q^2}{3}$, рассмотренного ранее.

При $\Lambda < 0$ и $M > 0$ необходимо просто изменить сигнатуру всей метрики.

33.6 Итоги главы

В двух последних главах были найдены и классифицированы все глобальные вакуумные решения уравнений Эйнштейна с космологической постоянной, которые

имеют вид сплетенного произведения двух поверхностей. Явное построение и классификация решений проведена в зависимости от значений постоянной скалярной кривизны одной из поверхностей, значения космологической постоянной Λ и единственной постоянной интегрирования M , которая для решения Шварцшильда имеет физический смысл массы черной дыры. Мы видим, что требование максимального продолжения решений практически однозначно определяет глобальную структуру пространства-времени. Важно отметить, что при решении уравнений движения мы не ставим никаких граничных условий. Подчеркнем, что решение уравнений Эйнштейна в какой-то фиксированной системе координат само по себе дает не так уж много. Для того, чтобы дать физическую интерпретацию решений необходимо знать глобальную структуру пространства-времени. Эта задача сложна, но обойти ее, по-видимому, невозможно.

Предположение о том, что метрика пространства-времени имеет вид сплетенного произведения метрик двух поверхностей влечет за собой симметрию метрики, если потребовать выполнения вакуумных уравнений Эйнштейна. Например, для решения Шварцшильда мы не требовали сферической симметрии метрики – она возникла в процессе решения уравнений Эйнштейна.

Построенные решения представляют значительный физический интерес. Мы показали, что вакуумные решения уравнений Эйнштейна описывают черные дыры, космические струны, кротовые норы, доменные стенки сингулярностей кривизны и другие. В настоящей монографии мы лишь кратко обсудили свойства построенных глобальных решений.

Глава 34

Решение Шварцшильда

В настоящем разделе мы детально рассмотрим решение Шварцшильда, т.к. оно играет важную роль в физических приложениях. Фактически, основные наблюдательные данные, подтверждающие общую теорию относительности, основаны на предсказаниях, полученных для этого решения.

Вскоре после создания общей теории относительности К. Шварцшильд нашел точное статическое сферически симметричное решение вакуумных уравнений Эйнштейна [218]. Это же решение было также независимо найдено Дж. Дросте [219, 220, 221], который был учеником Г. Лоренца. Позже было доказано, что любое сферически симметричное решение вакуумных уравнений Эйнштейна вне горизонта имеет дополнительный времениподобный вектор Киллинга, т.е. существует такая система координат, что вне горизонта решение является статическим. Это утверждение означает, что условие статичности при получении решения Шварцшильда является излишним и известно, как теорема Бирхгоффа [253].

Решение Шварцшильда уже было найдено нами в разделе 33.4.1, где также было построено его максимальное продолжение вдоль экстремалей. В том числе было доказано существование дополнительного вектора Киллинга, который вне горизонта времениподобен, а под горизонтом пространственноподобен. Глобальное (максимально продолженное вдоль экстремалей) решение Шварцшильда описывает черную дыру – новый объект, предсказанный общей теорией относительности, который отсутствует в теории гравитации Ньютона. В настоящее время астрономические наблюдения говорят о возможном существовании черных дыр во вселенной, в частности, в центре нашей галактики. Метрика для решения Шварцшильда вне горизонта успешно используется для описания гравитационного поля звезд, а также солнечной системы.

34.1 Координаты Шварцшильда

Метрика Шварцшильда в координатах Шварцшильда $\{x^\alpha\} = (x^0, x^1, x^2, x^3) = (t, \rho, \theta, \varphi)$, имеет хорошо известный вид

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) dt^2 - \frac{d\rho^2}{1 - \frac{2M}{\rho}} - \rho^2(d\theta^2 + \sin^2\theta d\varphi^2), \quad (34.1)$$

где $M = \text{const}$ – интеграл движения, который имеет физический смысл массы, сосредоточенной в некоторой области, вне которой решаются вакуумные уравнения Эйнштейна. Это – сферически симметричное решение вакуумных уравнений Эйнштейна с нулевой космологической постоянной, которое было получено в разделе

33.4.1. Последнее слагаемое представляет собой метрику сферы радиуса ρ , которая индуцирована вложением $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ и взята с обратным знаком. Группа вращений $\mathbb{O}(3)$ действует обычным образом на азимутальный и полярный углы θ и φ . Метрика (34.1) определена при

$$-\infty < t < \infty, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi.$$

При $\rho = \rho_s := 2M$ компоненты метрики вырождаются. Радиус ρ_s называется *радиусом Шварцшильда* или *гравитационным радиусом*. Ему соответствует сфера радиуса ρ_s , которая называется *горизонтом*. При этом координата ρ , которая предполагается положительной, меняется либо в пределах $2M < \rho < \infty$ (вне горизонта ρ_s), либо $0 < \rho < 2M$ (под горизонтом).

Метрика (34.1) вырождена при $\rho = 0$, $2M$ и $\theta = 0, \pi$. Вырождение метрики при $\theta = 0, \pi$ связано с выбором координат на сфере $\mathbb{S}^2$. Это – хорошо известная особенность сферической системы координат, которую мы обсуждать не будем. Особенности метрики при $\rho = 0$, $2M$ мы обсудим подробно.

Несмотря на то, что компоненты метрики на сфере Шварцшильда вырождаются, ее определитель не имеет нулей при $\rho > 0$:

$$\det g_{\alpha\beta} = -\rho^4 \sin^2 \theta,$$

за исключением особенности сферической системы координат при $\theta = 0, \pi$.

Координата ρ при $\rho > 2M$ является аналогом радиальной координаты в сферической системе координат трехмерного евклидова пространства $\mathbb{R}^3$. Координата t при $\rho > 2M$ играет роль времени. Вне горизонта метрика Шварцшильда в координатах Шварцшильда является статической. При $0 < \rho < 2M$ смысл координат t и r меняется: t становится пространственноподобной координатой, а ρ – времениподобной. Поэтому координата ρ под горизонтом никакого отношения к радиусу не имеет.

Если на многообразии задана какая-либо метрика, то первое, что хочется сделать, – это вычислить компоненты соответствующего тензора кривизны и его инварианты. Поскольку метрика Шварцшильда является решением вакуумных уравнений Эйнштейна с нулевой космологической постоянной, то тензор Риччи и скалярная кривизна пространства-времени равны нулю. Однако полный тензор кривизны отличен от нуля. Поскольку тензор Риччи и скалярная кривизна равны нулю, то тензор кривизны в рассматриваемом случае совпадает с тензором Вейля. Перейдем к вычислениям.

Метрика (34.1) задает в пространстве-времени символы Кристоффеля (6.24), из которых выпишем только линейно независимые и нетривиальные:

$$\begin{aligned} \Gamma_{00}^1 &= \frac{M}{\rho^2} \left(1 - \frac{2M}{\rho} \right), & \Gamma_{22}^1 &= -\rho \left(1 - \frac{2M}{\rho} \right), \\ \Gamma_{01}^0 &= \frac{M}{\rho^2 \left(1 - \frac{2M}{\rho} \right)}, & \Gamma_{33}^1 &= -\rho \left(1 - \frac{2M}{\rho} \right) \sin^2 \theta, \\ \Gamma_{11}^1 &= -\frac{M}{\rho^2 \left(1 - \frac{2M}{\rho} \right)}, & \Gamma_{23}^3 &= \frac{\cos \theta}{\sin \theta}, \\ \Gamma_{12}^2 &= \Gamma_{13}^3 = \frac{1}{\rho}, & \Gamma_{33}^2 &= -\sin \theta \cos \theta. \end{aligned} \tag{34.2}$$

“Линейно независимые” означает, что мы не выписали нетривиальные компоненты, которые получаются из приведенных выше простой перестановкой индексов. Например, $\Gamma_{10}^0 = \Gamma_{01}^0$.

Прямые довольно громоздкие вычисления приводят к следующим нетривиальным линейно независимым компонентам тензора кривизны (6.74) для метрики (34.1):

$$\begin{aligned} R_{0101} &= -\frac{2M}{\rho}, & R_{1212} &= -\frac{M}{\rho \left(1 - \frac{2M}{\rho}\right)}, \\ R_{0202} &= \frac{M}{\rho} \left(1 - \frac{2M}{\rho}\right), & R_{1313} &= -\frac{M}{\rho \left(1 - \frac{2M}{\rho}\right)} \sin^2 \theta, \\ R_{0303} &= \frac{M}{\rho} \left(1 - \frac{2M}{\rho}\right) \sin^2 \theta, & R_{2323} &= 2M\rho \sin^2 \theta. \end{aligned} \quad (34.3)$$

Компоненты тензора кривизны можно либо вычислить в лоб, воспользовавшись формулой (6.74), либо воспользоваться формулами из раздела 33.1 для сплетенного произведения двух многообразий, т.к. метрика Шварцшильда (34.1) имеет именно такой вид. В обозначениях этого раздела следует положить

$$\begin{aligned} k &= 1, & m &= -\rho^2, \\ g_{\alpha\beta} &= \begin{pmatrix} 1 - \frac{2M}{\rho} & 0 \\ 0 & -\frac{1}{1 - \frac{2M}{\rho}} \end{pmatrix}, & h_{\mu\nu} &= \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix}. \end{aligned}$$

В последнем случае вычисления несколько упрощаются.

Мы видим, что компоненты тензора кривизны вырождены в нуле $\rho = 0$ и на горизонте $\rho = \rho_s$. Само по себе это только указывает на возможные особенности геометрических объектов (в рассматриваемом случае метрики), т.к. компоненты тензора не являются геометрическими инвариантами. Чтобы прояснить ситуацию, вычислим квадрат тензора кривизны, который инвариантен относительно преобразований координат:

$$\begin{aligned} R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} &= 4 \left[R_{0101}^2 (g^{00} g^{11})^2 + R_{0202}^2 (g^{00} g^{22})^2 + R_{0303}^2 (g^{00} g^{33})^2 + \right. \\ &\quad \left. + R_{1212}^2 (g^{11} g^{22})^2 + R_{1313}^2 (g^{11} g^{33})^2 + R_{2323}^2 (g^{22} g^{33})^2 \right] = \frac{48M^2}{\rho^6}. \end{aligned} \quad (34.4)$$

Хотя тензор Риччи и скалярная кривизна равны нулю, геометрия нетривиальна, т.к. квадрат тензора кривизны, который в данном случае совпадает с квадратом тензора Вейля, отличен от нуля. Он определен при всех $\rho > 0$ и имеет только одну особенность при $\rho = 0$.

При $\rho \rightarrow \infty$ кривизна пространства-времени стремится к нулю. Это значит, что пространство-время является асимптотически плоским на больших расстояниях. Этот факт следовал также из вида метрики, поскольку метрика (34.1) стремится к метрике Лоренца при $\rho \rightarrow 0$.

На горизонте, $\rho = \rho_s$, кривизна никакой особенности не имеет. Это указывает на то обстоятельство, что особенность метрики (34.1) на горизонте может быть связана с неудачным выбором координат, т.е. является координатной. Действительно, ниже мы увидим, что в координатах Крускала–Секереша метрика Шварцшильда на горизонте регулярна. Конечно, преобразование координат Шварцшильда к координатам Крускала–Секереша вырождено при $\rho = \rho_s$.

Особенность при $\rho = 0$ является неустранимой. Подчеркнем, что значение $\rho = 0$ не соответствует началу сферической системы координат, т.к. координата ρ под горизонтом является времениподобной. Ниже мы увидим, что пространственноподобная особенность $\rho = 0$ соответствует черной и белой дырам.

Метрика (34.1) сферически симметрична, при этом группа вращений $\mathbb{O}(3)$ действует только на угловые координаты θ и φ . Сферической симметрии метрики (34.1) соответствуют три векторных поля Киллинга, которые имеют нетривиальные компоненты только относительно базисных векторных полей ∂_θ и ∂_φ . Это – обычные поля Киллинга на сфере (29.6). Кроме этого имеется четвертое векторное поле Киллинга, которое в координатах Шварцшильда имеет вид ∂_t . Оно соответствует статичности метрики Шварцшильда вне горизонта в координатах Шварцшильда.

Наличие четвертого векторного поля Киллинга для любого сферически симметричного решения вакуумных уравнений Эйнштейна известно как *теорема Бирхгоффа* [253]. В разделе 33 мы доказали аналогичное утверждение в более общем случае. А именно, четвертый вектор Киллинга появляется в любом решении вакуумных уравнений Эйнштейна с космологической постоянной, которое имеет вид сплетенного произведения двух поверхностей. При этом мы не требовали симметрии метрики относительно какой-либо группы преобразований. Группа симметрии возникает автоматически при решении уравнений Эйнштейна. Это может быть либо группа вращений $\mathbb{SO}(3)$, либо группа Лоренца $\mathbb{SO}(1, 2)$, либо группа Пуанкаре $\mathbb{IO}(2)$, которые действуют на сечениях $t = \text{const}$. Глобальная структура всех таких решений была найдена и классифицирована в разделе 33.

Ясно, что в метрике Шварцшильда (34.1) можно произвольно менять радиальную координату $\rho \mapsto \rho'(\rho)$, при этом метрика также будет удовлетворять вакуумным уравнениям Эйнштейна. Обычно в пользу выбора радиальной координаты Шварцшильда ρ приводят следующий “аргумент”. Площадь поверхности сферы $\mathbb{S}^2$, определяемой уравнением $\rho = \text{const} > 2M$, такая же, как и в евклидовом пространстве $4\pi\rho^2$. Соответственно, длина окружности, лежащей в экваториальной плоскости сферы, равна $2\pi\rho$. У этого “аргумента” нет никаких математических или физических оснований. Заметим также, что “центр” у сферы отсутствует, т.к. при $\rho = 0$ сфера вырождается в точку, которой соответствует пространственноподобная, а не времениподобная кривая в плоскости t, ρ .

Важным является то обстоятельство, что постоянная M , входящая в решение Шварцшильда, является постоянной интегрирования, и вообще не содержится в вакуумных уравнениях Эйнштейна. С математической точки зрения постоянная M может принимать произвольные значения $M \in \mathbb{R}$. Однако, если предположить, что решение Шварцшильда описывает гравитационное поле вне точечной массивной частицы или некоторого сферически симметричного распределения масс, то сравнение с законом всемирного тяготения Ньютона на больших расстояниях, которое было проведено в разделе 23.10, говорит о том, что постоянная интегрирования M – это масса частицы, которая, следовательно, должна быть положительной.

Допустим, что решение Шварцшильда описывает солнечную систему, когда в качестве источника гравитационного поля рассматривается Солнце, а планеты и другие тела – в качестве пробных частиц, движущихся по экстремалиям (геодезическим) в гравитационном поле Солнца. Тогда возникает естественный и нетривиальный вопрос: какому значению координаты ρ соответствует орбита, например, Земли? Более общо, какое отношение имеет координата ρ , входящая в решение Шварцшильда, к расстоянию до Солнца, которое мы измеряем? Как было отмечено выше, в решении Шварцшильда мы можем произвести произвольную замену радиальной координаты $\rho \mapsto \rho'(\rho)$, не меняющую асимптотики на бесконечности. Новая метрика также будет удовлетворять вакуумным уравнениям Эйнштейна и будет асимптотически плоской. При этом с законом тяготения Ньютона будет все в порядке, поскольку он определяется только асимптотикой при $\rho \rightarrow \infty$. Возможный ответ на этот вопрос заключает-

ся в измерении площади поверхности сферы S_r достаточно большого радиуса r или длины окружности L_r того же радиуса. Площадь поверхности и длина окружности являются геометрическими инвариантами и, следовательно, не зависят от выбора системы координат. После этого надо вычислить отношения S_r/r^2 и L_r/r , где r – измеренный радиус. Если отношения будут равны 4π и 2π , соответственно, то можно утверждать, что измеряемый радиус – это действительно радиальная координата Шварцшильда, $r = \rho$.

Другой способ. Посылаем космический корабль со скоростью света к горизонту (Солнцу) и обратно. После возвращения, смотрим на показания часов, находящихся на корабле. Это будет собственное время, которое равно удвоенному значению канонического параметра для светоподобной экстремали от Земли до горизонта, и которое является геометрическим инвариантом. Из этого условия можно найти значение наблюдаемого радиуса r .

В общей теории относительности мы требуем, чтобы пространство-время было максимально продолжено вдоль экстремалей (геодезических). Это означает, что любую экстремаль на псевдоримановом многообразии (M, g) можно либо продолжить до бесконечного значения канонического параметра в обе стороны, либо при конечном значении канонического параметра она попадет в сингулярную точку, где какой-либо из геометрических инвариантов обратится в бесконечность. Более подробно этот вопрос обсуждался в разделе 23.2. Метрика Шварцшильда в координатах Шварцшильда (34.1) геодезически неполна на горизонте, и многообразие (пространство-время) можно продолжить, например, перейдя к координатам Крускала–Секереша.

В сферически симметричном случае четырехмерное пространство-время представляет собой топологическое произведение $M = U \times S^2$ двумерной лоренцевой поверхности U на сферу S^2 . Максимально продолженные вдоль геодезических поверхности U удобно изображать в виде диаграмм Картера–Пенроуза, которые были описаны в главе 30. Чтобы пояснить смысл диаграмм и ввести обозначения, рассмотрим

Пример 34.1.1. Плоскость Минковского. Рассмотрим плоскость Минковского $\mathbb{R}^{1,1}$ с декартовыми координатами t, x . Введем светоподобные координаты

$$u := t + x, \quad v := t - x$$

и отобразим всю плоскость на квадрат вдоль светоподобных направлений, например, положим

$$U := \arctg u \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \quad V := \arctg v \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \quad (34.5)$$

который изображен на рис.34.1.

Этот квадрат представляет собой диаграмму Картера–Пенроуза для плоскости Минковского $\mathbb{R}^{1,1}$. На диаграмме буквы i^+ и i^- обозначают пространственноподобные бесконечности будущего и прошлого, соответственно. Они изображены закрашенными кружками, что означает их геодезическую полноту. Буквы $J^\pm$ обозначают светоподобные бесконечности будущего и прошлого. Они изображены жирными прямыми линиями, т.к. геодезически полны. Буквы i^0 обозначают левую и правую времениподобные пространственные бесконечности. Они также геодезически полны и поэтому изображены закрашенными кружками.

Пространство-время Минковского $\mathbb{R}^{1,1}$ является геодезически полным многообразием. Геодезическими являются прямые линии и только они на плоскости t, x .

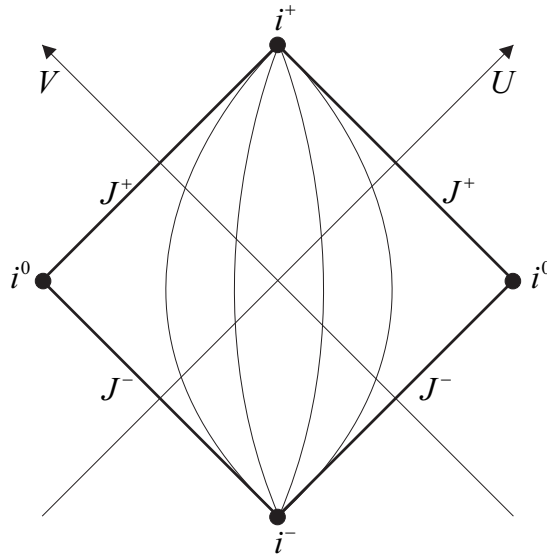


Рис. 34.1: Диаграмма Карттера–Пенроуза для плоскости Минковского $\mathbb{R}^{1,1}$

Поскольку отображение (34.5) проводится вдоль светоподобных направлений, то светоподобные геодезические на диаграмме Карттера–Пенроуза являются прямыми отрезками, идущими как и прежде под углами $\pm\pi/4$. Они начинаются и заканчиваются на сторонах квадрата. Это свойство позволяет легко представить причинную структуру пространства-времени, т.к. световые конусы в каждой точке такие же как и на исходной плоскости Минковского. После отображения (34.5) времени- и пространственноподобные геодезические уже не будут прямыми линиями на диаграмме Карттера–Пенроуза за исключением осей абсцисс и ординат. Все времениподобные геодезические начинаются в пространственноподобной бесконечности прошлого i^- и заканчиваются в пространственноподобной бесконечности будущего i^+ . Все пространственноподобные геодезические соединяют левую и правую пространственные бесконечности i^0 .

Плоскость Минковского имеет много векторных полей Киллинга, соответствующих трансляциям и вращениям. Трансляциям по времени соответствует поле Киллинга ∂_t . Его интегральные кривые на плоскости t, x – это прямые, параллельные оси ординат t . На диаграмме Карттера–Пенроуза они показаны в виде тонких сплошных линий, соединяющих бесконечность прошлого i^- и будущего i^+ .

Хотя бесконечности прошлого и будущего $i^\pm$ изображены на диаграмме Карттера–Пенроуза в виде точек (кружков), следует помнить, что в пространстве-времени $\mathbb{R}^{1,1}$ им соответствуют пространственноподобные прямые бесконечной длины. Это же относится и к пространственным бесконечностям i^0 . Если на плоскости Минковского $\mathbb{R}^{1,1}$ заданы какие-либо функции, то на диаграмме Карттера–Пенроуза бесконечности $i^\pm$ и i^0 будут являться существенно особыми точками для этих функций, т.к. пределы в общем случае будут зависеть от пути, по которому мы к ним приближаемся. $\square$

Максимально продолженное пространство-время $\mathbb{M}$, соответствующее метрике Шварцшильда (34.1), представляет собой топологическое произведение некоторой лоренцевой поверхности $t, \rho \in \mathbb{U}$ на сферу $\theta, \varphi \in \mathbb{S}^2$: $\mathbb{M} = \mathbb{U} \times \mathbb{S}^2$. Сфера является геодезически полным многообразием без каких-либо особенностей. Поэтому в дальнейшем мы сосредоточимся на исследовании лоренцевой поверхности $\mathbb{U}$ с координатами

t, ρ , которые входят в метрику (34.1), и соответствующей двумерной частью метрики

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) dt^2 - \frac{d\rho^2}{1 - \frac{2M}{\rho}}. \quad (34.6)$$

Эта двумерная метрика имеет один вектор Киллинга ∂_t . Общий подход к построению максимально продолженных лоренцевых поверхностей с одним вектором Киллинга был описан в главе 30, где координата ρ обозначалась буквой q поскольку далеко не всегда ей можно было придать смысл какого-либо радиуса. В частности, двумерная метрика (34.6) была рассмотрена в разделе 30.5.1.

Повторим результат. Максимально продолженная вдоль экстремалей лоренцева поверхность $\mathbb{U}$ изображается в виде диаграммы Картера–Пенроуза, показанной на рис.34.2. В квадратных статических конформных блоках I и III координата ρ ме-

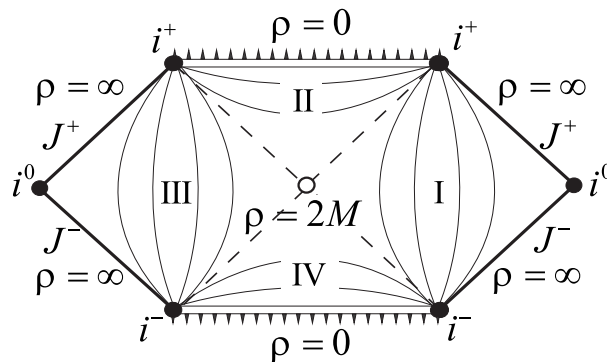


Рис. 34.2: Диаграмма Картера–Пенроуза для решения Шварцшильда.

няется в пределах $2M < \rho < \infty$, пространственноподобна и является аналогом радиальной координаты сферической системы координат. В блоке I координата ρ возрастает слева направо, а в блоке III – справа налево. В треугольных однородных конформных блоках II и IV координата ρ меняется в пределах $0 < \rho < 2M$ и является времениподобной. В блоке II она возрастает сверху вниз, а в блоке IV – снизу вверх. Конформные блоки на диаграмме Картера–Пенроуза склеены вдоль горизонтов, соответствующих радиусу Шварцшильда $\rho_s = 2M$, которые показаны на диаграмме пунктирными линиями. Центральная точка (место пересечения горизонтов), которая на диаграмме Картера–Пенроуза помечена окружностью, является седловой точкой для координаты ρ и геодезически не полна.

На диаграмме Картера–Пенроуза пространственноподобные бесконечности будущего и прошлого i^+ и i^- , левая и правая времениподобные пространственные бесконечности i^0 являются геодезически полными и изображены закрашенными кружками. Светоподобные бесконечности будущего и прошлого J^+ и J^- геодезически полны и показаны жирными прямыми. Горизонты обозначены пунктирными линиями внутри диаграммы, поскольку не являются геодезически полными. Сингулярные границы, соответствующие $\rho = 0$, геодезически не полны.

Наглядность диаграммы Картера–Пенроуза заключается в том, что световые конусы в каждой внутренней точке образованы двумя пересекающимися прямыми, идущими под углами $\pm\pi/4$, так же, как и на плоскости Минковского $\mathbb{R}^{1,1}$. Отсюда сразу следует, что никакая времениподобная или светоподобная экстремаль (геодезическая), которая начинается внутри блока II, не может попасть в области I или III, и с необходимостью попадет на сингулярность $\rho = 0$, расположенную сверху, при

конечном значении канонического параметра. По этой причине верхнюю сингулярность $\rho = 0$ называют сингулярностью черной дыры, а область II под горизонтом – внутренностью черной дыры.

Максимально продолженное решение Шварцшильда описывает две вселенные: I и III. Мы находимся либо во вселенной I, либо в III. Эти вселенные между собой не связаны причинно: никакой наблюдатель из вселенной I не может попасть во вселенную III и наоборот, т.к. никакие их точки нельзя соединить времени- или светоподобной кривой. При этом у наблюдателя из вселенной I есть две возможности: либо жить вечно, чему соответствуют времениподобные кривые, уходящие в пространственноподобную бесконечность будущего i^+ (она изображена закрашенным кружком в верхнем правом углу диаграммы), либо за конечное собственное время угодить в черную дыру $\rho = 0$ после пересечения горизонта.

Заметим, что с точки зрения внешнего наблюдателя, который измеряет координаты Шварцшильда, приближение к горизонту займет бесконечное время t . Это легко видеть для радиальных светоподобных экстремалей, которые определяются условием $ds^2 = 0$. Действительно, для метрики (34.6) радиальные светоподобные экстремали определяются уравнением

$$\left(1 - \frac{2M}{\rho}\right) dt^2 - \frac{d\rho^2}{1 - \frac{2M}{\rho}} = 0.$$

Отсюда следует равенство

$$t = \int \frac{\rho d\rho}{\rho - 2M}.$$

Этот интеграл логарифмически расходится на горизонте при $\rho = 2M$. Следовательно, радиальные светоподобные экстремали с точки зрения внешнего наблюдателя приближаются к горизонту бесконечно долго. Аналогичный анализ можно провести и для времениподобных экстремалей. То есть внешний наблюдатель никогда не увидит прохождение чего-либо через горизонт. С точки зрения же падающего наблюдателя он благополучно пересечет горизонт при конечном значении собственного времени и продолжит падение на черную дыру, пока не достигнет сингулярности $\rho = 0$ при конечном значении канонического параметра (собственного времени).

Светоподобные экстремали из вселенной I попадают либо в правую светоподобную бесконечность будущего J^+ при бесконечном значении канонического параметра, либо достигают сингулярности черной дыры при конечном значении канонического параметра.

Из внутренности белой дыры (область IV) времениподобные экстремали могут попасть во все три области I, II и III. Однако никакой наблюдатель из вселенной I или III не может попасть внутрь белой дыры.

Времениподобные пространственные бесконечности слева и справа i^0 можно соединить с внутренней точкой диаграммы Картера–Пенроуза только пространственноподобной кривой при условии, что тип кривой не меняется от точки к точке.

С точки зрения задачи Коши, решение волнового уравнения в области I определяется начальными данными вблизи сингулярности белой дыры $\rho = 0$ и правой пространственноподобной бесконечности прошлого i^- .

В разделе 30.5.1 мы доказали, что диаграмма Картера–Пенроуза для решения Шварцшильда представляет собой гладкую двумерную лоренцеву поверхность с гладкой метрикой. При этом лоренцева поверхность с метрикой (34.6) при $-\infty < t < \infty$

и $2M < \rho < \infty$ изометрична области I и области III. Лоренцева поверхность с координатами $-\infty < t < \infty$ и $0 < \rho < 2M$ изометрична области II и области IV.

Векторное поле Киллинга ∂_t времениподобно в блоках I и III (метрика статична) и пространственноподобно в блоках II и IV (метрика однородна). Соответствующие траектории Киллинга показаны на диаграмме Картера–Пенроуза тонкими сплошными линиями внутри блоков. Горизонты на поверхности $\mathbb{U}$ являются светоподобными траекториями Киллинга. Это становится ясным, когда метрика Шварцшильда (34.6) записана в координатах Эддингтона–Финкельстейна или Крускала–Секереша.

Таким образом, координаты Шварцшильда не покрывают всю лоренцеву поверхность $\mathbb{U}$ и, следовательно, не являются глобальными.

34.2 Координаты Эддингтона–Финкельстейна

Для того, чтобы доказать дифференцируемость метрики при склейке конформных блоков вдоль горизонтов в разделе 30.6 мы использовали координаты Эддингтона–Финкельстейна, которые первоначально были введены именно для решения Шварцшильда [230, 231]. Перейдем от координат Шварцшильда к координатам Эддингтона–Финкельстейна $t, \rho \mapsto \xi, \rho$, где

$$t := \xi - \int_{\frac{2M}{\rho}}^{\rho} \frac{dr}{1 - \frac{2M}{r}} = \xi - \rho - 2M \ln|\rho - 2M|. \quad (34.7)$$

Это преобразование временной координаты определено при $0 < \rho < 2M$ и $2M < \rho < \infty$. Постоянная интегрирования соответствует сдвигу t , и мы, для простоты, положили ее равной нулю. Для соответствующих дифференциалов справедливо равенство

$$dt = d\xi - \frac{d\rho}{1 - \frac{2M}{\rho}},$$

и метрика Шварцшильда (точнее, ее двумерная t, ρ часть) в новых координатах принимает вид

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) d\xi^2 - 2d\xi d\rho. \quad (34.8)$$

Эта метрика невырождена, $\det g_{\alpha\beta} = -1$, и определена при всех значениях координат

$$-\infty < \xi < \infty, \quad 0 < \rho < \infty, \quad (34.9)$$

т.е. на полуплоскости ξ, ρ .

Преобразование координат Шварцшильда к координатам Эддингтона–Финкельстейна (34.7) хорошо определено при $0 < \rho < 2M$ и $2M < \rho < \infty$, однако вырождено на горизонте ρ_s . Таким образом внутренность конформных блоков I и II отображается на две непересекающиеся области полуплоскости (34.9), которые гладко склеены вдоль горизонта ρ_s . Эта полуплоскость на диаграмме Картера–Пенроуза (см. рис.34.3 слева) изображена в виде диагональной цепочки конформных блоков, идущих сверху слева вниз направо. Вдоль этой цепочки координата ρ увеличивается от нуля до бесконечности, а ξ меняется от $-\infty$ до $+\infty$. Напомним, что горизонт ρ_s достигается при бесконечном значении t , как было показано в предыдущем разделе.

Координаты Шварцшильда не различают конформные блоки типа I и III, а также блоки типа II и IV. Поэтому метрике (34.8) соответствует также вторая цепочка

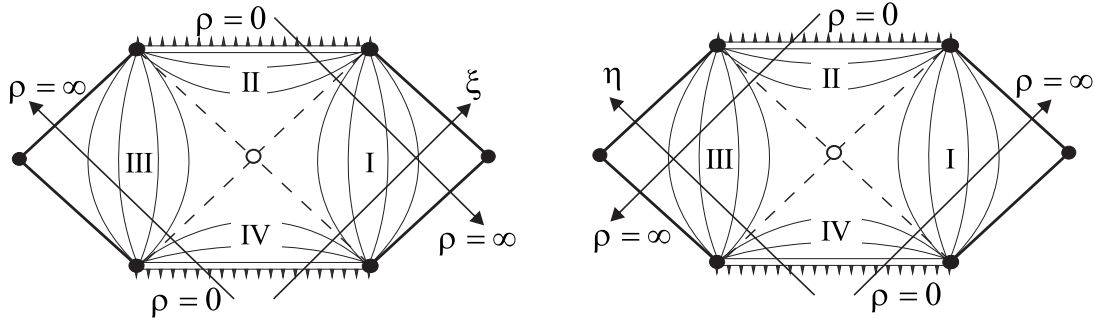


Рис. 34.3: Координаты Эддингтона–Финкельштейна ξ, ρ (слева) и η, ρ (справа) на диаграмме Картера–Пенроуза для решения Шварцшильда.

конформных блоков, идущая снизу справа наверх налево (возрастание координаты ρ).

Введенные координаты Эддингтона–Финкельштейна склеивают конформные блоки I–II и III–IV. Можно также ввести другие координаты Эддингтона–Финкельштейна $t, \rho \mapsto \eta, \rho$, которые склеивают блоки I–IV и II–III. Формула соответствующего преобразования имеет вид

$$t := \eta + \int^{\rho} \frac{dr}{1 - \frac{2M}{r}} = \eta + \rho + 2M \ln|\rho - 2M|,$$

где, как и раньше, $0 < \rho < 2M$ или $2M < \rho < \infty$. Метрика в этих координатах примет вид

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) d\eta^2 + 2d\eta d\rho.$$

Полученная метрика невырождена и определена на полуплоскости

$$-\infty < \eta < \infty, \quad 0 < \rho < \infty.$$

На диаграмме Картера–Пенроуза (см. рис.34.3 справа) эта полуплоскость изображается в виде диагональных цепочек конформных блоков, которые идут либо снизу слева вверх направо, либо сверху справа вниз налево, что соответствует возрастанию координаты ρ .

Двумерная поверхность с метрикой (34.8) не является максимально продолженной. Можно показать, что экстремали на полуплоскости (34.9) неполны при $\xi \rightarrow -\infty$ для цепочки конформных блоков I–II и при $\xi \rightarrow \infty$ для цепочки конформных блоков II–IV. Чтобы ввести глобальные координаты на всей диаграмме Картера–Пенроуза введем координаты Крускала–Секереша.

34.3 Координаты Крускала–Секереша

Координаты Крускала–Секереша U, V являются глобальными, покрывают всю лоренцеву поверхность $\mathbb{U}$ и вводятся следующим образом. Для ясности, введем координаты Крускала–Секереша в несколько этапов. Во-первых, запишем двумерную часть метрики Шварцшильда (34.6) в *черепаших координатах* $t, \rho \mapsto t, r$ (для определенности мы рассмотрим область вне горизонта $\rho > 2M$),

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) (dt^2 - dr^2), \quad (34.10)$$

где новая координата r связана со старой дифференциальным уравнением

$$\frac{d\rho}{dr} = \pm \left(1 - \frac{2M}{\rho}\right). \quad (34.11)$$

В этом уравнении знак “+” соответствует статическому блоку I, а знак “−” – статическому блоку III. Это соответствует преобразованию $r \mapsto -r$. В обоих случаях новая координата r , если ее изобразить на диаграмме Картера–Пенроуза, возрастает слева направо. Уравнение (34.11) легко решается:

$$r = \pm (\rho + 2M \ln|\rho - 2M|), \quad (34.12)$$

где мы положили несущественную постоянную интегрирования, которая соответствует сдвигу r , равной нулю.

Если радиальная координата $\rho \in (2M, \infty)$, то пространственная координата r меняется от $-\infty$ до ∞ . Тем самым внешнее решение Шварцшильда изометрично всей плоскости t, r с метрикой (34.10).

Теперь на плоскости t, r введем конусные координаты $t, r \mapsto \xi, \eta$, где

$$\xi := t + r, \quad \eta := t - r.$$

Тогда метрика примет вид

$$ds^2 = \Phi d\xi d\eta,$$

где

$$\Phi := 1 - \frac{2M}{\rho}$$

– конформный множитель, в котором ρ рассматривается как неявная функция от ξ и η .

Область I

Введем координаты Крускала–Секереша, совершив конформное преобразование $\xi, \eta \mapsto U, V$, где

$$U := 2e^{\frac{\xi}{4M}} > 0, \quad V := -2e^{-\frac{\eta}{4M}} < 0. \quad (34.13)$$

Отсюда следуют связи между дифференциалами:

$$d\xi = 4M \frac{dU}{U}, \quad d\eta = -4M \frac{dV}{V}.$$

Поэтому метрика (34.10) принимает следующий вид

$$ds^2 = -16M^2 \Phi \frac{dU dV}{UV}. \quad (34.14)$$

Поскольку

$$r = \frac{1}{2}(\xi - \eta) = 2M \ln \left(-\frac{UV}{4} \right), \quad (34.15)$$

то

$$-\frac{UV}{4} = e^{\frac{r}{2M}}.$$

Учитывая связь радиальных координат (34.12), где для области I выбирается знак “+”, метрику (34.14) запишем в виде

$$ds^2 = \frac{4M^2}{\rho} e^{-\frac{\rho}{2M}} dU dV, \quad (34.16)$$

где радиальная координата Шварцшильда ρ является неявной функцией произведения UV , которая определяется формулами (34.12) и (34.15):

$$-\frac{UV}{4} = |\rho - 2M| e^{\frac{\rho}{2M}}. \quad (34.17)$$

Простые вычисления приводят к следующему выражению для временной координаты Шварцшильда

$$-\frac{U}{V} = e^{\frac{t}{2M}}. \quad (34.18)$$

Это значит, что постоянному времени $t = \text{const}$ в координатах Крускала–Секереша соответствуют прямые линии на плоскости U, V , проходящие через начало координат.

Поскольку координаты ξ, η меняются на всей плоскости, то координаты Крускала–Секереша (34.13) меняются в квадранте $U > 0, V < 0$. Таким образом внешнее решение Шварцшильда изометрично отображено на квадрант I, изображенный на рис.34.4.

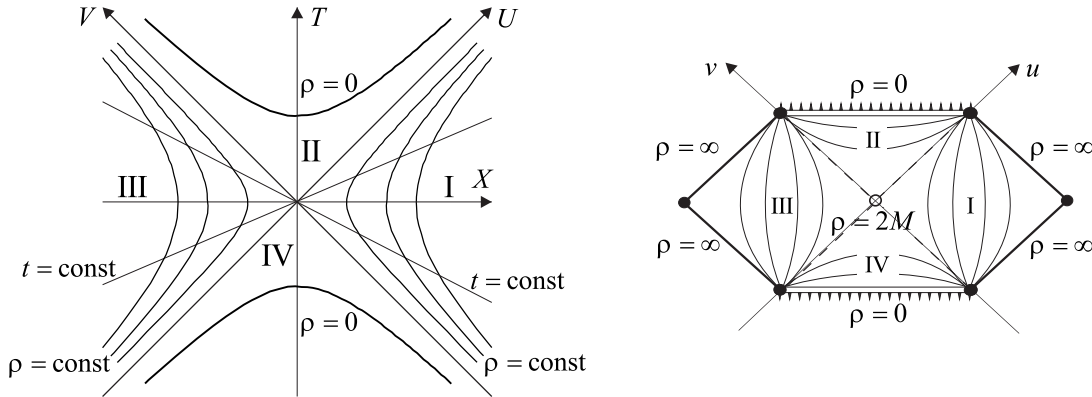


Рис. 34.4: Координаты Крускала–Секереша U, V для решения Шварцшильда (слева) и диаграмма Картера–Пенроуза (справа).

Область III

Теперь на плоскости ξ, η введем новые координаты

$$U := -2e^{-\frac{\xi}{4M}} < 0, \quad V := 2e^{\frac{\eta}{4M}} > 0.$$

Такие же вычисления, как и в области I, приводят к той же метрике (34.16). Однако теперь, на плоскости U, V она определена в квадранте III. Области III в уравнении (34.12) соответствует знак “−” и поэтому связь радиальной координаты Шварцшильда ρ с координатами Крускала–Секереша задается тем же уравнением (34.17). Таким образом мы изометрично отобразили внешнее решение Шварцшильда на квадрант III.

Простые вычисления приводят к тому же выражению (34.18) для временной координаты Шварцшильда, что и в области I.

Область II

Преобразование координаты (34.12) определено также при $0 < \rho < 2M$, что соответствует внутреннему решению Шварцшильда. Разница заключается в том, что теперь координаты r, t определены не на всей плоскости, а только на половине. Если в уравнении выбран знак “+”, то координата r меняется от $r = 2M \ln 2M$ при $\rho = 0$ до $-\infty$ при $\rho = 2M$. Если выбран знак “–”, то $\rho \in (-2M \ln 2M, \infty)$.

Выберем в уравнении (34.12) знак “+”. Поскольку под горизонтом временная и пространственная координаты меняются местами, то введем конусные координаты следующим образом:

$$\xi := r + t, \quad \eta := r - t. \quad (34.19)$$

что соответствует повороту плоскости r, t на 90° градусов. Теперь перейдем к координатам Крускала–Секереша

$$U := 2e^{\frac{\xi}{4M}} > 0, \quad V := 2e^{\frac{\eta}{4M}} > 0.$$

Тогда внутренняя область отобразится в квадрант II. Причем сверху образ внутренней области на U, V плоскости будет определяться уравнением $UV < 8M$, т.е. будет ограничен гиперболой $UV = 8M$. Таким образом внутренняя область решения Шварцшильда отображена в квадрант II. При этом метрика, как легко проверить будет иметь тот же вид, что и в первом квадранте (34.16). Связь радиальной координаты Шварцшильда с координатами Крускала–Секереша задается равенством

$$\frac{UV}{4} = |\rho - 2M| e^{\frac{\rho}{2M}}. \quad (34.20)$$

Это уравнение отличается знаком от уравнения (34.17) для статичных областей, поскольку в квадранте II произведение $UV > 0$.

Простые вычисления приводят к следующему выражению для временной координаты Шварцшильда

$$\frac{U}{V} = e^{\frac{t}{2M}}. \quad (34.21)$$

Область IV

Выберем в уравнении (34.12) знак “–”. Поскольку нас интересует внутренняя область решения Шварцшильда, то введем координаты светового конуса (34.19) и перейдем к координатам Крускала–Секереша:

$$U := -2e^{-\frac{\xi}{4M}} < 0, \quad V := -2e^{-\frac{\eta}{4M}} < 0.$$

Поскольку координаты r, t определены на полуплоскости, то внутренняя область решения Шварцшильда отображается в квадрант IV, причем на координаты Крускала–Секереша наложено условие $UV < 8M$, т.е. образ внутренней области ограничен снизу гиперболой $UV = 8M$. Метрика, как и прежде, будет иметь вид (34.16). Нетрудно также проверить, что координата ρ связана с координатами Крускала–Секереша тем же соотношением, что и в области II (34.20).

Простые вычисления приводят к тому же выражению (34.21) для временной координаты Шварцшильда, что и в области II.

Заметим, что связь координаты ρ с координатами Крускала–Секереша во всех четырех областях можно записать в виде одного уравнения

$$-\frac{UV}{4} = (\rho - 2M) e^{\frac{\rho}{2M}}, \quad (34.22)$$

опустив знак модуля.

Таким образом, мы отобразили два экземпляра внешней области решения Шварцшильда $\rho > 2M$ на квадранты I, III и два экземпляра внутреннего решения – в квадранты II, IV, которые ограничены условием $UV < 8M$. Во всех четырех квадрантах метрика имеет вид (34.16). При этом связь координаты ρ с координатами Крускала–Секереша определяется уравнением (34.22). Метрика в координатах Крускала–Секереша (34.16) определена на части плоскости U, V , которая ограничена сверху и снизу двумя ветвями гиперболы $UV = 8M$, как показано на рис.34.4. Горизонтам соответствуют координатные линии $U = 0$ и $V = 0$.

На плоскости U, V можно ввести координаты T, X :

$$U := T + X, \quad V := T - X, \quad (34.23)$$

где T – глобальное время и X – глобальная пространственная координата.

Для того, чтобы из метрики Шварцшильда в координатах Крускала–Секереша получить диаграмму Картера–Пенроуза, необходимо отобразить конусные координаты $U \in \mathbb{R}$ и $V \in \mathbb{R}$ на конечные интервалы вещественной оси $(-a, a) \subset \mathbb{R}$. Это можно сделать, например, с помощью преобразования координат

$$u := \operatorname{arctg} U \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \quad v := \operatorname{arctg} V \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Тогда на плоскости u, v мы получим диаграмму Картера–Пенроуза. Для полного счастья функции $u(U)$ и $v(V)$ нужно подобрать таким образом, чтобы гиперболы $UV = 8M$ отображались в горизонтальные прямые на плоскости u, v . Ясно, что это всегда можно сделать.

34.4 Мост Эйнштейна–Розена

В 1935 году Эйнштейн и Розен предложили следующую модификацию решения Шварцшильда (34.1) [256]. Если решение Шварцшильда рассматривать вне горизонта $\rho > 2M$, то можно ввести новую “радиальную” координату $\rho \mapsto u(\rho)$ с помощью уравнения

$$\frac{1}{2}u^2 = \rho - 2M > 0. \quad (34.24)$$

Прямые вычисления приводят к следующей метрике в новых координатах

$$ds^2 = \frac{u^2}{u^2 + 4M} dt^2 - (u^2 + 4M) du^2 - \frac{1}{4}(u^2 + 4M)^2 (d\theta^2 + \sin^2 \theta d\varphi^2). \quad (34.25)$$

Преобразование радиальной координаты (34.24) проведено таким образом, что если ρ увеличивается от $2M$ до бесконечности, то новая координата либо увеличивается от 0 до ∞ , либо уменьшается от 0 до $-\infty$.

Поскольку метрика (34.25) получена преобразованием координат из решения Шварцшильда, то она также удовлетворяет уравнениям Эйнштейна по крайней мере в областях $u > 0$ и $u < 0$. Эта метрика определена на всей вещественной оси $u \in \mathbb{R}$. При этом два экземпляра внешнего решения Шварцшильда, которым соответствуют положительные и отрицательные значения u сшиты на гиперповерхности $u = 0$, которая представляет собой топологическое произведение прямой $t \in \mathbb{R}$ на сферу $\theta, \varphi \in \mathbb{S}^2$. Компоненты метрики (34.25) являются гладкими функциями, поэтому можно предположить, что уравнения Эйнштейна удовлетворяются при всех

значениях u . Однако здесь надо проявить осторожность. Дело в том, что при $u = 0$ метрика Эйнштейна–Розена вырождена, ее обратная неопределена, и поэтому сами уравнения Эйнштейна требуют доопределения. Мы отложим этот вопрос до конца раздела.

Метрика (34.25), очевидно, инвариантна относительно инверсии $u \mapsto -u$.

В своей оригинальной работе Эйнштейн и Розен дали следующую физическую интерпретацию метрики (34.25). Они предположили, что точечная массивная частица массы M покоится в “центре” сферической системы координат $u = 0$, а пространство-время вокруг нее описывается двумя листами $u < 0$ и $u > 0$ с метрикой (34.25). При этом оба листа описывают одну и ту же вселенную.

В настоящее время мосту Эйнштейна–Розена принято давать другую физическую интерпретацию. С этой целью проведем следующее построение. Мы видим, что метрика (34.25) описывает два внешних решения Шварцшильда, которые гладко сшиты вдоль гиперповерхности $u = 0$. Поэтому рассмотрим движение пробных частиц в экваториальной плоскости $\theta = \pi/2$. Они движутся в трехмерном пространстве-времени с метрикой

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) dt^2 - \frac{d\rho^2}{1 - \frac{2M}{\rho}} - \rho^2 d\varphi^2, \quad \rho > 2M,$$

которую мы записали в координатах Шварцшильда. Для таких пробных частиц пространством являются сечения $t = \text{const}$. То есть в данном случае пространство – это двумерное многообразие (поверхность) с метрикой

$$dl^2 = \frac{d\rho^2}{1 - \frac{2M}{\rho}} + \rho^2 d\varphi^2, \quad \rho > 2M, \quad (34.26)$$

где мы, для наглядности, изменили знак метрики. Будем считать, что ρ и φ – это обычные полярные координаты на евклидовой плоскости.

Для того, чтобы лучше представить себе поверхность с метрикой (34.26) вложим ее в трехмерное евклидово пространство [257]. Пусть ρ, φ, z – цилиндрические координаты в $\mathbb{R}^3$ с евклидовой метрикой

$$ds^2 := d\rho^2 + \rho^2 d\varphi^2 + dz^2.$$

Допустим, что вращательно симметричное вложение задается некоторой функцией $z(\rho)$. Тогда индуцированная метрика на вложенной поверхности будет иметь вид

$$dl^2 = \left[1 + \left(\frac{dz}{d\rho}\right)^2\right] d\rho^2 + \rho^2 d\varphi^2.$$

Для того, чтобы эта метрика совпала с метрикой (34.26) необходимо положить

$$\left(\frac{dz}{d\rho}\right)^2 = \frac{1}{\frac{\rho}{2M} - 1}.$$

Это уравнение может быть выполнено только для внешнего решения Шварцшильда $\rho > 2M$. Его общее решение имеет вид

$$z = \int \frac{d\rho}{\sqrt{\frac{\rho}{2M} - 1}} = 4M \sqrt{\frac{\rho}{2M} - 1} + \text{const}. \quad (34.27)$$

Постоянная интегрирования соответствует сдвигу координаты z , и мы положим ее равной нулю. Отсюда следует, что поверхность с метрикой (34.26) можно вложить в трехмерное евклидово пространство $\mathbb{R}^3$, которая задана уравнением

$$\frac{\rho}{2M} = 1 + \left(\frac{z}{4M} \right)^2.$$

Это – параболоид вращения, который называется *параболоидом Фламма* и изображен на рис.34.5а. При этом вложении верхняя ($z > 0$) и нижняя ($z < 0$) части параболоида изометричны экваториальным сечениям ($\theta = \pi/2$, $t = \text{const}$) внешнего решения Шварцшильда. Эти сечения гладко сшиты вдоль горловины $z = 0$, которая соответствует горизонту $\rho_s = 2M$.

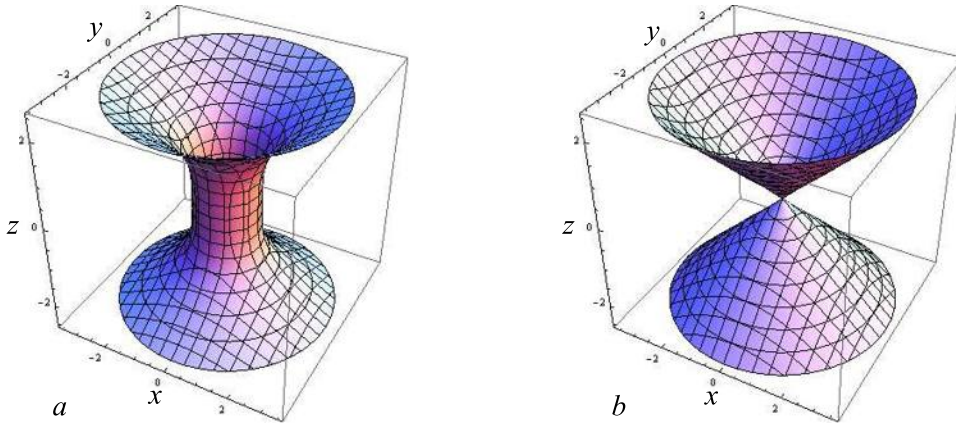


Рис. 34.5: Параболоид Фламма в координатах $\rho^2 = x^2 + y^2$ (а). Та же поверхность в координатах $u^2 = x^2 + y^2$ (b).

Параболоид Фламма глобально изометричен экваториальным сечениям $\theta = \pi/2$, $t = \text{const}$ моста Эйнштейна–Розена. В координатах u, φ (34.24) вложение задается уравнением

$$z^2 = 4Mu^2 \quad \Leftrightarrow \quad z = \pm \sqrt{4Mu}.$$

То есть параболоид превращается в два конуса, и горловина моста стягивается в точку (см. рис.34.5b).

Проведенное построение позволяет дать следующую физическую интерпретацию моста Эйнштейна–Розена. Мы имеем две вселенные, которые соответствуют положительным и отрицательным значениям координаты z . Обе вселенные являются асимптотически плоскими на больших расстояниях $\rho \rightarrow \infty$. В “центре” $\rho = 2M \Leftrightarrow u = 0$ вселенные склеены. Такую конфигурацию называют *кротовой норой*, поскольку через нее мы имеем возможность попасть из одной вселенной в другую.

Теперь рассмотрим как ведет себя кривизна моста Эйнштейна–Розена. Матрица Якоби преобразования координат $(t, u, \theta, \varphi) \mapsto (t, \rho, \theta, \varphi)$ диагональна,

$$\frac{\partial(t, \rho, \theta, \varphi)}{\partial(t, u, \theta, \varphi)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & u & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Поэтому при преобразовании координат Шварцшильда к координатам Эйнштейна–Розена меняются только те компоненты тензора кривизны, которые содержат индекс

$1 = u$. Из формул (34.3) следуют явные выражения для компонент тензора кривизны моста Эйнштейна–Розена:

$$\begin{aligned} R_{0u0u} &= -\frac{2Mu^2}{\left(\frac{1}{2}u^2 + 2M\right)^3}, & R_{u2u2} &= -2M, \\ R_{0202} &= \frac{Mu^2}{2\left(\frac{1}{2}u^2 + 2M\right)^2}, & R_{u3u3} &= -2M \sin^2 \theta, \\ R_{0303} &= \frac{Mu^2}{2\left(\frac{1}{2}u^2 + 2M\right)^2} \sin^2 \theta, & R_{2323} &= 2M \left(\frac{1}{2}u^2 + 2M\right) \sin^2 \theta. \end{aligned} \quad (34.28)$$

Квадрат тензора кривизны (=тензора Вейля) для моста Эйнштейна–Розена имеет вид

$$R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} = \frac{48M^2}{\left(\frac{1}{2}u^2 + 2M\right)^6}.$$

Отметим, что и компоненты тензора кривизны, и квадрат тензора кривизны являются гладкими функциями при $u = 0$, где сшиваются два внешних решения Шварцшильда.

Мы видим, что, несмотря на то, что метрика (34.25) вырождена при $u = 0$ и компонента обратной метрики g^{00} в этой точке имеет полюс, компоненты ковариантного тензора кривизны определены (полюс сокращается). Сокращение полюса происходит также в тензоре Риччи и скалярной кривизне. Поэтому можно считать, что вакуумные уравнения Эйнштейна

$$R_{\alpha\beta} = 0$$

выполнены во всем пространстве-времени $\mathbb{R}^2 \times \mathbb{S}^2$, где $t, u \in \mathbb{R}$.

В заключение раздела отметим, что аналогичным образом можно сшить два внутренних решения Шварцшильда. При $0 < \rho < 2M$ совершим преобразование координаты

$$\frac{1}{2}u^2 := 2M - \rho > 0. \quad (34.29)$$

В новых координатах метрика примет вид

$$ds^2 = \frac{u^2}{u^2 - 4M} dt^2 - (u^2 - 4M) du^2 - \frac{1}{4}(u^2 - 4M)(d\theta^2 + \sin^2 \theta d\varphi^2). \quad (34.30)$$

При этом два внутренних решения Шварцшильда отображаются на положительные и отрицательные значения координаты u :

$$0 < \rho < 2M \quad \Leftrightarrow \quad 0 < |u| < 2\sqrt{M}.$$

Для метрики (34.30) координата t является пространственноподобной, а u – времениподобной.

34.5 Изотропные координаты

Приведем пространственную часть метрики Шварцшильда

$$ds^2 = -\frac{d\rho^2}{1 - \frac{2M}{\rho}} - \rho^2(d\theta^2 + \sin^2 \theta d\varphi^2)$$

к конформно плоскому виду. С этой целью совершим преобразование радиальной координаты $\rho \mapsto \rho(r)$. Тогда пространственная часть метрики примет вид

$$ds^2 = -\frac{\rho'^2 dr^2}{1 - \frac{2M}{\rho}} - \rho^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

где штрих обозначает дифференцирование по r . Для того, чтобы эта метрика была конформно плоской необходимо и достаточно, чтобы функция $\rho(r)$ удовлетворяла уравнению

$$\frac{\rho^2}{r^2} = \frac{\rho'^2}{1 - \frac{2M}{\rho}}.$$

Рассмотрим область вне горизонта $2M < \rho < \infty$ и предположим, что $\rho' > 0$, т.е. радиальные координаты возрастают одновременно. Тогда функция $\rho(r)$ удовлетворяет уравнению

$$\frac{\rho}{r} = \frac{\rho'}{\sqrt{1 - \frac{2M}{\rho}}}.$$

Общее решение этого уравнения имеет вид

$$Cr = \sqrt{\rho^2 - 2M\rho} + \rho - M \quad \Leftrightarrow \quad \rho = \frac{(Cr + M)^2}{2Cr},$$

где $C > 0$ – постоянная интегрирования.

Пространственная часть метрики Шварцшильда в новых координатах является конформно евклидовой:

$$ds^2 = -\Phi[dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)],$$

где

$$\Phi(r) := \frac{(Cr + M)^4}{4C^2r^4}$$

– конформный множитель. Потребуем, чтобы полученная метрика была асимптотически евклидовой, т.е.

$$\lim_{r \rightarrow \infty} \Phi(r) \rightarrow 1.$$

Отсюда находится постоянная интегрирования: $C = 2$, что мы и предположим. Тогда преобразование радиальной координаты примет вид

$$\rho := r \left(1 + \frac{M}{2r}\right)^2. \quad (34.31)$$

Теперь нетрудно вычислить компоненту g_{00} метрики Шварцшильда:

$$1 - \frac{2M}{\rho} = \left(\frac{1 - \frac{M}{2r}}{1 + \frac{M}{2r}}\right)^2.$$

Таким образом, метрика Шварцшильда в новой системе координат принимает вид

$$ds^2 = \left(\frac{1 - \frac{M}{2r}}{1 + \frac{M}{2r}}\right)^2 dt^2 - \left(1 + \frac{M}{2r}\right)^4 [dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)]. \quad (34.32)$$

Если перейти к декартовой системе координат на пространственных сечениях, то при $M > 0$ эта метрика определена всюду в $\mathbb{R}^4$ за исключением мировой линии начала сферической системы координат,

$$-\infty < t < \infty, \quad 0 < r < \infty, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \varphi < 2\pi.$$

Очевидно, что полученная метрика является гладкой в остальной части $\mathbb{R}^4$.

Поскольку данная метрика получена из решения Шварцшильда преобразованием координат (34.31), то она удовлетворяет вакуумным уравнениям Эйнштейна всюду в $\mathbb{R}^4$ за исключением мировой линии начала сферической системы координат.

Каждое сечение постоянного времени $t = \text{const}$ является конформно плоским и, следовательно, изотропным. Поэтому координаты t, r, θ, φ называют *изотропными*. Изотропные координаты для метрики Шварцшильда давно и хорошо известны (см., например, [169], §100, задача 4).

Определитель метрики (34.32) равен

$$\det g_{\alpha\beta} = - \left(1 - \frac{M}{2r}\right)^2 \left(1 + \frac{M}{2r}\right)^{10} r^4 \sin^2 \theta. \quad (34.33)$$

Следовательно, эта метрика вырождена на сфере радиуса

$$r_* := \frac{M}{2}, \quad (34.34)$$

который соответствует горизонту $\rho_s = 2M$ в решении Шварцшильда, и на оси z ($\theta = 0, \pi$). Вырождение на оси z связано со сферической системой координат при $z \neq 0$.

Мы уже показали, что все геометрические инварианты для метрики Шварцшильда, построенные из тензора кривизны, не имеют особенностей на горизонте, и, следовательно, на сфере радиуса r_* в изотропных координатах.

Преобразование координат (34.31) показано на рис.34.6. Когда радиус r убывает от бесконечности до критического значения $r_* := M/2$, то радиальная координата Шварцшильда ρ убывает от ∞ до горизонта $\rho_s := 2M$. Затем, по мере убывания r от r_* до нуля, радиус ρ возрастает от ρ_s до ∞ . Таким образом два экземпляра метрики Шварцшильда вне горизонта $2M < \rho < \infty$ отображаются на две различные области в $\mathbb{R}^3$: $0 < r < r_*$ и $r_* < r < \infty$. На сфере $r = r_*$ они гладко сшиваются. При этом пространственная часть метрики невырождена. Заметим, что площадь поверхности сферы при приближении к началу координат стремится к бесконечности. Это связано с тем, что компоненты пространственной части метрики (34.32) расходятся при $r \rightarrow 0$.

Легко вычислить пространственную асимптотику нулевой компоненты метрики (34.32)

$$g_{00} \approx 1 - \frac{2M}{r}, \quad r \rightarrow \infty.$$

Она такая же, как и у метрики Шварцшильда в координатах Шварцшильда. Это неудивительно, т.к. $\rho \rightarrow r$ при $r \rightarrow \infty$. Это означает, что с законом тяготения Ньютона для метрики (34.32) все в порядке, поскольку он определяется асимптотикой при $r \rightarrow \infty$.

Метрика Шварцшильда в изотропных координатах (34.32) инвариантна относительно следующего преобразования радиальной координаты

$$r \mapsto r' := \frac{M^2}{4r} \quad \Leftrightarrow \quad \frac{2r'}{M} = \frac{M}{2r}. \quad (34.35)$$

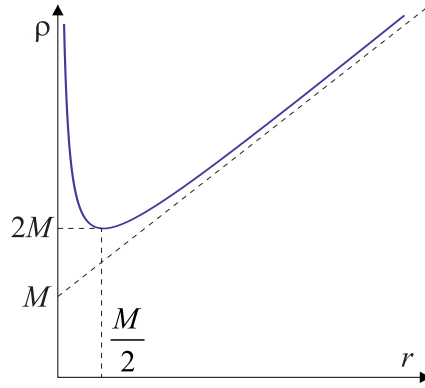


Рис. 34.6: Преобразование от координат Шварцшильда к изотропным координатам.

Это утверждение легко проверяется. При таком преобразовании координат точки пространства-времени отражаются относительно критической сферы. Сама критическая сфера $r_* = M/2$ остается неподвижной. Начало координат $r = 0$ отображается в бесконечно удаленную точку $r' = \infty$, и, наоборот, бесконечно удаленная точка $r = \infty$ отображается в начало координат $r' = 0$.

Заметим, что шварцшильдовская радиальная координата (34.31) при преобразовании (34.35) не меняется:

$$\rho = r \left(1 + \frac{M}{2r} \right)^2 = r' \left(1 + \frac{M}{2r'} \right)^2.$$

То есть одной и той же точке внешнего решения Шварцшильда соответствует две точки пространства-времени с метрикой (34.32): одна – внутри критической сферы, а другая – снаружи. Преобразование координат (34.35) меняет эти точки местами.

Метрика Шварцшильда в изотропных координатах глобально изометрична метрике моста Эйнштейна–Розена. Изометрия задается преобразованием

$$u = \sqrt{2r} \left(1 - \frac{M}{2r} \right), \quad r > 0. \quad (34.36)$$

При этом внешние решения Шварцшильда, соответствующее положительным и отрицательным значениям u , отображаются во внешность и внутренность сферы критического радиуса $r_* := M/2$:

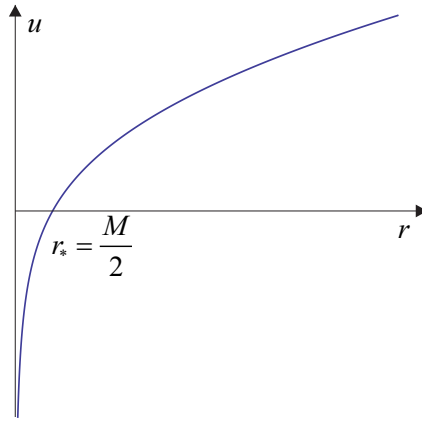
$$\begin{aligned} u &> 0, & r_* < r < \infty, \\ u &= 0, & r = r_*, \\ u &< 0, & 0 < r < r_*. \end{aligned}$$

Качественное поведение функции $u(r)$ показано на рис.34.7.

Инвариантности метрики Эйнштейна–Розена относительно инверсии $u \mapsto u' := -u$ соответствует инвариантность метрики Шварцшильда в изотропных координатах относительно преобразования радиальной координаты (34.35), поскольку справедлива формула

$$u = \sqrt{2r} \left(1 - \frac{M}{2r} \right) = -\sqrt{2r'} \left(1 - \frac{M}{2r'} \right),$$

в чем можно убедиться прямой проверкой.

Рис. 34.7: Качественное поведение функции $u(r)$.

34.6 Координаты Пенлеве–Гулстранда

Пусть в пространстве-времени задана метрика Шварцшильда в координатах Шварцшильда (34.1). Совершим преобразование временной координаты $(t, \rho, \theta, \varphi) \mapsto (T, \rho, \theta, \varphi)$, где функция $T = T(t, \rho)$ имеет вид [258, 259]

$$T := t \mp \left(4M \operatorname{arth} \sqrt{\frac{2M}{\rho}} - 2\sqrt{2M\rho} \right). \quad (34.37)$$

Отсюда вытекает связь дифференциалов

$$dT = dt \pm \frac{\sqrt{2M\rho}}{\rho - 2M} d\rho = dt \pm \frac{\sqrt{\frac{2M}{\rho}}}{1 - \frac{2M}{\rho}} d\rho. \quad (34.38)$$

Разрешив это соотношение относительно dt и подставив в метрику Шварцшильда (34.1), получим выражение для метрики в координатах Пенлеве–Гулстранда

$$\begin{aligned} ds^2 &= dT^2 - \left(d\rho \pm \sqrt{\frac{2M}{\rho}} dT \right)^2 - \rho^2 (d\theta^2 + \sin^2 \theta d\varphi^2) = \\ &= \left(1 - \frac{2M}{\rho} \right) dT^2 \mp 2\sqrt{\frac{2M}{\rho}} dT d\rho - d\rho^2 - \rho^2 (d\theta^2 + \sin^2 \theta d\varphi^2). \end{aligned} \quad (34.39)$$

Характерной особенностью данной метрики является то, что все сечения постоянного времени $T = \text{const}$ являются трехмерными евклидовыми пространствами $\mathbb{R}^3$. За эту наглядность пришлось заплатить наличием в метрике недиагонального слагаемого $dT d\rho$.

Преобразование координат (34.37) определено только вне горизонта, т.к. функция arth не определена при $\rho \in (0, 2M)$.

Матрица Якоби преобразования координат $(t, \rho) \mapsto (T, \rho)$ имеет вид

$$\frac{\partial(T, \rho)}{\partial(t, \rho)} = \begin{pmatrix} 1 & \pm \frac{\sqrt{2M\rho}}{\rho - 2M} \\ 0 & 1 \end{pmatrix}.$$

Отсюда следует, что якобиан преобразования координат (34.37) равен единице.

Если в выражении для метрики (34.39) перейти от сферических координат к декартовым ($\rho, \theta, \varphi \mapsto x, y, z$), то метрика будет определена всюду в $\mathbb{R}^4$ за исключением мировой линии начала координат $\rho = 0$. То есть она определена в большей области, чем преобразование координат (34.37) и покрывает не только область вне горизонта, $\rho \in (2M, \infty)$, но и под горизонтом, $\rho \in (0, 2M)$. При этом на горизонте $\rho = 2M$ она невырождена. Это значит, что метрика Шварцшильда в координатах Пенлеве–Гулстранда покрывает области I и II или III и IV на диаграмме Картера–Пенроуза аналогично координатам Эддингтона–Финкельштейна.

Действительно, под горизонтом $\rho \in (0, 2M)$ определено преобразование координат

$$T := t \mp \left(4M \operatorname{arth} \sqrt{\frac{\rho}{2M}} - 2\sqrt{2M\rho} \right). \quad (34.40)$$

Нетрудно проверить, что для дифференциалов преобразования координат мы получаем прежнее выражение (34.38).

Преобразования координат (34.37) и (34.40) определены, соответственно, вне и под горизонтом при всех $t \in \mathbb{R}$. При фиксированном значении радиуса ρ – это просто сдвиг временной координаты. Если зафиксировать t , например, положив $t = 0$, и выбрать знак минус в выражениях (34.37) и (34.40), тогда функция $T(t = 0, \rho)$ будет убывать от 0 до $-\infty$ при $\rho \in (0, 2M)$ и возрастать от $-\infty$ до ∞ при $\rho \in (2M, \infty)$. Качественное поведение этой функции показано на рис.34.8.

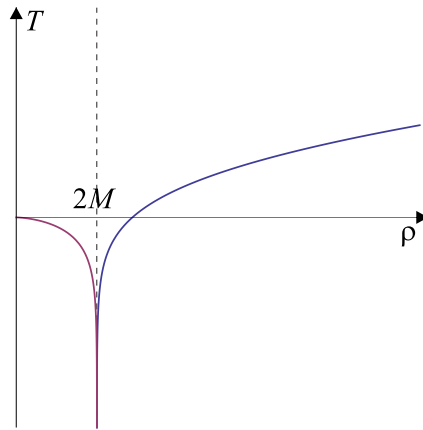


Рис. 34.8: Качественное поведение функции $T(t, \rho)$ при $t = 0$.

Метрика Пенлеве–Гулстранда (34.39) является асимптотически плоской на бесконечности $\rho \rightarrow \infty$. При этом компонента g_{00} имеет ту же асимптотику, что и в координатах Шварцшильда.

34.7 Гравитационное поле точечной частицы

Решение Шварцшильда описывает гравитационное поле вне компактного сферически симметричного распределения масс в общей теории относительности. В частности, оно описывает гравитационное поле вне точечной частицы. В литературе часто встречается утверждение, что метрика Шварцшильда в координатах Шварцшильда (34.1) описывает гравитационное поле точечной массивной частицы. Однако это утверждение является неточным. Метрика в координатах Шварцшильда имеет одну “физическую” особенность, сосредоточенную при $\rho = 0$. На плоскости t, ρ эта особенность представляет собой пространственноподобную линию, что никак не может

соответствовать точечной частице, мировая линия которой является времениподобной. Здесь возникает вопрос о том, что, с одной стороны, гравитационное поле вокруг точечной частицы должно описываться решением Шварцшильда, а с другой стороны неясно, где находится δ -функция, которая соответствует точечной частице.

В настоящем разделе мы найдем точное решение уравнений Эйнштейна с δ -образным источником, соответствующим точечной массивной частице. Оказывается, что этим решением является метрика Шварцшильда в изотропных координатах [260]. Решение Шварцшильда в изотропных координатах локально изометрично решению Шварцшильда в координатах Шварцшильда, однако существенно отличается глобально.

Рассмотрим четырехмерное топологически тривиальное многообразие $M \approx \mathbb{R}^4$ (пространство-время) с декартовой системой координат x^α , $\alpha = 0, 1, 2, 3$, и метрикой $g_{\alpha\beta}(x)$ лоренцевой сигнатуры $\text{sign } g_{\alpha\beta} = (+ - - -)$. Пусть мировая линия точечной частицы описывается функциями $\{q^\alpha(\tau)\}$, где τ – канонический параметр вдоль мировой линии.

В общей теории относительности точечная частица массы M описывается следующим действием

$$S = \frac{1}{16\pi} \int dx \sqrt{|g|} R - M \int d\tau \sqrt{\dot{q}^\alpha \dot{q}^\beta g_{\alpha\beta}}, \quad (34.41)$$

где $g := \det g_{\alpha\beta}$ и $\dot{q} := dq/d\tau$.

Вариация действия (34.41) по компонентам метрики $g_{\alpha\beta}$ и координатам частицы q^α приводит к уравнениям Эйнштейна и уравнениям для экстремалей (геодезических) (см. раздел 23.9):

$$R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R = -\frac{1}{2} T^{\alpha\beta}, \quad (34.42)$$

$$\left(\ddot{q}^\alpha + \Gamma_{\beta\gamma}^\alpha|_{x=q} \dot{q}^\beta \dot{q}^\gamma \right) g_{\alpha\delta} = 0, \quad (34.43)$$

где

$$T^{\alpha\beta} := \frac{16\pi M \dot{q}^\alpha \dot{q}^\beta}{\sqrt{|g|} \dot{q}^0} \delta(\mathbf{x} - \mathbf{q}) \quad (34.44)$$

– тензор энергии-импульса, $\dot{q}_\alpha := \dot{q}^\beta g_{\beta\alpha}$ и

$$\delta(\mathbf{x} - \mathbf{q}) := \delta(x^1 - q^1) \delta(x^2 - q^2) \delta(x^3 - q^3)$$

– трехмерная δ -функция на сечениях $x^0 = \text{const}$. Мы предполагаем, что для мировой линии частицы выполнено неравенство $\dot{q}^0 \neq 0$ (иначе мировая линия не будет времениподобной).

Систему уравнений движения (34.42)–(34.43) мы будем решать в гамильтоновой форме. Для этой цели используем АДМ параметризацию метрики (см. раздел 24.2)

$$g_{\alpha\beta} = \begin{pmatrix} N^2 + N^\rho N_\rho & N_\nu \\ N_\mu & g_{\mu\nu} \end{pmatrix}, \quad (34.45)$$

где N, N_μ – функции хода и сдвига. Выше $g_{\mu\nu}$ – метрика на пространственных сечениях $x^0 = \text{const}$ и $N^\rho := \hat{g}^{\rho\mu} N_\mu$, где $\hat{g}^{\rho\mu}$ – 3×3 матрица, обратная к $g_{\mu\nu}$:

$$\hat{g}^{\rho\mu} g_{\mu\nu} = \delta_\nu^\rho.$$

В дальнейшем подъем и опускание пространственных индексов, которые занумерованы греческими буквами из середины алфавита $\mu, \nu, \dots = 1, 2, 3$, проводятся с помощью метрики $g_{\mu\nu}$ (отрицательно определенной) и ее обратной $\hat{g}^{\mu\nu}$.

Переход от лагранжевого действия (34.41) к гамильтоновой форме является трудоемким. К счастью гамильтонова формулировка общей теории относительности была подробно проделана в главе 24, а гамильтониан точечной частицы, движущейся в псевдоримановом многообразии, построен в разделе 22.8. Обозначим через $p^{\mu\nu}$ и p_α импульсы, сопряженные обобщенным координатам $g_{\mu\nu}$ и q^α . Поскольку кинетические слагаемые в действии для точечной частицы в гравитационном поле (34.41) не перемешаны, то выражения для импульсов через обобщенные координаты и скорости остаются прежними (22.71), (24.59), и действия в гамильтоновой форме (22.86), (24.70) нужно просто сложить. В результате действие (34.41) для точечной частицы в общей теории относительности в гамильтоновой форме примет вид

$$S = \frac{1}{16\pi} \int dx (p^{\mu\nu} \dot{g}_{\mu\nu} - H_{\text{ADM}}) + \int dt (p_\mu \dot{q}^\mu - N \sqrt{\hat{p}^2 + M^2} + N^\mu p_\mu), \quad (34.46)$$

где гамильтонова плотность H_{ADM} определена формулой (24.65), $\hat{p}^2 := -\hat{g}^{\mu\nu} p_\mu p_\nu \geq 0$ – неотрицательный квадрат пространственного импульса частицы, и мы использовали временную калибровку для точечной частицы $q^0 = x^0 = t$ для упрощения последующих вычислений. Заметим, что параметр t вдоль мировой линии частицы не является каноническим. Вторую часть действия можно записать в виде интеграла по объему, вставив четырехмерную δ -функцию и один раз проинтегрировав, как это было сделано в разделе 23.9. В результате суммарное действие в гамильтоновой форме примет вид

$$S = \frac{1}{16\pi} \int dx (p^{\mu\nu} \dot{g}_{\mu\nu} + 16\pi p_\mu \dot{q}^\mu - N H_\perp - N^\mu H_\mu), \quad (34.47)$$

где

$$H_\perp = \frac{1}{\hat{e}} (p^{\mu\nu} p_{\mu\nu} - p^2) - \hat{e} \hat{R} + 16\pi \sqrt{M^2 + \hat{p}^2} \delta(\mathbf{x} - \mathbf{q}), \quad (34.48)$$

$$H_\mu = -2\hat{\nabla}_\nu p^\nu{}_\mu - 16\pi p_\mu \delta(\mathbf{x} - \mathbf{q}), \quad (34.49)$$

где $p := p^{\mu\nu} g_{\mu\nu}$ – след пространственных компонент импульсов и $p^\mu := \hat{g}^{\mu\nu} p_\nu$. Кроме того приняты следующие обозначения: $\hat{e} := \sqrt{|\det g_{\mu\nu}|}$, $\hat{R}$ – трехмерная скалярная кривизна для метрики $g_{\mu\nu}$ и $\hat{\nabla}_\mu$ – трехмерная ковариантная производная. В этом действии динамическими переменными являются пространственные компоненты метрики $g_{\mu\nu}$, $p^{\mu\nu}$ и пространственные координаты частицы q^μ , p_μ . Функции хода N и сдвига N^μ представляют собой множители Лагранжа.

Вариация действия (34.47) по множителям Лагранжа дает четыре связи:

$$H_\perp = 0, \quad H_\mu = 0. \quad (34.50)$$

Эти связи являются связями первого рода и соответствуют инвариантности действия (34.47) относительно общих преобразований координат. Вариация гамильтонова действия по компонентам метрики и сопряженным импульсам дает гамильтоновы уравнения движения для пространственных компонент метрики:

$$\dot{g}_{\mu\nu} = \frac{2N}{\hat{e}} p_{\mu\nu} - \frac{N}{\hat{e}} g_{\mu\nu} p + \hat{\nabla}_\mu N_\nu + \hat{\nabla}_\nu N_\mu, \quad (34.51)$$

$$\begin{aligned} \dot{p}^{\mu\nu} = & \frac{N}{2\hat{e}} \hat{g}^{\mu\nu} \left(p^{\rho\sigma} p_{\rho\sigma} - \frac{1}{2} p^2 \right) - \frac{2N}{\hat{e}} \left(p^{\mu\rho} p^\nu{}_\rho - \frac{1}{2} p^{\mu\nu} p \right) + \hat{e} (\hat{\Delta} N \hat{g}^{\mu\nu} - \hat{\nabla}^\mu \hat{\nabla}^\nu N) - \\ & - \hat{e} N \left(\hat{R}^{\mu\nu} - \frac{1}{2} g^{\mu\nu} \hat{R} \right) - p^{\mu\rho} \hat{\nabla}_\rho N^\nu - p^{\nu\rho} \hat{\nabla}_\rho N^\mu + \hat{\nabla}_\rho (N^\rho p^{\mu\nu}) - \frac{8\pi N p^\mu p^\nu}{\sqrt{M^2 + \hat{p}^2}} \delta(\mathbf{x} - \mathbf{q}), \end{aligned} \quad (34.52)$$

где $\hat{\Delta} := \hat{\nabla}^\mu \hat{\nabla}_\mu$ – трехмерный оператор Лапласа–Бельтрами. Гамильтоновы уравнения для геодезических имеют вид

$$\dot{q}^\mu = - \frac{N}{\sqrt{M^2 + \hat{p}^2}} \Big|_{x=q} p^\mu - N^\mu|_{x=q}, \quad (34.53)$$

$$\dot{p}_\mu = -\partial_\mu \left[N \sqrt{M^2 + \hat{p}^2} - N^\nu p_\nu \right]_{x=q}. \quad (34.54)$$

Теперь необходимо решить полную систему уравнений (34.50)–(34.54) для покоящейся частицы. Предположим, что частица покоится в начале системы координат:

$$q^0 = x^0 = t, \quad q^\mu = 0, \quad \mu = 1, 2, 3.$$

Мы ищем статическое сферически симметричное решение уравнений движения (34.50)–(34.54). В этом случае систему координат можно выбрать таким образом, что функции сдвига равны нулю, $N_\mu = 0$. Мы также выберем вейлевски плоскую калибровку для пространственной части метрики,

$$g_{\mu\nu} = -f^2 \delta_{\mu\nu}, \quad (34.55)$$

где $\delta_{\mu\nu} := \text{diag}(+, +, +)$ – плоская евклидова метрика и

$$f = f(r), \quad r := \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2},$$

– некоторая неизвестная функция от радиуса r . Для статических решений все импульсы обращаются в нуль: $p^{\mu\nu} = 0$ и $p_\mu = 0$. Производные по времени $\dot{g}_{\mu\nu} = 0$ также равны нулю. Поэтому в рассматриваемом случае у нас есть только две неизвестные функции $N(r)$ и $f(r)$, зависящие от радиуса.

Если сделанные предположения верны, то уравнения $H_\mu = 0$, (34.51) и (34.53) тождественно удовлетворяются, а уравнения $H_\perp = 0$, (34.52) и (34.54) принимают вид

$$-\hat{e}\hat{R} + 16\pi M\delta(\mathbf{x}) = 0, \quad (34.56)$$

$$\hat{e} \left(\hat{\Delta} N \hat{g}^{\mu\nu} - \hat{\nabla}^\mu \hat{\nabla}^\nu N \right) - \hat{e} N \left(\hat{R}^{\mu\nu} - \frac{1}{2} g^{\mu\nu} \hat{R} \right) = 0, \quad (34.57)$$

$$\partial_\mu N|_{x=0} = 0, \quad (34.58)$$

где мы разделили уравнение (34.54) на M .

Важно отметить, что уравнение (34.57) нельзя разделить на $\hat{e}$ и индексы μ, ν нельзя опустить, т.к. метрика имеет особенность при $r = 0$, а тензор Риччи и скалярная кривизна содержат δ -функцию.

Основное уравнение

Уравнение (34.56) является основным уравнением в рассматриваемой задаче, которое необходимо решить. Запишем его в виде

$$\hat{e}\hat{R} = 16\pi M\delta(\mathbf{x}). \quad (34.59)$$

Трехмерный элемент объема $\hat{e}$ в левой части необходим для ковариантности, потому что δ -функция представляет собой скалярную плотность относительно преобразования координат.

Это уравнение является простейшим ковариантным дифференциальным уравнением второго порядка с δ -функцией в правой части. В двумерном случае в конформной калибровке уравнение (34.59) сводится к уравнению Пуассона и дает фундаментальное решение двумерного уравнения Лапласа (см. раздел ??). В большем числе измерений оно становится нелинейным, и мы предъявим его решение в трехмерном случае. Вполне вероятно, что данное уравнение будет иметь приложения в дифференциальной геометрии само по себе.

Будем искать сферически симметричное решение уравнения (34.59), что согласуется с симметрией правой части. Каждая сферически симметричная трехмерная метрика является вейлевски плоской, что означает существование системы координат, в которой метрика имеет вид (34.55). Тогда основное уравнение (34.59) для вейлевски плоской сферически симметричной метрики примет вид

$$\Delta f - \frac{\partial f^2}{2f} = -4\pi M \delta(\mathbf{x}), \quad (34.60)$$

где $\Delta := \partial_1^2 + \partial_2^2 + \partial_3^2$ – плоский оператор Лапласа, и введено обозначение $\partial f^2 := \delta^{\mu\nu} \partial_\mu f \partial_\nu f$.

Уравнение (34.60) нелинейно из-за второго слагаемого в левой части. Решение этого уравнения необходимо понимать в слабом смысле после интегрирования с пробной функцией (см, например, [3]), т.к. в правой части стоит δ -функция. Если искать решение в классе функционалов (обобщенных функций) $\mathcal{D}'(\mathbb{R}^3)$, то возникают две серьезные проблемы. Во-первых, поскольку уравнение нелинейно, то обобщенные функции необходимо умножать, а умножение в $\mathcal{D}'$ отсутствует. Во-вторых, в дальнейшем мы увидим, что некоторые слагаемые в уравнении (34.60) не являются локально интегрируемыми функциями, и поэтому необходима регуляризация возникающих интегралов. Хотя умножение функционалов в $\mathcal{D}'$ нельзя определить в общем случае, для некоторых функционалов оно все таки хорошо определено, например, если они заданы гладкими функциями. Если функции к тому же не обращаются в нуль, то для таких функций левая часть уравнения (34.60) имеет хорошо определенный математический смысл.

Будем искать сферически симметричное решение уравнения (34.60) в сферической системе координат r, θ, φ .

Теорема 34.7.1. Функция

$$f := 1 + \frac{M}{r} + \frac{M^2}{4r^2} = \left(1 + \frac{M}{2r}\right)^2 \quad (34.61)$$

удовлетворяет уравнению (34.60), чьи решения понимаются в слабом смысле после интегрирования с пробными функциями.

Доказательство. Используем аналитическую регуляризацию обобщенной функции r^λ , $\lambda \in \mathbb{C}$ [261]. Для $\operatorname{re} \lambda > -3$ функция r^λ является локально интегрируемой, и функционал

$$(r^\lambda, \varphi) := \int_{\mathbb{R}^3} dx r^\lambda \varphi$$

для всех пробных функций из $\varphi \in \mathcal{D}(\mathbb{R}^3)$ (гладкие функции на $\mathbb{R}^3$ с компактным носителем). Этот функционал можно аналитически продолжить на всю комплексную плоскость $\lambda \in \mathbb{C}$, за исключением точек на вещественной оси $\lambda = -3, -5, -7, \dots$, где

он имеет простые полюса [261], глава I, §3, раздел 9. В рассматриваемом случае полюса не являются препятствием, т.к. мы в них не попадем.

Рассмотрим функционал

$$f_\nu := \left(1 + \frac{M}{2}r^\nu\right)^2, \quad \text{re } \nu \geq 1, \quad \in \mathcal{D}'(\mathbb{R}^3),$$

зависящий от параметра $\nu \in \mathbb{C}$. Простые вычисления показывают, что выполнено следующее равенство

$$\frac{\partial f_\nu^2}{2f_\nu} = \frac{f_\nu'^2}{2f_\nu} = \frac{M^2}{2}\nu^2 r^{2\nu-2}, \quad (34.62)$$

где штрих обозначает дифференцирование по радиусу. Это равенство заведомо выполнено при $\text{re } \nu \geq 1$. Его можно аналитически продолжить до точки $\nu = -1$, поскольку мы не попадаем в полюс, упомянутый выше. В результате получаем следующее выражение для второго слагаемого в левой части уравнения (34.60)

$$\frac{\partial f^2}{2f} = \frac{M^2}{2r^4}. \quad (34.63)$$

Это равенство понимается в слабом смысле после интегрирования с пробными функциями. Поэтому уравнение (34.60) для обобщенной функции (34.61) принимает вид

$$\Delta \left(1 + \frac{M}{r} + \frac{M^2}{4r^2}\right) - \frac{M^2}{2r^4} = -4\pi M\delta(\mathbf{x}).$$

Функция M/r является единственным фундаментальным решением уравнения Лапласа, которое убывает на бесконечности (см., например, [3]):

$$\Delta \left(1 + \frac{M}{r}\right) = \Delta \frac{M}{r} = -4\pi M\delta(\mathbf{x}). \quad (34.64)$$

Функция $M^2/2r^4$ не является локально интегрируемой и требует регуляризации. Мы понимаем ее как аналитическое продолжение функционала r^λ в точку $\lambda = -4$, которое было упомянуто в начале.

Равенство

$$\Delta r^\mu = \frac{1}{r^2} \partial_r (r^2 \partial_r r^\mu) = \mu(\mu + 1)r^{\mu-2}$$

просто проверяется для $\text{re } \mu > 2$. Оно остается также в силе после аналитического продолжения в точку $\mu = -2$. При $\mu = -2$ оно выглядит следующим образом

$$\Delta \frac{1}{r^2} = 2 \frac{1}{r^4}.$$

Таким образом, после аналитического продолжения мы ставим в соответствие локально неинтегрируемой функции $1/r^4$ хорошо определенный функционал

$$\left(\frac{1}{r^4}, \varphi\right) := \frac{1}{2} \left(\Delta \frac{1}{r^2}, \varphi\right) = \frac{1}{2} \left(\frac{1}{r^2}, \Delta \varphi\right).$$

Отсюда следует, что нелинейное слагаемое в левой части уравнения (34.60) сокращается после интегрирования с пробной функцией

$$\left(\Delta \frac{M^2}{4r^2} - \frac{M^2}{2r^4}, \varphi\right) = 0, \quad (34.65)$$

если для функции $M^2/2r^4$ использована аналитическая регуляризация. $\square$

Доказательство теоремы можно провести и в других переменных. Пусть $f := \chi^2$. Тогда уравнение (34.60) перепишется в эквивалентном виде

$$\chi \Delta \chi = -2\pi M \delta(\mathbf{x}). \quad (34.66)$$

Положим

$$\chi = 1 + \frac{M}{2r}. \quad (34.67)$$

Подстановка этого решения в уравнение (34.66) приводит к равенствам:

$$\left(1 + \frac{M}{2r}\right) \Delta \left(1 + \frac{M}{2r}\right) = \left(1 + \frac{M}{2r}\right) (-2\pi M) \delta(\mathbf{x}) = -2\pi M \delta(\mathbf{x}). \quad (34.68)$$

Выражение $\frac{1}{r} \delta(\mathbf{x})$ не определено, и мы придадим ему математический смысл следующим образом. Равенство

$$\Delta \frac{r^\lambda}{r} = \lambda(\lambda - 1) r^{\lambda-3} + r^\lambda \Delta \frac{1}{r}$$

легко проверяется при $\operatorname{Re} \lambda > 3$. Продолжим его аналитически в точку $\lambda = -1$, где его можно переписать в виде

$$\frac{1}{r} \Delta \frac{1}{r} = \Delta \frac{1}{r^2} - \frac{2}{r^4}.$$

Поэтому мы определим функционал $\frac{1}{r} \delta(\mathbf{x})$ следующими равенствами:

$$\left(\frac{1}{r} \delta(\mathbf{x}), \varphi\right) := -\frac{1}{4\pi} \left(\frac{1}{r} \Delta \frac{1}{r}, \varphi\right) := -\frac{1}{4\pi} \left(\Delta \frac{1}{r^2} - \frac{2}{r^4}, \varphi\right).$$

Для аналитически регуляризованного функционала $\frac{1}{r^4}$ правая часть равна нулю, и мы получаем равенство

$$\frac{1}{r} \delta(\mathbf{x}) = 0,$$

которое надо понимать в слабом смысле. Другими словами, равенство $r^\lambda \delta(\mathbf{x}) = 0$, которое выполнено при $\lambda > 0$, аналитически продолжается в точку $\lambda = -1$.

Таким образом, решение (34.61) можно получить либо решая уравнение (34.60), либо (34.66). В обоих случаях использована аналитическая регуляризация функционалов. Подчеркнем, что полученное решение явно зависит от массы M .

Функция (34.61) является локально интегрируемой и принадлежит пространству обобщенных функций $\mathcal{D}'(\mathbb{R}^3)$. Она удовлетворяет уравнению (34.60) в слабом смысле, и левая часть уравнения для нее хорошо определена. К сожалению мы не можем описать то подпространство обобщенных функций в $\mathcal{D}'(\mathbb{R}^3)$, для которого левая часть основного уравнения (34.59) является математически определенной в общем случае. Поэтому вопрос о единственности полученного решения остается без ответа.

Оставшиеся уравнения

Теперь осталось решить уравнения (34.57) и (34.58). Как и для основного уравнения (34.59) мы не можем в общем случае описать то подпространство в $\mathcal{D}'(\mathbb{R}^3)$, для которого нелинейная часть уравнения (34.57) определена. Вместо этого мы предъядим функцию в явном виде и проверим, что для нее все слагаемые определены и равенство выполнено в слабом смысле.

Теорема 34.7.2. Пусть трехмерная метрика имеет вид (34.55), где конформный множитель определен равенством (34.61). Тогда функция хода

$$N := \frac{1 - \frac{M}{2r}}{1 + \frac{M}{2r}}. \quad (34.69)$$

удовлетворяет уравнению (34.57) в $\mathbb{R}^3$ в слабом смысле.

Доказательство. Функция хода положительна при $r > M/2$ и отрицательна при $0 < r < M/2$. Легко проверить, что она имеет конечный предел в нуле,

$$\lim_{r \rightarrow 0} N = -1,$$

и бесконечно дифференцируема при $r > 0$. Поэтому функция хода локально интегрируема и, следовательно, лежит в $\mathcal{D}'(\mathbb{R}^3)$.

Нетрудно видеть, что лапласиан от функции хода равен нулю при $r > 0$,

$$\hat{\Delta} N = \frac{1}{\hat{e}} \partial_\mu (\hat{e} \hat{g}^{\mu\nu} \partial_\nu N) = 0.$$

Функция N является непрерывной, но ее производные в нуле имеют скачки. Поскольку $N \in \mathcal{D}'(\mathbb{R}^3)$, то ее слабые производные определены всюду, и равенство

$$\hat{\Delta} N = 0,$$

также определено всюду в $\mathbb{R}^3$, т.к. функция хода имеет конечный предел в нуле. Это равенство можно умножить на $\hat{e} g^{\mu\nu}$, поскольку вблизи нуля справедлива следующая асимптотика

$$\hat{e} \sim r^{-4}, \quad g^{\mu\nu} \sim r^4 \delta^{\mu\nu}. \quad (34.70)$$

Поэтому первое слагаемое в левой части уравнения (34.57) равно нулю всюду в $\mathbb{R}^3$ в слабом смысле.

Можно проверить, что при $r > 0$ справедливо следующее равенство

$$\hat{e} \hat{\nabla}^\mu \hat{\nabla}^\nu N + \hat{e} N \left(\hat{R}^{\mu\nu} - \frac{1}{2} g^{\mu\nu} \hat{R} \right) = 0.$$

Поэтому у нас остались только δ -функции, которые содержатся в тензоре Риччи и скалярной кривизне. Они не сокращаются, и уравнение (34.57) принимает вид

$$\frac{1}{4} g^{\mu\nu} N \hat{e} \hat{R} = 4\pi M g^{\mu\nu} N \delta(\mathbf{x}) = 0.$$

Это уравнение удовлетворено, поскольку вблизи нуля обратная метрика имеет асимптотику

$$g^{\mu\nu} N \sim \delta^{\mu\nu} r^4, \quad \text{и} \quad r^4 \delta(\mathbf{x}) = 0.$$

Таким образом уравнение (34.57) выполнено в слабом смысле. $\square$

Теперь осталось решить только уравнение (34.58) для мировой линии точечной частицы. Так как

$$\partial_r N = \frac{M}{r^2} \frac{1}{\left(1 + \frac{M}{2r}\right)^2},$$

то оно не удовлетворяется. На этом этапе мы привлекаем физический аргумент. Правая часть уравнения (34.54) представляет собой силу, которая действует на частицу со стороны собственного гравитационного поля. Поэтому будем понимать уравнение (34.54) после усреднения по сфере, окружающей частицу. Тогда оно выполнено ввиду наличия сферической симметрии. Другими словами, точечная частица не движется под действием собственного гравитационного поля.

Аналогичная ситуация возникает в классической механике и электродинамике. Например, в гравитации Ньютона потенциал имеет вид $\varphi = -1/r$, и сила, действующая на частицу со стороны собственного гравитационного поля равна

$$\partial_\mu \frac{1}{r} = \frac{x_\mu}{r^3}.$$

Это выражение не определено при $r = 0$, и, следовательно, уравнения движения не выполнены. Однако они будут удовлетворены после усреднения по сфере.

Объединяя теоремы 34.7.1 и 34.7.2 получаем метрику

$$ds^2 = \left(\frac{1 - \frac{M}{2r}}{1 + \frac{M}{2r}} \right)^2 dt^2 - \left(1 + \frac{M}{2r} \right)^4 [dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)]. \quad (34.71)$$

Это есть метрика Шварцшильда в изотропных координатах, которая была подробно рассмотрена в разделе 34.5. Таким образом, доказано, что метрика (34.71) удовлетворяет системе уравнений Эйнштейна с δ -образным источником, соответствующим точечной частице. Тем самым найдено гравитационное поле точечной массивной частицы в рамках общей теории относительности.

Метрика Шварцшильда в изотропных координатах определена всюду в $\mathbb{R}^4$ за исключением мировой линии частицы, покоящейся в начале координат. Тем самым топология пространства-времени тривиальна $M \approx \mathbb{R}^4$.

Хотя метрика Шварцшильда в изотропных координатах давно и хорошо известна, в настоящем разделе ей дана существенно новая интерпретация. Постоянная M в решении Шварцшильда является постоянной интегрирования и в действии не содержится. Ей придается физический смысл массы только после рассмотрения асимптотики на больших расстояниях $r \rightarrow \infty$ и сравнения с законом Ньютона (см. раздел 23.9.1). При рассмотрении точечной частицы в общей теории относительности постоянная M входит непосредственно в действие (34.41) и поэтому имеет смысл массы частицы с самого начала. То, что в конце концов получилась одна и та же метрика свидетельствует о самосогласованности теории. Кроме того, множитель $1/(16\pi)$ перед действием Гильберта–Эйнштейна также играет двоякую роль. В разделе 24.8.4 было показано, что этот множитель необходим для того, чтобы полная энергия гравитационного поля, определенная через поверхностный интеграл на пространственной бесконечности, была равна Mc^2 . В настоящем разделе этот множитель привел к “правильному” коэффициенту в основном уравнении (34.59), которое воспроизвело решение Шварцшильда.

Метрика Шварцшильда в изотропных координатах имеет ту же асимптотику на пространственной бесконечности, что и метрика Шварцшильда в координатах Шварцшильда. Поэтому с законом всемирного тяготения проблем нет. При $r \rightarrow \infty$ метрика, которую создает точечная частица, стремится к метрике Лоренца и, следовательно, является геодезически полной. Существенно новое поведение возникает вблизи частицы. Начало координат $r = 0$ соответствует бесконечному значению шварцшильдовской радиальной координате ρ (см. рис.34.6). Это значит, что

пространство-время в непосредственной близости от частицы является также асимптотически плоским и геодезически полным. Хотя метрика и имеет особенность в начале координат, полный тензор кривизны $R_{\alpha\beta\gamma\delta}$ стремится к нулю при $r \rightarrow \infty$ (асимптотическая свобода).

Точечная частица, которая создает метрику (34.71), обладает необычными свойствами по сравнению с классической механикой, т.к. компоненты метрики расходятся на мировой линии частицы. Например, площадь сферы $t = \text{const}$, $r = \text{const}$ равна

$$S_r = 4\pi r^2 \left(1 + \frac{M}{4r}\right)^4. \quad (34.72)$$

На больших расстояниях площадь сферы приближается к евклидову значению $4\pi r^2$. Однако вблизи частицы площадь сферы расходится:

$$\lim_{r \rightarrow 0} S_r = \frac{M^4}{64r^2} \rightarrow \infty.$$

Кроме того, длина радиальной пространственноподобной линии $t, \theta, \varphi = \text{const}$ от произвольной точки $r_0 > 0$ до начала координат (расстояние до точечной частицы) равна бесконечности, поскольку интеграл вблизи нуля

$$\int_0 ds = \int_0 dr \left(1 + \frac{M}{2r}\right)^2 \rightarrow \int_0 dr \frac{M^2}{4r^2} \rightarrow \infty$$

расходится. Эти свойства метрики весьма необычны с точки зрения точечной частицы, движущейся в евклидовом пространстве. Тем не менее метрику (34.71) нельзя отбросить как нефизическую, т.к. это единственное известное в настоящее время точное решение уравнений Эйнштейна с источником в виде точечной частицы.

Рассмотрим движение пробных частиц в гравитационном поле точечной частицы массы M , т.е. геодезические (экстремали) в пространстве-времени с метрикой (34.71). Область пространства-времени на больших расстояниях $M/2 < r < \infty$ изометрична внешнему решению Шварцшильда и поэтому мы имеем гравитационное притяжение (ускорение направлено к горизонту), которое на больших расстояниях совпадает с законом всемирного тяготения. Область пространства-времени вблизи точечной частицы $0 < M < M/2$ также изометрична внешнему решению Шварцшильда, но вывернутому наизнанку. То есть гравитационное притяжение вблизи частицы меняется на отталкивание. При этом отталкивание становится тем сильнее, чем ближе пробная частица приближается к точечной массе. Поскольку геодезические полны при $r \rightarrow 0$, то никакая пробная частица не может упасть на точечную массу.

С математической точки зрения поведение геодезических при $r \rightarrow 0$ и $r \rightarrow \infty$ не вызывает никаких трудностей. Проблема возникает на критической сфере радиуса $r_* = M/2$, на которой сшиваются два внешних решения Шварцшильда вдоль горизонта. В следующем разделе будет показано, что с точки зрения внешнего наблюдателя пробные частицы приближаются к критической сфере за бесконечное наблюдаемое время, но при конечном значении канонического параметра. То есть они не полны, и их необходимо продолжить. Заметим, что все геометрические инварианты на критической сфере r_* конечны, поскольку она соответствует горизонту. Продолжение геодезических за горизонт для решения Шварцшильда обсуждалось в координатах Крускала–Секереша в разделе 34.3. Однако при таком продолжении необходимо приклеить внутреннее решение Шварцшильда, а его в метрике (34.71) просто нет. Возможный выход из этой ситуации предложен в следующем разделе.

34.8 Проход по мосту Эйнштейна–Розена

В предыдущем разделе мы установили, что метрика, которую создает вокруг себя точечная частица массы M в общей теории относительности, является метрикой Шварцшильда в изотропных координатах (34.71). Пространство-время с такой метрикой обладает существенным недостатком: оно геодезически неполно на критической сфере радиуса $r_* := M/2$. Если для продолжения решения приклеить внутреннее решение Шварцшильда, то мы получим максимальное продолжение Крускала–Секереша (см., раздел 34.3), описывающее черную и белые дыры. Однако в этом случае метрика перестанет удовлетворять уравнениям Эйнштейна с точечным источником, чего хотелось бы избежать. В настоящем разделе предложен другой выход из создавшейся ситуации. А именно, мы ослабим требование геодезической полноты.

Напомним, что уравнения геодезической $q^\alpha(\tau)$, $\alpha = 0, 1, 2, 3$, в пространстве-времени $(\mathbb{M}, g)$ описывают три физические степени свободы q^μ , $\mu = 1, 2, 3$, которые соответствуют пространственным значениям индексов, и одну нефизическую степень свободы q^0 (см., раздел 22.8). Потребуем, чтобы пространство-время было полным только для физических степеней свободы, т.е. чтобы решения $q^\mu(\tau)$ уравнений геодезических были достаточно гладкими и определены при всех значениях $\tau \in \mathbb{R}$, где τ – канонический параметр вдоль геодезических. Это требования является более слабым, чем требование полноты самих геодезических, т.к. у нефизической степени свободы $q^0(\tau)$ допускается наличие особенностей. Ниже будет показано, что метрика Шварцшильда в изотропных координатах является полной в этом более слабом смысле.

В разделе 34.5 было показано, что пространство-время с метрикой Шварцшильда в изотропных координатах глобально изометрично мосту Эйнштейна–Розена. Это означает, что геодезические линии для обеих метрик связаны преобразованием координат (34.36). Мы будем исследовать геодезические для метрики Эйнштейна–Розена, потому что картинка симметрична относительно преобразования $u \mapsto -u$. Кроме того, попутно предложено решение еще одной проблемы, которая стоит в общей теории относительности много лет: может ли наблюдатель пройти из одной вселенной в другую, если мост Эйнштейна–Розена интерпретировать как кротовую нору? Ответ на этот вопрос положителен: пробные частицы свободно проходят по мосту Эйнштейна–Розена в обоих направлениях, если требовать полноту геодезических только для физических степеней свободы. При этом все особенности сдвигаются в нефизический сектор.

Метрика Эйнштейна–Розена имеет вид (34.25). Для простоты мы рассмотрим только “радиальные” геодезические $t(\tau), u(\tau)$ ($\theta, \varphi = \text{const}$), где τ – канонический параметр. Поскольку метрика диагональна, то для исследования радиальных геодезических достаточно рассмотреть плоскость с координатами $t, u \in \mathbb{R}^2$, на которой задана двумерная статическая метрика лоренцевой сигнатуры

$$g_{\alpha\beta} = \begin{pmatrix} \frac{u^2}{u^2 + 4M} & 0 \\ 0 & -(u^2 + 4M) \end{pmatrix}, \quad g^{\alpha\beta} = \begin{pmatrix} \frac{u^2 + 4M}{u^2} & 0 \\ 0 & -\frac{1}{(u^2 + 4M)} \end{pmatrix}, \quad (34.73)$$

где $\{x^\alpha\} = \{t, u\}$. Простые вычисления показывают, что только четыре символа Кристоффеля отличны от нуля:

$$\Gamma_{tt}^u = \frac{4Mu}{(u^2 + 4M)^3}, \quad \Gamma_{tu}^t = \Gamma_{ut}^t = \frac{4M}{u(u^2 + 4M)}, \quad \Gamma_{uu}^u = \frac{u}{u^2 + 4M}.$$

Тогда уравнения геодезических (16.3) принимают вид

$$\begin{aligned}\ddot{t} &= -\frac{8M}{u(u^2 + 4M)}\dot{t}\dot{u}, \\ \ddot{u} &= -\frac{4Mu}{(u^2 + 4M)^3}\dot{t}^2 - \frac{u}{u^2 + 4M}\dot{u}^2.\end{aligned}$$

Эти уравнения имеют два первых интеграла:

$$C_0 = g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = \frac{u^2}{u^2 + 4M}\dot{t}^2 - (u^2 + 4M)\dot{u}^2, \quad (34.74)$$

$$C_1 = K_\alpha\dot{x}^\alpha = \frac{u^2}{u^2 + 4M}\dot{t}. \quad (34.75)$$

Первый интеграл движения – это сохранение длины касательного вектора. Вторым интегралом возникает благодаря наличию векторного поля Киллинга $K = \partial_t$ (статичность решения). Теперь нетрудно получить уравнение, определяющее форму радиальных геодезических

$$\frac{dt}{du} = \pm \frac{u^2 + 4M}{u\sqrt{1 - C\frac{u^2}{u^2 + 4M}}}, \quad (34.76)$$

где $C := C_0/C_1^2$. Так как постоянные C_0 и C пропорциональны и их знаки совпадают, то времениподобные, нулевые и пространственноподобные геодезические описываются, соответственно, уравнением (34.76) с постоянными $C > 0$, $C = 0$ и $C < 0$. Вблизи вертикальной прямой $u = 0$ решение этого уравнения дается логарифмически расходящимся интегралом

$$t \sim \int_0 \frac{du}{u}$$

не зависимо от типа геодезической. Это значит, что никакая геодезическая, в том числе пространственноподобная, не пересекает линию $u = 0$. В частности, пробная частица ($C > 0$) достигает прямую $u = 0$ при бесконечном значении наблюдаемого времени t .

В то же время уравнение (34.74) приводит к равенству

$$C_0 = \left[\frac{u^2}{u^2 + 4M} \left(\frac{dt}{du} \right)^2 - (u^2 + 4M) \right] \dot{u}^2.$$

Подстановка производной (34.76) дает следующее уравнение вблизи $u = 0$ для ненулевых геодезических

$$C_0 = \left[\frac{u^2 + 4M}{1 - C\frac{u^2}{u^2 + 4M}} - (u^2 + 4M) \right] \dot{u}^2 \approx Cu^2\dot{u}^2.$$

Его решение дается сходящимся интегралом

$$\tau \sim \int_0 duu. \quad (34.77)$$

Это означает, что все ненулевые радиальные геодезические сингулярны в точке $u = 0$: они уходят в бесконечность $t \rightarrow \pm\infty$ с обеих сторон $u > 0$ и $u < 0$ при конечном значении канонического параметра.

Для нулевых геодезических ситуация такая же. Если $C_0 = 0$, то первое уравнение (34.74) дает

$$\dot{t} = \pm \frac{u^2 + 4M}{u} \dot{u}.$$

Подстановка этого равенства во второе уравнение (34.74) приводит к уравнению

$$C_1 = \pm u \dot{u},$$

Общее решение которого имеет вид

$$C_1(\tau + \text{const}) = \pm \frac{u^2}{2}. \quad (34.78)$$

Поэтому нулевые геодезические при $u \rightarrow 0$ также уходят в бесконечность $t \rightarrow \pm\infty$ при конечном значении канонического параметра τ , следовательно, не полны.

Таким образом, мы показали, что мост Эйнштейна–Розена является геодезически неполным при $u = 0$. Качественное поведение времениподобных геодезических показано на рис.34.9.

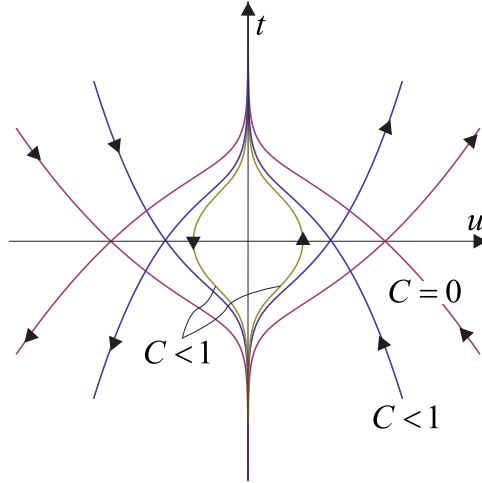


Рис. 34.9: Качественное поведение времениподобных геодезических, $C > 0$. Нулевые геодезические соответствуют значению $C = 0$. Каждую геодезическую можно сдвигать вдоль оси времени t .

Теперь покажем, что времениподобные геодезические хорошо определены и полны в физическом секторе. Это дает естественный способ отождествить геодезические, находящиеся по разные стороны прямой $u = 0$. Для того, чтобы разделить физические и нефизические степени свободы нам понадобится гамильтонова формулировка теории, которая была рассмотрена в разделе 22.8. Для простоты рассмотрим только времениподобные геодезические. Обозначим координаты и импульсы пробной частицы через q^α, p_α . Тогда действие для пробной частицы (22.82) в гамильтоновой форме примет вид

$$S = \int dt (p_\alpha \dot{q}^\alpha - H) = \int dt \left(p_\mu \dot{q}^\mu - |\dot{q}^0 N| \sqrt{\hat{p}^2 + m^2} + \dot{q}^0 N^\mu p_\mu \right), \quad (34.79)$$

где N и N^μ – функции хода и сдвига в АДМ параметризации метрики (34.45) и $\hat{p}^2 := -\hat{g}^{\mu\nu} p_\mu p_\nu$. В дальнейшем мы будем следить за возможными знаками функции хода

$N > 0$ и $N < 0$. Временная нефизическая компонента импульса $p_0(\tau)$ определяется связью первого рода

$$G := \sqrt{\hat{p}^2 + m^2} - \left| \frac{p_0 - N^\mu p_\mu}{N} \right| = 0. \quad (34.80)$$

Нефизическая степень свободы $q^0(\tau)$ является пока произвольной функцией, которую надо зафиксировать калибровочным условием.

В общем случае параметр τ при гамильтоновой формулировке теории не является каноническим. В дальнейшем он будет фиксирован калибровочным условием.

Рассмотрим случай $\dot{q}^0 N > 0$. Тогда гамильтоновы уравнения движения для физических степеней свободы имеют вид

$$\dot{q}^\mu = \dot{q}^0 \frac{\partial H_{\text{eff}}}{\partial p_\mu}, \quad \dot{p}_\mu = -\dot{q}^0 \frac{\partial H_{\text{eff}}}{\partial q^\mu}, \quad (34.81)$$

где введен эффективный гамильтониан

$$H_{\text{eff}} := N \sqrt{\hat{p}^2 + m^2} - N^\mu p_\mu.$$

Нетрудно проверить, что энергия $E := H_{\text{eff}}$ сохраняется при произвольной функции $q^0(\tau)$, если метрика $g_{\alpha\beta}(x)$ не зависит от времени t . Сохранение энергии является следствием того, что функция $q^0(\tau)$ описывает свободу в выборе параметризации мировой линии пробной частицы, и ее необходимо фиксировать калибровочным условием.

Функция хода определяется метрикой (34.25) с точностью до знака. Положим

$$N := \frac{u}{\sqrt{u^2 + 4M}}. \quad (34.82)$$

Эта функция положительна и отрицательна, соответственно, на правой, $u > 0$, и левой, $u < 0$, полуплоскостях. Это важно, т.к. если мы выберем всюду положительную функцию хода $N = |N|$, то она будет иметь скачок производной, который приведет к появлению нетривиального тензора энергии-импульса в правой части уравнений Эйнштейна, пропорционального $\delta(u)$. В последнем случае мост Эйнштейна–Розена уже не будет решением вакуумных уравнений Эйнштейна.

Теперь зафиксируем калибровку и проанализируем уравнения движения (34.81) для метрики (34.25). Для упрощения формул положим $q^1 := q$ и $p_1 := p$. Тогда уравнения (34.81) примут вид

$$\dot{q} = \frac{\dot{q}^0 N p}{(q^2 + 4M) \sqrt{\hat{p}^2 + m^2}}, \quad (34.83)$$

$$\dot{p} = -\dot{q}^0 \sqrt{\hat{p}^2 + m^2} \partial_q N + \frac{\dot{q}^0 N q p^2}{(q^2 + 4M)^2 \sqrt{\hat{p}^2 + m^2}}, \quad (34.84)$$

где функция хода задана уравнением (34.82) с заменой $u \mapsto q$ и

$$\hat{p}^2 = \frac{p^2}{q^2 + 4M}.$$

Уравнения движения (34.83) можно решить благодаря закону сохранения энергии:

$$E = N \sqrt{\hat{p}^2 + m^2} = \text{const.}$$

Из этого равенства следует выражение для импульса

$$p = \pm \frac{1}{N} \sqrt{(q^2 + 4M)(E^2 - m^2 N^2)}.$$

Подстановка полученного выражения в уравнение (34.83) приводит к уравнению первого порядка для координаты

$$\dot{q} = \pm \frac{\dot{q}^0 N}{E(q^2 + 4M)} \sqrt{(q^2 + 4M)(E^2 - m^2 N^2)}. \quad (34.85)$$

Это уравнение определяет форму геодезической $dt/du = \dot{q}^0/\dot{q}$, которое было получено ранее (34.76), где $C := m^2/E^2$. Мы видим, что и лагранжев, и гамильтонов формализм приводят к одним и тем же геодезическим, как и должно быть.

Наконец, зафиксируем калибровку $\dot{q}^0 N = 1$. Тогда равенство (34.85) сведется к простому уравнению

$$\dot{q} = \pm \frac{1}{\sqrt{q^2 + 4M}} \sqrt{1 - C \frac{q^2}{q^2 + 4M}}. \quad (34.86)$$

Это и есть окончательный ответ, определяющий физическую степень свободы. Решения данного уравнения являются гладкими функциями и определены при всех $\tau \in \mathbb{R}$. Качественное поведение физических траекторий $q(\tau)$ показано на рис.34.10. Нулевые ($C = 0$) и времениподобные ($C < 1$) геодезические свободно начинаются и заканчиваются при $u = \pm\infty$. При $C > 1$ времениподобные траектории осциллируют вблизи линии $u = 0$. Все пробные частицы при этом свободно проходят через линию $u = 0$ как слева направо, так и справа налево.

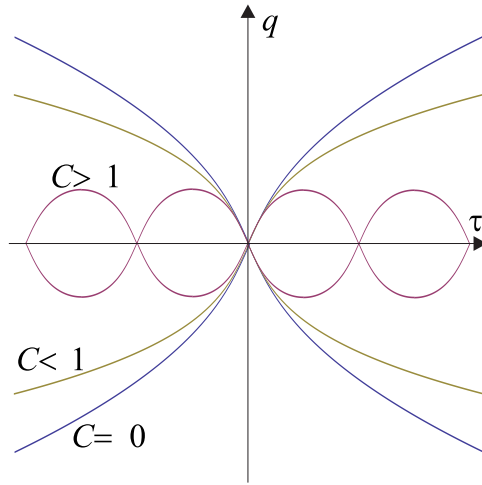


Рис. 34.10: Качественное поведение физических траекторий. Если $C > 1$, то времениподобные траектории являются осциллирующими.

Наиболее интересной является точка $u = q = 0$, вблизи которой уравнение (34.86) принимает простой вид $\dot{q} = \pm 1/\sqrt{4M}$. Оно имеет гладкое решение

$$q = \pm \frac{\tau}{\sqrt{4M}}, \quad (34.87)$$

где, без ограничения общности, постоянная интегрирования положена равной нулю. Таким образом, пробная частица свободно проходит через линию $u = 0$, если не

учитывать нефизическую степень свободы. С физической точки зрения это вполне оправдано, поскольку поведение калибровочных (нефизических) степеней свободы не так важно.

В выбранной калибровке параметр τ не является каноническим. Необходимым и достаточным условием того, что параметр канонический, является закон сохранения $C_0 = \text{const}$. Если сделать замену параметра $\tau \mapsto \tau_c(\tau)$, то закон сохранения для решения (34.86) примет вид

$$C_0 = \left[N^2 (\dot{q}^0)^2 - (q^2 + 4M) \dot{q}^2 \right] \left(\frac{d\tau}{d\tau_c} \right)^2 = \frac{C_0}{C_1^2} \frac{q^2}{q^2 + 4M} \left(\frac{d\tau}{d\tau_c} \right)^2.$$

Отсюда следует, что модуль канонического параметра τ_c вблизи точки $q = 0$ дается интегралом

$$|\tau_c| = \frac{1}{|C_1|} \int d\tau \sqrt{\frac{q^2}{q^2 + 4M}} \sim \int_0 d\tau \tau \sim \tau^2,$$

где использована асимптотика (34.87). Можно выбрать $\tau_c \sim \tau^2$ для $u > 0$ и $\tau_c \sim -\tau^2$ для $u < 0$ с тем, чтобы функция $\tau_c(\tau)$ была монотонной. Следовательно, полнота траекторий по отношению к параметру τ влечет за собой полноту траекторий по отношению к каноническому параметру.

На рис.34.9 стрелки показывают направление возрастания параметра τ , которое совпадает с направлением увеличения собственного времени. На левой полуплоскости выполнено неравенство $\dot{q}^0 < 0$, и поэтому наблюдатель видит античастицу (частицу той же массы, но противоположного заряда, как показано в разделе 22.8).

Особенность метрики Эйнштейна–Розена сдвинута в нефизический сектор, благодаря калибровочному условию, которое можно записать в виде

$$q^0 = \int^\tau d\tau' \frac{q(\tau')^2 + 4M}{q(\tau')}.$$

Отсюда вытекает, что $q \rightarrow 0$ для решений уравнения (34.86).

Физическая интерпретация полученного решения довольно странная. Предположим, что внешний наблюдатель может “видеть” координаты x^0 и x^μ . Тогда в данный момент времени x^0 он будет одновременно видеть две частицы: частицу справа и античастицу слева или наоборот. При $C < 1$ эти частицы движутся симметрично либо к, либо от линии $u = 0$. Если $C > 1$, то наблюдатель видит рождение пары частица-античастица в бесконечно удаленном прошлом $x^0 = -\infty$ и аннигиляцию частицы и античастицы в бесконечном будущем $x^0 = \infty$. Для наблюдателя, движущегося вместе с частицей, рождение пары и ее аннигиляция происходят за конечное собственное время, т.е. возникают осцилляции, изображенные на рис. 34.10. Конечно, реальные наблюдения требуют пересечения линии $u = 0$ светом, поэтому то, что наблюдается в действительности, остается под вопросом.

В предыдущем разделе мы показали, что гравитационное поле точечной массивной частицы описывается метрикой Шварцшильда в изотропных координатах. Это пространство-время глобально изометрично мосту Эйнштейна–Розена. Поэтому поведение геодезических одинаково, необходимо только преобразовать координаты. Метрика Шварцшильда в изотропных координатах геодезически неполна на критической сфере радиуса $r_* = M/2$, которая соответствует линии $u = 0$ моста Эйнштейна–Розена. Выше мы показали, что если ослабить требование геодезической полноты, заменив его на требование полноты геодезических только для физических

степеней свободы, то пространство-время вокруг точечной частицы будет полным в этом более слабом смысле, и ничего продолжать не надо. Это произошло, благодаря отождествлению геодезических с разных сторон критической сферы.

34.9 Вложения решения Шварцшильда

Для того, чтобы наглядно представить себе топологически нетривиальное многообразие, его можно вложить в плоское пространство более высокой размерности. Например, сферу обычно представляют в виде поверхности в трехмерном евклидовом пространстве, $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. При этом становится очевидной гладкость сферы как многообразия и ее симметрии. Вскоре после нахождения решения Шварцшильда появились попытки его вложения в плоское пространство-время более высокой размерности. Было доказано, что локальное вложение решения Шварцшильда в пятимерное плоское пространство-время невозможно [262]. Кроме этого, Казнер построил вложение решения Шварцшильда в шестимерное пространство-время $\mathbb{R}^{2,4}$. Таким образом, минимальная размерность плоского пространства-времени, в которое локально вкладывается решение Шварцшильда, равна шести. В настоящее время известно шесть минимальных вложений [263]. Ниже мы опишем три из них, которые можно назвать гиперболическим, эллиптическим и параболическим.

Вложение Фронсдела

Метрику решения Шварцшильда (34.1) можно вложить в плоское шестимерное пространство. Точнее, в шестимерное пространство Минковского $\mathbb{R}^{1,5}$ можно вложить четырехмерное многообразие $\mathbb{M} \hookrightarrow \mathbb{R}^{1,5}$ таким образом, что индуцированная на нем метрика совпадает с метрикой (34.1) в какой-то части подмногообразия $\mathbb{M}$. Это подмногообразие оказывается максимально продолженным и, по существу, совпадает с расширением Крускала–Секереша решения Шварцшильда [264]. Вложение Фронсдела является гиперболическим.

Рассмотрим шестимерное пространство Минковского $\mathbb{R}^{1,5}$ с декартовыми координатами $\{Z^1, \dots, Z^6\}$. Метрика этого пространства, по определению, имеет вид

$$ds^2 := (dZ^1)^2 - (dZ^2)^2 - (dZ^3)^2 - (dZ^4)^2 - (dZ^5)^2 - (dZ^6)^2. \quad (34.88)$$

Будем рассматривать это многообразие, как топологическое произведение $\mathbb{R}^{1,5} = \mathbb{R}^{1,2} \times \mathbb{R}^3$ трехмерного пространства Минковского $\mathbb{R}^{1,2}$ с декартовыми координатами $\{Z^1, Z^2, Z^3\}$ и трехмерного евклидова пространства $\mathbb{R}^3$ с декартовыми координатами $\{Z^4, Z^5, Z^6\}$.

Четырехмерное многообразие с метрикой Шварцшильда представляет собой топологическое произведение двумерного лоренцева многообразия и двумерной сферы, $\mathbb{M} = \mathbb{U} \times \mathbb{S}^2$. Будем считать, что на этом многообразии в некоторой окрестности заданы координаты t, ρ, θ, φ , которые мы отождествим с теми координатами, которые входят в решение Шварцшильда (34.1). При этом мы считаем, что координаты t, ρ заданы на $\mathbb{U}$, а θ, φ – обычные сферические координаты на сфере $\mathbb{S}^2$.

Мы умеем вкладывать сферу в трехмерное евклидово пространство и поэтому положим

$$Z^4 := \rho \sin \theta \cos \varphi, \quad Z^5 := \rho \sin \theta \sin \varphi, \quad Z^6 := \rho \cos \theta, \quad (34.89)$$

где ρ, θ, φ – три (из четырех) координат на $\mathbb{M}$. Тогда справедливо равенство (7.7)

$$-(dZ^4)^2 - (dZ^5)^2 - (dZ^6)^2 = -d\rho^2 - \rho^2(d\theta^2 + \sin^2 \theta d\varphi^2).$$

Это есть евклидова метрика в $\mathbb{R}^3$ в сферических координатах, взятая с обратным знаком. Второе слагаемое в правой части является метрикой на сфере, которая индуцирована вложением $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. Вложение $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ можно также описать с помощью алгебраического уравнения

$$(Z^4)^2 + (Z^5)^2 + (Z^6)^2 = \rho^2. \quad (34.90)$$

Это вложение, очевидно, симметрично относительно $\mathbb{SO}(3)$ вращений. Такое описание вложения является глобальным, в то время как параметризация (34.89) локальна.

Теперь нам необходимо добавить четвертую координату t и вложить двумерную лоренцеву поверхность $\mathbb{U}$ в трехмерное пространство Минковского $\mathbb{R}^{1,2}$. Для внешнего решения Шварцшильда при $\rho > 2M$ это можно сделать следующим образом

$$\begin{aligned} Z^1 &:= \pm 4M \sqrt{1 - \frac{2M}{\rho}} \operatorname{sh} \frac{t}{4M}, \\ Z^2 &:= \pm 4M \sqrt{1 - \frac{2M}{\rho}} \operatorname{ch} \frac{t}{4M}, \\ Z^3 &:= f(\rho), \end{aligned} \quad (34.91)$$

где $f(\rho)$ – некоторая дифференцируемая пока неизвестная функция от ρ . Функции Z^1 и Z^2 подобраны таким образом, чтобы в индуцированной метрике отсутствовали перекрестные члены $dt d\rho$ и коэффициент при dt^2 имел тот же вид, что и в метрике Шварцшильда. Если теперь подставить выражения для всех дифференциалов dZ в исходную метрику Лоренца (34.88), то получим следующее выражение для индуцированной метрики

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) dt^2 - \frac{16M^4 d\rho^2}{\rho^4 \left(1 - \frac{2M}{\rho}\right)} - f'^2 d\rho^2 - d\rho^2 - \rho^2 (\sin^2 d\theta^2 + d\varphi^2), \quad (34.92)$$

где $f' := df/d\rho$. Для того, чтобы полученная метрика совпала с метрикой Шварцшильда (34.1) необходимо и достаточно, чтобы функция $f(\rho)$ была произвольным решением следующего дифференциального уравнения

$$f'^2 = \frac{2M(\rho^2 + 2M\rho + 4M^2)}{\rho^3}. \quad (34.93)$$

Таким образом, если подмногообразие $\mathbb{M} \hookrightarrow \mathbb{R}^{1,5}$ задано параметрически (вернее, какая то его пока неизвестная часть, на которой определены координаты t, ρ, θ, φ), то функции, определяющие вложение, могут иметь вид (34.89), (34.91), где функция $f(\rho)$ является произвольным решением обыкновенного дифференциального уравнения (34.93).

Заметим, что вложение $\mathbb{U} \hookrightarrow \mathbb{R}^{1,2}$, заданное формулами (34.91), (34.93) индуцирует на $\mathbb{U}$ двумерную лоренцеву метрику

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) dt^2 - \frac{2M}{\rho - 2M} d\rho^2,$$

которая отличается от двумерной части метрики Шварцшильда (34.6). Метрика (34.6) возникает только после добавления четвертого слагаемого $-d\rho^2$ в выражении

(34.92). Поэтому нельзя говорить, что двумерная часть метрики Шварцшильда (34.6) индуцирована вложением $\mathbb{U} \hookrightarrow \mathbb{R}^{1,2}$. Только четырехмерная метрика Шварцшильда (34.1) индуцируется вложением $\mathbb{U} \times \mathbb{S}^2 \hookrightarrow \mathbb{R}^{1,5}$.

Поскольку вложение двумерной сферы в трехмерное евклидово пространство $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ хорошо известно, то проанализируем вложение двумерной лоренцевой поверхности в трехмерное пространство Минковского $\mathbb{U} \hookrightarrow \mathbb{R}^{1,2}$, которое задано параметрически уравнениями (34.91) и (34.93). Для того, чтобы понять, как устроено подмногообразие $\mathbb{U}$ в целом, необходимо проанализировать решения дифференциального уравнения (34.93). Этот анализ затруднен тем, что уравнение не интегрируется в явном виде. Тем не менее качественное поведение решений довольно просто.

Чтобы упростить формулы совершим растяжку радиальной координаты и искомой функции:

$$\rho \mapsto 2M\rho, \quad f \mapsto 2Mf. \quad (34.94)$$

Тогда уравнение (34.93) примет вид

$$f'^2 = \frac{\rho^2 + \rho + 1}{\rho^3}. \quad (34.95)$$

Первое, что бросается в глаза, это то, что правая часть уравнения положительно определена при $\rho > 0$, а при $\rho \rightarrow 0$ мы имеем степенную особенность. Для определенности мы выберем положительный корень $f' > 0$. Отрицательный корень $f' < 0$ соответствует простому отражению координаты $Z^3 \mapsto -Z^3$. Тогда вблизи особенности уравнение принимает вид

$$df = \frac{d\rho}{\rho^{3/2}}$$

и легко интегрируется:

$$f + \text{const} = -\frac{2}{\sqrt{\rho}}.$$

Постоянная интегрирования соответствует сдвигу координаты Z^3 и не существенна.

Асимптотика уравнения (34.95) на бесконечности имеет вид

$$df = \frac{d\rho}{\sqrt{\rho}}.$$

Откуда следует, что при $\rho \rightarrow \infty$ решение стремится к функции

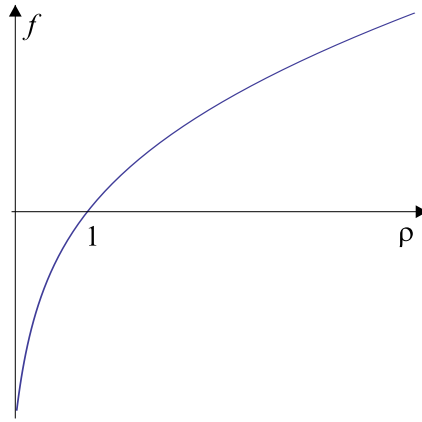
$$f + \text{const} = 2\sqrt{\rho}.$$

Таким образом, функция f монотонно возрастает от $-\infty$ до ∞ , когда ρ возрастает от 0 до ∞ .

После растяжки координат (34.94) горизонту соответствует значение $\rho = 1$. Выберем постоянную интегрирования в уравнении (34.95) таким образом, чтобы было выполнено равенство $f(1) = 0$. Тогда решение уравнения (34.95) качественно ведет себя так, как показано на рис.34.11.

Поскольку мы рассматриваем вложение внешнего решения Шварцшильда, то необходимо ограничиться положительной частью решения, которая соответствует $\rho > 1$. В дальнейшем мы увидим, что отрицательная часть решения описывает вложение внутреннего решения Шварцшильда $0 < \rho < 1$.

Заметим, что на горизонте решение уравнения (34.95) никакой особенности не имеет.

Рис. 34.11: Решение уравнения (34.95) с начальным условием $f(1) = 0$.

Чтобы проанализировать вложение $\mathbb{U} \hookrightarrow \mathbb{R}^{1,2}$ исключим координату t из описания. Тогда поверхность $\mathbb{U}$ будет задана двумя уравнениями

$$\begin{aligned} (Z^2)^2 - (Z^1)^2 &= 4 \left(1 - \frac{1}{\rho}\right), \\ Z^3 &= \int_a^\rho d\rho' \sqrt{\frac{\rho'^2 + \rho' + 1}{\rho'^3}}, \quad a > 0. \end{aligned} \quad (34.96)$$

При $\rho \in (0, \infty)$ координата Z^3 хорошо определена и меняется в интервале $Z^3 \in (-\infty, \infty)$. Постоянная интегрирования a соответствует сдвигу Z^3 и является несущественной. Положим $a = 1$. Тогда горизонты $\rho = 1$ будут лежать в плоскости $Z^3 = 0$.

Напомним, что мы рассматриваем решение Шварцшильда вне горизонта, т.е. $\rho > 1$. Каждому сечению поверхности $\rho = \text{const} \Leftrightarrow Z^3 = \text{const}$ соответствует две ветви гиперболы, которые на плоскости Z^1, Z^2 лежат в I и III квадрантах, если ось Z^2 выбрана в качестве оси абсцисс, а ось Z^1 (время) – ординат. Если во втором уравнении (34.91) выбран знак плюс, то мы имеем ветвь из квадранта I. Если минус, то из квадранта III. Знак “ $\pm$ ” в первом уравнении (34.91) соответствует взаимной ориентации координат Z^1 и t . На горизонте, $\rho = 1$, гиперболы вырождаются в две пересекающиеся прямые $Z^1 = \pm Z^2$. При $\rho \rightarrow \infty$ возникают предельные гиперболы

$$(Z^2)^2 - (Z^1)^2 = 4. \quad (34.97)$$

Таким образом, внешнее решение Шварцшильда отображается на ту часть поверхности $\mathbb{U}$, которая проектируется либо в первый, либо в третий квадрант. При этом проекции поверхности $\mathbb{U}$ на квадранты I и III ограничены ветвями гиперболы (34.97) и прямыми $Z^1 = \pm Z^2$.

Для внутреннего решения Шварцшильда, $0 < \rho < 2M$, вложение определяется функциями

$$\begin{aligned} Z^1 &:= \pm 4M \sqrt{\frac{2M}{\rho} - 1} \operatorname{ch} \frac{t}{4M}, \\ Z^2 &:= \pm 4M \sqrt{\frac{2M}{\rho} - 1} \operatorname{sh} \frac{t}{4M}, \\ Z^3 &:= f(\rho), \end{aligned} \quad (34.98)$$

Анализ проводится аналогично случаю внешнего решения Шварцшильда. После растяжки (34.94) возникает та же поверхность (34.96), что и для внешнего решения Шварцшильда. Отличие заключается в том, что теперь внутреннее решение Шварцшильда отображается на ту часть поверхности, которая лежит либо над вторым, либо над четвертым квадрантом плоскости Z^1, Z^2 . При этом проекция поверхности $\mathbb{U}$ на плоскость Z^1, Z^2 полностью покрывает квадранты II и IV.

Вложение $\mathbb{U} \hookrightarrow \mathbb{R}^{1,2}$, заданное формулами (34.96) при $a = 1$, показано на рис.34.12.

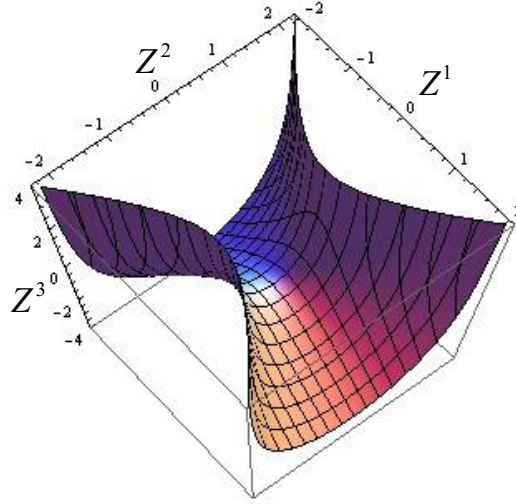


Рис. 34.12: Вложение Фронсдела $\mathbb{U} \hookrightarrow \mathbb{R}^{1,2}$.

Система уравнений (34.96) и, следовательно, поверхность $\mathbb{U}$ инвариантна относительно лоренцевых $\mathbb{SO}_0(1, 1)$ вращений в плоскости Z^1, Z^2 . Эта инвариантность соответствует статичности решения Шварцшильда вне горизонта. Кроме этого имеется симметрия относительно обращения времени $Z^1 \mapsto -Z^1$ и пространства $Z^2 \mapsto -Z^2$.

Таким образом, построена гладкая поверхность $\mathbb{U}$, вложенная в трехмерное пространство Минковского $\mathbb{R}^{1,2}$. Экстремали полны при $\rho \rightarrow \infty$ и неполны при $\rho \rightarrow 0$. Но здесь их продолжать не надо, т.к. кривизна имеет особенность. Поскольку других граничных точек у поверхности $\mathbb{U}$ нет, то построенное вложение определяет максимальное продолжение решения Шварцшильда. Поверхность $\mathbb{U}$ взаимно однозначно проектируется на ту часть плоскости Z^1, Z^2 , которая ограничена двумя ветвями предельной гиперболы (34.97). Поэтому координаты Z^1, Z^2 являются глобальными координатами для максимального продолжения решения Шварцшильда. Они аналогичны глобальным координатам T, X (34.23) для максимального расширения Крускала–Секереша.

Сечениями поверхности $\mathbb{U} \hookrightarrow \mathbb{R}^{1,2}$, соответствующими постоянному значению ρ , являются гиперболы при $0 < \rho < 2M$ и $2M < \rho < \infty$. На горизонте $\rho = 2M$ гиперболы вырождаются в две пересекающиеся прямые. Абелева группа сдвигов по времени $\mathbb{R} \ni t \mapsto t + \text{const} \in \mathbb{R}$, которая является подгруппой группы симметрии исходной метрики Шварцшильда (34.1), при вложении Фронсдела реализуется в виде группы Лоренца $\mathbb{SO}_0(1, 1)$ действующей в плоскости Z^1, Z^2 объемлющего пространства-времени $\mathbb{R}^{1,5}$. То есть мы имеем гомоморфизм группы сдвигов на группу Лоренца $\mathbb{R} \rightarrow \mathbb{SO}_0(1, 1)$. Поэтому описанное вложение назовем гиперболическим. Оно оказалось не только локальным, но и глобальным, т.к. поверхность $\mathbb{U}$ является максимально продолженной вдоль геодезических.

Вложение Казнера

Исторически, первое вложение решения Шварцшильда в шестимерное пространство-время было предложено Казнером [265]. Мы увидим, что вложение Казнера является вложением внешней части решения Шварцшильда и потому является только локальным. Оно является эллиптическим.

Рассмотрим плоское пространство-время $\mathbb{R}^{2,4}$ с двумя временными координатами. Обозначим декартовы координаты через $\{Y^1, \dots, Y^6\}$. Метрика этого пространства, по определению, имеет вид

$$ds^2 := (dY^1)^2 + (dY^2)^2 - (dY^3)^2 - (dY^4)^2 - (dY^5)^2 - (dY^6)^2. \quad (34.99)$$

Будем рассматривать это многообразие, как топологическое произведение $\mathbb{R}^{2,4} = \mathbb{R}^{2,1} \times \mathbb{R}^3$ трехмерного псевдоевклидова пространства $\mathbb{R}^{2,1}$ с декартовыми координатами $\{Y^1, Y^2, Y^3\}$ и трехмерного евклидова пространства $\mathbb{R}^3$ с декартовыми координатами $\{Y^4, Y^5, Y^6\}$.

Таким образом, нам необходимо построить вложение

$$\mathbb{U} \times \mathbb{S}^2 \hookrightarrow \mathbb{R}^{2,1} \times \mathbb{R}^3.$$

В настоящем разделе $\mathbb{U}$ обозначает не максимально продолженную лоренцеву поверхность, изображаемую диаграммой Картера–Пенроуза, а только внешнюю часть решения Шварцшильда.

Как и в случае вложения Фронсдела сферу $\mathbb{S}^2$ мы вложим в трехмерное евклидово пространство $\mathbb{R}^3$ с помощью формул (34.89) или (34.90). Вложение $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$ при $\rho > 2M$ определим формулами

$$\begin{aligned} Y^1 &:= \pm 4M \sqrt{1 - \frac{2M}{\rho}} \sin \frac{t}{4M}, \\ Y^2 &:= \pm 4M \sqrt{1 - \frac{2M}{\rho}} \cos \frac{t}{4M}, \\ Y^3 &:= f(\rho), \end{aligned} \quad (34.100)$$

которые отличаются от вложения Фронсдела тем, что гиперболические функции заменены на тригонометрические. Здесь, по-прежнему, $f(\rho)$ – некоторая неизвестная дифференцируемая функция. Функции Y^1 и Y^2 подобраны таким образом, чтобы перекрестные члены $dt d\rho$ отсутствовали, а коэффициент при dt^2 имел тот же вид, что и в метрике Шварцшильда. При таком вложении метрика, индуцированная на $\mathbb{U} \times \mathbb{S}^2$, примет вид

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) dt^2 + \frac{16M^4 d\rho^2}{\rho^4 \left(1 - \frac{2M}{\rho}\right)} - f'^2 d\rho^2 - d\rho^2 - \rho^2 (\sin^2 d\theta^2 + d\varphi^2). \quad (34.101)$$

Она отличается от метрики (34.101) только знаком перед вторым слагаемым. Для того, чтобы эта метрика совпала с метрикой Шварцшильда, в качестве неизвестной функции f можно взять произвольное решение дифференциального уравнения

$$f'^2 = \frac{2M}{\rho - 2M} \left[1 + \left(\frac{2M}{\rho} \right)^3 \right]. \quad (34.102)$$

Вне горизонта $\rho > 2M$ правая часть положительно определена и, следовательно, вложение существует. Для определенности мы выберем положительный корень $f' > 0$. Отрицательный корень $f' < 0$ соответствует простому отражению координаты $Y^3 \mapsto -Y^3$.

Если $\rho \rightarrow 2M$, то уравнение (34.102) в главном приближении принимает вид

$$f' = \sqrt{\frac{4M}{\rho - 2M}}.$$

Отсюда следует асимптотика решения уравнения (34.102) вблизи горизонта

$$f + \text{const} = 4\sqrt{M(\rho - 2M)}, \quad \rho \rightarrow 2M. \quad (34.103)$$

Постоянная интегрирования соответствует сдвигу координаты Y^3 , является несущественной, и, для простоты, положим ее равной нулю.

На бесконечности $\rho \rightarrow \infty$ уравнение (34.102) имеет вид

$$f' = \sqrt{\frac{2M}{\rho}},$$

и мы получаем асимптотику

$$f + \text{const} = 2\sqrt{2M\rho}, \quad \rho \rightarrow \infty.$$

Таким образом, при увеличении радиуса от горизонта $\rho = 2M$ до бесконечности решение уравнения (34.102) монотонно возрастает от нуля до бесконечности.

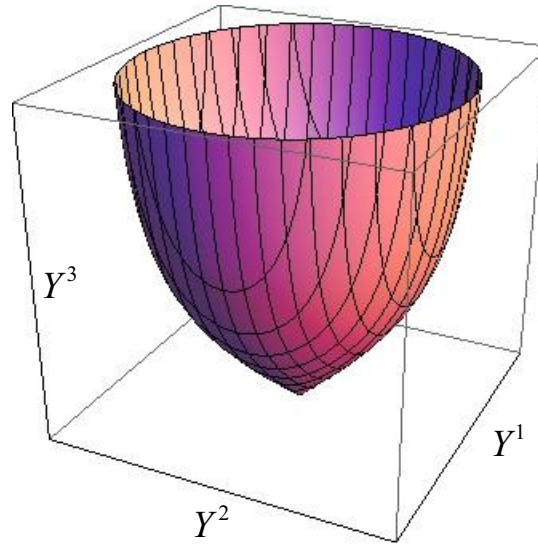
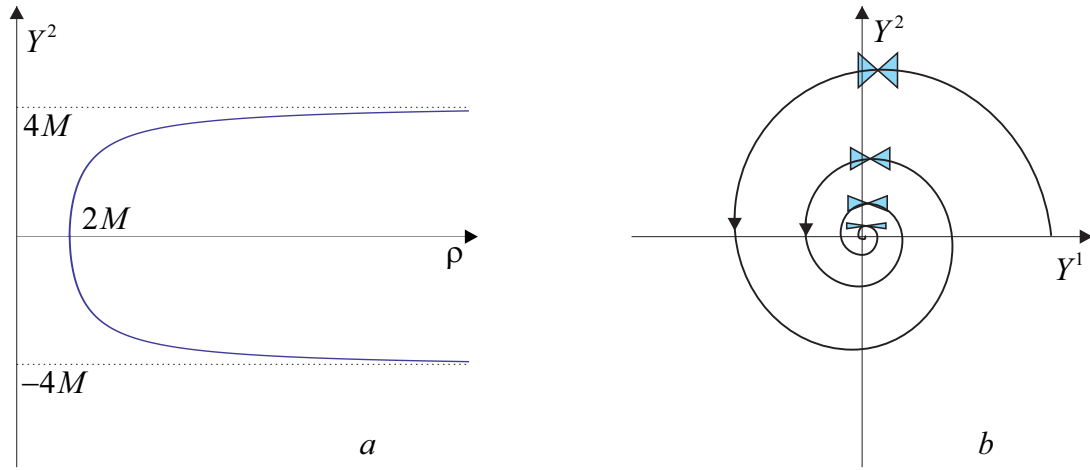
Вложение $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$, которое определяется формулами (34.100), можно записать в более прозрачном виде, исключив параметр t :

$$\begin{aligned} (Y^1)^2 + (Y^2)^2 &= (4M)^2 \left(1 - \frac{2M}{\rho}\right), \\ Y^3 &= f(\rho), \end{aligned} \quad (34.104)$$

где $f(\rho)$ – произвольное решение уравнения (34.102). Это – поверхность вращения во временной плоскости Y^1, Y^2 . Сечениями этой поверхности плоскостями, перпендикулярными оси Y^3 , являются окружности радиус которых возрастает от нуля при $\rho = 2M \Leftrightarrow Y^3 = 0$ до $4M$ при $\rho \rightarrow \infty \Leftrightarrow Y^3 \rightarrow \infty$. Знаки $\pm$ в уравнениях (34.100) несущественны, т.к. соответствуют изменению направления вращения при увеличении времени t и его сдвигу $t \mapsto t + \text{const}$. Вложение поверхности $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$ показано на рис.34.13.

Проанализируем вложение $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$ подробнее. Для простоты заменим координату Y^3 на ρ . Тогда в системе координат Y^1, Y^2, ρ поверхность $\mathbb{U}$ задается первым алгебраическим уравнением в системе (34.104). На рисунке 34.14 показано сечение поверхности $\mathbb{U}$ плоскостью $Y^1 = 0$. Поверхность $\mathbb{U}$ является поверхностью вращения вокруг оси ρ и, как многообразие, является гладкой.

Поверхность $\mathbb{U}$ можно представить следующим образом. Внешнее решение Шварцшильда, точнее, его двумерная часть (34.6) определена на полуплоскости $t \in (-\infty, \infty)$, $\rho \in (2M, \infty)$. Ввиду статичности решения мы можем отождествить лучи $t \sim t + 8\pi M$, что было сделано в формулах (34.100). В результате получится полубесконечный цилиндр. Теперь заметим, что длина образующей окружности цилиндра стремится к

Рис. 34.13: Вложение Казнера $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$.Рис. 34.14: Сечение $Y^1 = 0$ вложения $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$ в координатах Y^2, ρ (a). Поведение светоподобных геодезических вблизи вершины (b).

нулю при приближении к горизонту $\rho \rightarrow 2M$. При вложении Казнера эта окружность стягивается в точку. Строго говоря, внешняя часть решения Шварцшильда определена только при $\rho > 2M$ и вершина поверхности вращения $\rho = 2M$ не принадлежит $\mathbb{U}$. Однако мы ее добавим, и тогда возникнет гладкая поверхность, которая диффеоморфна плоскости.

Поверхность $\mathbb{U}$ является гладкой, но метрика, индуцированная на ней имеет особую точку – вершину $Y^1 = Y^2 = 0$, $\rho = 2M$, которая соответствует горизонту. Это видно из вложения. Действительно, во всех точках вне вершины существуют как времениподобные касательные векторы, так и пространственноподобные. Это значит, что индуцированная метрика имеет лоренцеву сигнатуру. В то же время все касательные векторы в вершине времениподобны. То есть индуцированная метрика не имеет лоренцевой сигнатуры, и вершина является особой точкой для метрики, но не для поверхности как многообразия. Ниже мы увидим, что данная поверхность с индуцированной метрикой не является геодезически полной в вершине.

Чтобы проанализировать особую точку, рассмотрим малую окрестность вершины в координатах Y^1, Y^2 . Найдем метрику в этих координатах. Из формул (34.104) и

асимптотики (34.103) следует, что вблизи вершины уравнения вложения принимают вид

$$\begin{aligned}(Y^1)^2 + (Y^2)^2 &= 8M(\rho - 2M), \\ Y^3 &= 4\sqrt{M(\rho - 2M)}.\end{aligned}\tag{34.105}$$

После исключения ρ получим простое алгебраическое уравнение, определяющее вложение вблизи вершины

$$(Y^1)^2 + (Y^2)^2 = \frac{1}{2} (Y^3)^2.\tag{34.106}$$

Теперь вычислим метрику, индуцированную вложением $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$,

$$\begin{aligned}ds^2 &= (dY^1)^2 + (dY^2)^2 - (dY^3)^2 = \\ &= \frac{(Y^2)^2 - (Y^1)^2}{(Y^1)^2 + (Y^2)^2} (dY^1)^2 - 4 \frac{Y^1 Y^2}{(Y^1)^2 + (Y^2)^2} dY^1 dY^2 + \frac{(Y^1)^2 - (Y^2)^2}{(Y^1)^2 + (Y^2)^2} (dY^2)^2.\end{aligned}\tag{34.107}$$

В матричных обозначениях метрика имеет вид

$$g_{\alpha\beta} = \frac{1}{(Y^1)^2 + (Y^2)^2} \begin{pmatrix} (Y^2)^2 - (Y^1)^2 & -2Y^1 Y^2 \\ -2Y^1 Y^2 & (Y^1)^2 - (Y^2)^2 \end{pmatrix}.$$

Определитель этой метрики равен $\det g_{\alpha\beta} = -1$.

В вершине компоненты метрики (34.107) не определены. Если приближаться к вершине вдоль лучей $Y^2 = kY^1$, $k = \text{const}$, то метрика (34.107) становится хорошо определенной:

$$ds^2 = \frac{k^2 - 1}{k^2 + 1} (dY^1)^2 - \frac{4k}{k^2 + 1} dY^1 dY^2 - \frac{k^2 - 1}{k^2 + 1} (dY^2)^2.$$

Таким образом, вблизи вершины поверхности вращения $\mathbb{U}$ индуцированную метрику можно записать в виде (34.107). При этом в вершине ее можно определить по непрерывности вдоль лучей.

Для дальнейшего анализа введем полярные координаты:

$$Y^1 := r \cos \varphi, \quad Y^2 := r \sin \varphi.$$

Тогда индуцированная метрика примет простой вид

$$ds^2 = -dr^2 + r^2 d\varphi^2.$$

Она отличается от евклидовой метрики на плоскости в полярных координатах только знаком перед первым слагаемым и, следовательно, сигнатурой. В рассматриваемом случае координата r является пространственной, а φ – времениподобной.

Теперь рассмотрим светоподобные геодезические, которые определяются уравнением $ds^2 = 0$ или

$$dr^2 = r^2 d\varphi^2 \quad \Leftrightarrow \quad \frac{dr}{d\varphi} = \pm r.$$

Его общее решение имеет вид

$$\ln r = \pm \varphi + \text{const}.$$

При приближении к вершине $r \rightarrow 0$ и $\varphi \rightarrow \pm\infty$, т.е. светоподобные геодезические наматываются на центр $r = 0$ бесконечное число раз. Поскольку

$$\lim_{r \rightarrow 0} \frac{dr}{d\varphi} = 0,$$

то раствор светового конуса уменьшается по мере приближения к центру и в пределе световой конус вырождается в линию. Можно доказать, что светоподобные геодезические достигают вершины при конечном значении канонического параметра. Следовательно поверхность $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$ с индуцированной метрикой является гладкой, но геодезически неполной.

Метрика, индуцированная на поверхности $\mathbb{U}$ вложением $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$, отличается от той метрики, которая индуцирована полным вложением $\mathbb{U} \times \mathbb{S}^2 \hookrightarrow \mathbb{R}^{2,4}$ дополнительным слагаемым

$$ds^2 = -dr^2 + r^2 d\varphi^2 - d\rho^2,$$

что следует из формулы (34.101). Из асимптотики (34.105) вытекает, что вблизи вершины

$$\rho - 2M = \frac{1}{8M} \left[(Y^1)^2 + (Y^2)^2 \right] = \frac{1}{8M} r^2.$$

Это означает, что дополнительное слагаемое

$$d\rho^2 = \left(\frac{r}{4M} \right)^2 dr^2$$

вблизи вершины $r = 0$ не дает вклада в главное приближение для метрики. Поэтому поверхность $\mathbb{U}$ с метрикой, индуцированной полным вложением $\mathbb{U} \times \mathbb{S}^2 \hookrightarrow \mathbb{R}^{2,4}$ также геодезически неполна в вершине. Таким образом, вложение Казнера является локальным, т.к. поверхность $\mathbb{U}$ не является геодезически полной в вершине.

Аналогично можно вложить внутреннее решение Шварцшильда $0 < \rho < 2M$ в шестимерное евклидово пространство $\mathbb{R}^{2,4}$.

Сечениями поверхности $\mathbb{U}$ плоскостями $\rho = \text{const}$ являются окружности. При этом абелева группа сдвигов по времени $\mathbb{R} \ni t \mapsto t + \text{const} \in \mathbb{R}$ реализуется как группа вращений $\mathbb{SO}(2)$ временной плоскости с координатами Y^1, Y^2 . То есть мы имеем гомоморфизм группы сдвигов на группу вращений $\mathbb{R} \rightarrow \mathbb{SO}(2)$. Поэтому вложение Казнера называется эллиптическим.

Вложение Фуджитани–Икеды–Мацумото

Вскоре после Фронсдела еще одно минимальное вложение решения Шварцшильда в шестимерное пространство-время $\mathbb{R}^{2,4}$ было предложено Фуджитани, Икедой и Мацумото [266]. Это вложение локально и является параболическим.

Как и для вложения Казнера рассмотрим шестимерное плоское пространство $\mathbb{R}^{2,4}$ с декартовыми координатами $Y^1, \dots, Y^6$, где первые две координаты Y^1 и Y^2 являются временными. Метрика пространства $\mathbb{R}^{2,4}$, по определению, имеет вид (34.99). Как и прежде, вложение сферы $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ определяется формулами (34.89) или (34.90).

Вложение $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$ зададим следующим образом:

$$\begin{aligned} Y^1 &:= t \sqrt{1 - \frac{2M}{\rho}}, \\ Y^2 &:= \frac{M}{2} \left(\frac{t^2}{M^2} - 1 \right) \sqrt{1 - \frac{2M}{\rho}} + h(\rho), \\ Y^3 &:= \frac{M}{2} \left(\frac{t^2}{M^2} + 1 \right) \sqrt{1 - \frac{2M}{\rho}} + h(\rho), \end{aligned} \quad (34.108)$$

где $h(\rho)$ пока неизвестная дифференцируемая функция ρ . Это вложение определено вне горизонта $\rho > 2M$.

Прямые вычисления приводят к следующему выражению для метрики, индуцированной вложением $\mathbb{U} \times \mathbb{S}^2 \hookrightarrow \mathbb{R}^{2,4}$,

$$ds^2 = \left(1 - \frac{2M}{\rho}\right) dt^2 - \frac{2M^2 h' d\rho^2}{\rho^2 \sqrt{1 - \frac{2M}{\rho}}} - d\rho^2 - \rho^2 (\sin^2 \theta d\theta^2 + d\varphi^2), \quad (34.109)$$

где $h' := dh/d\rho$. Для того, чтобы эта метрика совпала с метрикой Шварцшильда (34.1) необходимо и достаточно, чтобы функция h удовлетворяла уравнению

$$h' = \frac{\rho}{M \sqrt{1 - \frac{2M}{\rho}}}. \quad (34.110)$$

Общее решение данного уравнения имеет вид

$$h + \text{const} = \frac{\rho(\rho + 3M)}{2M} \sqrt{1 - \frac{2M}{\rho}} + \frac{3M}{2} \ln \left[\frac{\rho}{M} \left(1 + \sqrt{1 - \frac{2M}{\rho}} \right) - 1 \right]. \quad (34.111)$$

Постоянная интегрирования соответствует сдвигу координат Y^2 и Y^3 , является несущественной, и мы положим ее равной нулю.

Поскольку правая часть уравнения (34.110) положительна, то функция $h(\rho)$ монотонно возрастает от нуля до бесконечности, $h \in (0, \infty)$, при возрастании радиуса от горизонта $\rho = 2M$ до бесконечности $\rho = \infty$. Качественное поведение функции $h(\rho)$ показано на рис.34.15

Проанализируем вложение $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$ подробнее. С этой целью введем конусные координаты

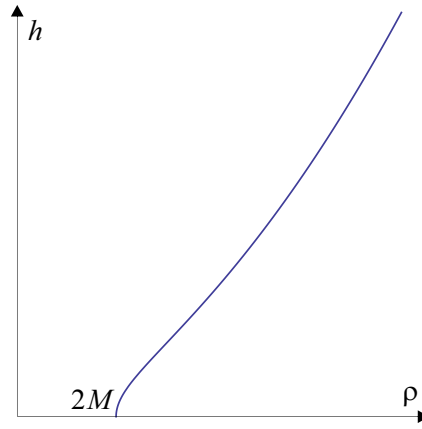
$$Y^\pm := Y^2 \pm Y^3.$$

Тогда вложение (34.108) запишется в виде

$$\begin{aligned} Y^+ &= \frac{t^2}{M} \sqrt{1 - \frac{2M}{\rho}} + 2h, \\ Y^- &= -M \sqrt{1 - \frac{2M}{\rho}}, \end{aligned} \quad (34.112)$$

при этом функция вложения $Y^1(t, \rho)$ остается прежней. Из выражений для функций вложения Y^1, Y^+ исключим время. Тогда для функций вложения будет выполнено алгебраическое уравнение

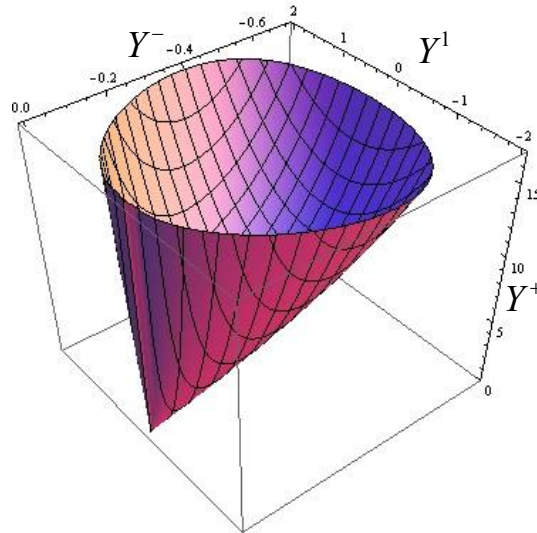
$$M \sqrt{1 - \frac{2M}{\rho}} Y^+ = (Y^1)^2 + 2Mh \sqrt{1 - \frac{2M}{\rho}}. \quad (34.113)$$

Рис. 34.15: Качественное поведение функции $h(\rho)$.

Функция вложения $Y^-(\rho)$ не зависит от времени t и определяется прежней формулой (34.112).

При фиксированном значении ρ и, следовательно, Y^1 в плоскости Y^1, Y^+ мы имеем параболу (34.113). Таким образом, трансляции по времени реализованы в виде трансляций вдоль парабол. Поэтому вложение Фуджитани–Икеды–Мацумото называется параболическим.

При $\rho \in (2M, \infty)$ функция вложения Y^- меняется от нуля до $-1/\sqrt{2}$. Следовательно, поверхность $\mathbb{U}$ представляет собой объединение парабол, которые параметризуются координатой $Y^- \in (0, -1/\sqrt{2})$. При $Y^- = -1/\sqrt{2}$ раствор параболы максимален и стремится к нулю при $Y^- \rightarrow 0$. Горизонту $\rho = 2M$ соответствует одна точка – начало координат $Y^1 = Y^2 = Y^3 = 0$.

Рис. 34.16: Вложение Фуджитани–Икеды–Мацумото $\mathbb{U} \hookrightarrow \mathbb{R}^{2,1}$.

Проанализируем поверхность $\mathbb{U}$ вблизи горизонта $\rho \rightarrow 2M$. С этой целью введем новую переменную

$$\rho = 2M + x, \quad x \ll M.$$

Тогда в главном порядке по x/M справедливы равенства:

$$\sqrt{1 - \frac{2M}{\rho}} \simeq \sqrt{\frac{x}{2M}},$$

$$h \simeq 5M^2 \sqrt{\frac{x}{2M}}.$$

Функции вложения в главном порядке имеют вид

$$\begin{aligned} Y^1 &\simeq t \sqrt{\frac{x}{2M}}, \\ Y^+ &\simeq t^2 \sqrt{\frac{x}{4M}} + 5M^2 \sqrt{\frac{x}{M}}, \\ Y^- &\simeq -\sqrt{\frac{x}{4M}}. \end{aligned} \tag{34.114}$$

Вблизи горизонта параболы определяются уравнением

$$\sqrt{\frac{x}{M}} Y^+ = (Y^1)^2 + 5Mx.$$

34.10 Эвристический “вывод” решения Шварцшильда

Эвристический подход к выводу метрики Шварцшильда описан в статье [267]. Рассмотрим плоское пространство-время Минковского $\mathbb{R}^{1,3}$ с декартовой системой координат $\{x^0, x^1, x^2, x^3\} = \{x^0, x^\mu\}$, $\mu = 1, 2, 3$. Согласно закону всемирного тяготения тело массы M , находящееся в начале координат, создает в точке $x \in \mathbb{R}^{1,3}$ гравитационный потенциал

$$\varphi(x) = -\frac{GM}{\rho}, \tag{34.115}$$

где G – гравитационная постоянная и $\rho := \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$ – евклидово расстояние от начала координат. Рассмотрим систему координат, которая связана с наблюдателями, свободно падающими на центр вдоль радиусов из бесконечно удаленной точки, где они имели нулевую скорость. Тогда из закона сохранения энергии для наблюдателя массы m следует равенство

$$\frac{mv^2}{2} - \frac{GmM}{\rho} = 0, \tag{34.116}$$

где $\mathbf{v} = \{v^\mu\}$ – скорость наблюдателя. Отсюда вытекает, что на расстоянии ρ от начала координат у наблюдателя будет скорость

$$v^\mu = -\sqrt{\frac{2GM}{\rho}} \frac{x^\mu}{\rho}, \tag{34.117}$$

где x^μ/ρ – компоненты единичного радиального вектора, и знак минус означает, что наблюдатель движется к центру. Поскольку свободно падающие наблюдатели не

чувствуют гравитационного поля, то в связанной с ними системе отсчета метрика будет совпадать с метрикой Лоренца:

$$ds^2 = c^2(dt_F)^2 - (dx_F^1)^2 - (dx_F^2)^2 - (dx_F^3)^2, \quad (34.118)$$

где $x^0 := ct$ и c – скорость света. Индекс F обозначает, что система координат связана со свободно падающими наблюдателями.

Предположим, что свободно падающая система координат связана с декартовой системой координат в пространстве-времени, которую можно интерпретировать, как систему координат бесконечно удаленного наблюдателя, преобразованием Галилея (1.174):

$$\begin{aligned} dt_F &= dt, \\ dx_F^\mu &= dx^\mu + v^\mu dt. \end{aligned} \quad (34.119)$$

Тогда простые вычисления приводят к следующей метрике в системе координат бесконечно удаленного наблюдателя:

$$\begin{aligned} ds^2 &= \left(c^2 - \frac{2GM}{\rho}\right) dt^2 + 2\sqrt{\frac{2GM}{\rho}} \frac{1}{\rho} x^\mu dx_\mu dt + dx^\mu dx_\mu = \\ &= \left(c^2 - \frac{2GM}{\rho}\right) dt^2 - 2\sqrt{\frac{2GM}{\rho}} dt d\rho - d\rho^2 - \rho^2(d\theta^2 + \sin^2\theta d\varphi^2) = \\ &= c^2 dt^2 - \left(d\rho + \sqrt{\frac{2GM}{\rho}} dt\right)^2 - \rho^2(d\theta^2 + \sin^2\theta d\varphi^2), \end{aligned} \quad (34.120)$$

где мы перешли в сферическую систему координат t, ρ, θ, φ . Напомним, что в наших обозначениях $x^\mu x_\mu = -(x^1)^2 - (x^2)^2 - (x^3)^2 = -\rho^2$.

Метрика (34.120) есть метрика Шварцшильда в координатах Пенлеве–Гулстранда [268, 259, 269], которые были рассмотрены ранее (34.39).

При “выводе” метрики Шварцшильда (34.120) было использовано четыре предположения: закон всемирного тяготения (34.115); закон сохранения энергии (34.116); предположение о том, что в инерциальной системе координат метрика плоская (34.118), и “преобразование Галилея” (34.119). Все четыре постулата настолько привычны в нерелятивистской механике, что “вывод” решения Шварцшильда кажется безукоризненным. Однако это не так. Мы не случайно поставили кавычки у слова “вывод”, поскольку приведенные рассуждения некорректны по следующей причине. “Преобразование Галилея” (34.119) записано для дифференциалов, а не самих координат и является неинтегрируемым. А именно, второе уравнение в (34.119) эквивалентно системе уравнений в частных производных

$$\begin{aligned} \frac{\partial x_F^\mu}{\partial x^\nu} &= \delta_\nu^\mu, \\ \frac{\partial x_F^\mu}{\partial t} &= -\sqrt{\frac{2GM}{\rho}} \frac{x^\mu}{\rho}. \end{aligned}$$

Эта система уравнений нетривиальна, т.к. скорость v^μ нетривиально зависит от точки пространства Минковского. Легко проверить, что условия интегрируемости для

этой системы уравнений не выполнены. Действительно, из первого уравнения следует равенство

$$\frac{\partial}{\partial t} \frac{\partial x_F^\mu}{\partial x^\nu} = 0.$$

В то же время, дифференцируя второе равенство, получаем нетривиальный ответ

$$\frac{\partial}{\partial x^\nu} \frac{\partial x_F^\mu}{\partial t} \neq 0.$$

Следовательно, не существует таких функций $x_F^\mu(x, t)$, что для дифференциалов выполнены равенства (34.119) даже локально. В остальном “вывод” решения Шварцшильда безупречен.

Отсутствие преобразования координат такого, которое переводит метрику Лоренца (34.118) в метрику Шварцшильда, следует также из общих соображений. Метрика Лоренца является плоской и соответствующий ей тензор кривизны равен нулю. Для метрики Шварцшильда тензор Риччи, как следствие уравнений Эйнштейна, равен нулю. Полный же тензор кривизны нетривиален (34.3). Следовательно, метрика Лоренца не может быть связана с метрикой Шварцшильда никаким преобразованием координат.

С физической точки зрения, в “оправдание” эвристического “вывода” можно привести следующий аргумент. Вдали от притягивающего тела скорости малы, и свободно падающий наблюдатель в небольшой окрестности и в течении небольшого промежутка времени может приближенно считать, что находится в инерциальной системе координат, которая движется равномерно и прямолинейно. Тогда равенства (34.119) можно считать выполненными приближенно. Это показывает, что в данном случае правдоподобные рассуждения приводят к неправильному ответу.

Метрику Шварцшильда в координатах Пенлеве–Гулстранда (34.120) можно переписать в эквивалентном виде

$$ds^2 = c^2 dt^2 + \eta_{\mu\nu} (dx^\mu + v^\mu dt)(x^\nu + v^\nu dt), \quad (34.121)$$

где компоненты скорости v^μ были определены ранее (34.116). В таком виде метрика Шварцшильда конформно эквивалентна эффективной метрике для распространения фононов в жидкости (23.157), которая была получена в разделе 23.12.1. Разница заключается в том, что для эффективной метрики, скорость v_0 является решением уравнений гидродинамики (23.140)–(23.142), в то время как в решении Шварцшильда выражение для скорости v следует из уравнений Эйнштейна.

34.10.1 Радиус Шварцшильда

Нулевая компонента g_{00} метрики Шварцшильда (34.1) обращается в нуль на шварцшильдовском радиусе

$$\rho_s := \frac{2GM}{c^2}, \quad (34.122)$$

где мы восстановили гравитационную постоянную G и скорость света c . Двумерная поверхность (сфера) для всех $t = \text{const}$ и $\rho = \rho_s$ является координатной сингулярностью, т.к. инварианты, построенные из компонент тензора кривизны, не имеют здесь никаких особенностей. Сфера радиуса ρ_s в координатах Шварцшильда (34.1) является горизонтом, потому что свет, испущенный внутри черной дыры ($\rho < \rho_s$), никогда не выйдет за пределы сферы $\rho = \rho_s$.

Удивительно, но значение радиуса Шварцшильда можно получить в рамках классической нерелятивистской ньютоновой механики. С этой целью напомним вывод формулы для величины второй космической скорости. По определению, *второй космической скоростью* v называется минимальная скорость, которую необходимо сообщить телу массы m , чтобы оно смогло улететь на бесконечно большое расстояние, находясь на поверхности сферического тела радиуса R и массы M . Очевидно, что скорость будет минимальна, если тело будет двигаться вдоль радиуса. Значение второй космической скорости сразу следует из закона сохранения энергии:

$$\frac{mv^2}{2} - \frac{GmM}{R} = 0.$$

При этом мы предположили, что кинетическая и потенциальная энергии тела на бесконечности равны нулю. Отсюда следует выражение для второй космической скорости:

$$v^2 = \frac{2GM}{R}. \quad (34.123)$$

Это выражение не зависит от массы тела. Выражение для второй космической скорости (34.123) было использовано также раньше, при построении системы координат, связанной со свободно падающими наблюдателями (34.116).

Теперь предположим, что полученное выражение для второй космической скорости справедливо также при $m = 0$, т.е. для фотонов, масса которых равна нулю. Тогда постановку задачи можно изменить. А именно, каков минимальный радиус сферического тела массы M , при котором фотоны не смогут его покинуть (уйти на бесконечно большое расстояние)? Тогда $v^2 = c^2$, и из выражения для второй космической скорости сразу следует значение радиуса Шварцшильда (34.122).

Историческое замечание. Выражение для радиуса Шварцшильда (34.122) было получено задолго до создания общей теории относительности Реверендом Джоном Мичеллом (Reverend John Michell) в 1784 году [270]. По существу, он использовал тот же метод, который приведен в настоящем разделе. Позже этот вывод был повторен Пьером Симоном Маркизом де Лапласом (Pierre Simon Marquis de Laplace) в 1796 году [271]. Перевод статьи Лапласа приведен в [272]. Мичелл в свое время был известным геологом, однако сделал вклад также в теорию гравитации. Еще одним его известным изобретением являются крутильные весы, которые позже использовал Кавендиш в своих экспериментах по измерению гравитационной постоянной. Исторические связи между Мичеллом, Кавендишем и Лапласом описаны в [273]. $\square$

Остановимся на физической интерпретации. Фотоны, находящиеся на поверхности тела радиуса $R < \rho_s$ могут удалиться от него на конечное расстояние, а потом вернуться. Согласно корпускулярной теории света, которая была распространена в XVIII веке, фотон рассматривается, как точечная частица. При этом по мере удаления от тела фотон теряет свою скорость, достигает точки поворота, в которой его скорость обращается в нуль, и возвращается обратно.

С квантовомеханической точки зрения фотон представляет собой квант электромагнитной волны с энергией E и частотой ω , которые связаны соотношением

$$E = \hbar\omega,$$

где $\hbar$ – постоянная Планка. Скорость распространения фотонов постоянна и равна c . При такой интерпретации частота фотона по мере удаления от тела уменьшается

(красное смещение) и в точке поворота становится равной нулю. При этом мы вынуждены предположить, что скорость фотона в точке поворота имеет скачок, т.к. меняется на противоположную.

Существует качественное отличие ньютоновой теории тяготения и общей теории относительности. В рамках ньютоновой механики горизонтом событий для сферически симметричного тела радиуса ρ_s является бесконечно удаленная сфера $\rho = \infty$. В то же время в общей теории относительности фотоны не могут покинуть область конечного радиуса $\rho < \rho_s$. Тем не менее выражение для радиуса Шварцшильда получено то же, что и в общей теории относительности.

Глава 35

Космология

В настоящей главе мы рассмотрим простейшие космологические модели в рамках общей теории относительности. Это направление исследований составляет классический раздел теории гравитации и в настоящее время стало очень актуальным в связи с накоплением большого количества наблюдательных данных.

Модели вселенной, которые рассматриваются в настоящей главе, предполагают, что вся вселенная заполнена сплошной средой (жидкостью или газом). В начале главы мы рассмотрим кинематические свойства сплошной среды в общей теории относительности. Будет получено, в частности, уравнение Райчаудхури, из которого следует, что при достаточно общих предположениях космологические решения уравнений Эйнштейна должны содержать особенности: большой взрыв или большое сжатие.

Основные космологические модели основаны на космологическом принципе, согласно которому вселенная однородна и изотропна. Для математической формулировки требований к таким моделям необходимо довольно подробное знание свойств трехмерных пространств постоянной кривизны с положительно определенной метрикой. Одно из этих пространств – это евклидово пространство $\mathbb{R}^3$ нулевой кривизны, которое было рассмотрено в разделе 1.8. Поэтому в разделах 35.2.1 и 35.2.2 мы рассмотрим трехмерную сферу S^3 (риманово пространство положительной кривизны) и двуполостный гиперboloид H^3 , который иногда называют псевдосферой, (риманово пространство отрицательной кривизны). Эти сведения необходимы для построения моделей, которые рассмотрены в последующих разделах.

В заключительных разделах рассмотрена классификация Бианки трехмерных алгебр Ли и модели однородных вселенных.

Заметим, что если евклидово пространство $\mathbb{R}^3$ заполнено веществом с однородной плотностью, то ньютонов гравитационный потенциал равен нулю. Это значит, что при отсутствии других взаимодействий однородная модель вселенной в механике Ньютона является статической. В общей теории относительности, как мы увидим ниже, ситуация другая: однородные и изотропные модели вселенной нестатичны, что хорошо согласуется с наблюдательными данными.

35.1 Сплошная среда в космологии

При построении космологических моделей в рамках общей теории относительности материя рассматривается как сплошная среда. В настоящем разделе мы рассмотрим кинематику сплошной среды и получим уравнение Райчаудхури. Из этого уравнения

вытекает, в частности, что если движение материи безвихревое, а плотность энергии и давление среды положительны, то в космологических моделях неизбежно имеет место либо большой взрыв, либо большое сжатие. Этот вывод важен и является довольно общим, т.к. не использует явный вид решений уравнений Эйнштейна.

35.1.1 Сплошная среда в классической механике

Напомним некоторые сведения из кинематики сплошной среды в классической механике.

Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовой системой координат x^μ , $\mu = 1, 2, 3$. Основной переменной в механике сплошной среды (жидкости или газа) является достаточно гладкое векторное поле скорости $u^\mu(t, \mathbf{x})$, зависящее от времени $t \in \mathbb{R}$ и точки пространства $\mathbf{x} \in \mathbb{R}^3$. Мы предполагаем, что система уравнений

$$\frac{dx^\mu(t, \mathbf{y})}{dt} = u^\mu(t, \mathbf{x}), \quad (35.1)$$

с начальным условием $x^\mu(t = 0, \mathbf{y}) = y^\mu$ имеет единственное решение при всех временах $t \in \mathbb{R}$ и при всех $\mathbf{y} \in \mathbb{R}^3$. Для единственности достаточно, чтобы правая часть системы уравнений (35.1) удовлетворяла условию Липшица (теорема 2.6.5). Полнота решений предполагается. Каждое решение данной системы уравнений определяет линию тока $x^\mu(t, \mathbf{y})$, вдоль которой движется частица среды, которая проходит через каждую точку пространства $\mathbf{y}$ в начальный момент времени $t = 0$. Линии тока не могут пересекаться, т.к. это противоречит единственности решений системы обыкновенных дифференциальных уравнений (35.1). Другими словами, линии тока заполняют все пространство $\mathbb{R}^3$, не имея лакун, и нигде не пересекаются.

Рассмотрим две близкие точки среды

$$p = \{x_p^\mu\}, \quad q = \{x_q^\mu := x_p^\mu + dx_q^\mu\}.$$

Скорости в этих точках в момент времени t равны

$$u_p = \{u_p^\mu\}, \quad u_q = \{u_q^\mu = u_p^\mu + \partial_\nu u^\mu|_p dx_q^\nu + O(dx_q^2)\}.$$

Если эти точки движутся вдоль линий тока, то через время dt они будут иметь новые координаты:

$$\begin{aligned} p &\mapsto \{x_p^\mu + u_p^\mu dt + O(dt^2, dx_q^2)\}, \\ q &\mapsto \{x_p^\mu + dx_q^\mu + u_p^\mu dt + \partial_\nu u^\mu|_p dx_q^\nu dt + O(dt^2, dx_q^2)\}. \end{aligned}$$

Разность декартовых координат конечных точек имеет вид

$$dx_q^\mu(t + dt) := x_q^\mu(t + dt) - x_p^\mu(t + dt) = \left(\delta_\nu^\mu + \partial_\nu u^\mu|_p dt \right) dx_q^\nu + O(dt^2, dx_q^2).$$

Таким образом, частные производные от поля скоростей $\partial_\nu u^\mu$ имеют смысл относительной скорости двух соседних частиц среды.

Разложим тензор второго ранга $\partial_\mu u_\nu$, где $u_\nu := u^\mu \delta_{\nu\mu}$, на неприводимые компоненты, выделив из него след θ и антисимметричную часть $\omega_{\mu\nu} = -\omega_{\nu\mu}$:

$$\partial_\mu u_\nu = \sigma_{\mu\nu} + \omega_{\mu\nu} + \frac{1}{3} \delta_{\mu\nu} \theta, \quad (35.2)$$

где

$$\begin{aligned}\theta &:= \partial_\mu u^\mu && - \text{скаляр расширения,} \\ \sigma_{\mu\nu} &:= \frac{1}{2}(\partial_\mu u_\nu + \partial_\nu u_\mu) - \frac{1}{3}\delta_{\mu\nu}\theta && - \text{тензор сдвига,} \\ \omega_{\mu\nu} &:= \frac{1}{2}(\partial_\mu u_\nu - \partial_\nu u_\mu) && - \text{тензор вращения.}\end{aligned}\tag{35.3}$$

Каждая неприводимая компонента имеет физический смысл и название по следующим причинам.

- Пусть $\sigma_{\mu\nu} = 0$ и $\omega_{\mu\nu} = 0$. Тогда смещение точки q относительно точки p в первом порядке имеет вид

$$dx_q^\mu(t+dt) = \left(1 + \frac{1}{3}\theta dt\right) dx_q^\mu(t).$$

То есть вектор $dx_q^\mu(t+dt)$ имеет то же направление, что и $dx_q^\mu(t)$, но другую длину. Такое движение среды является изотропным расширением ($\theta > 0$) или сжатием ($\theta < 0$).

- Пусть $\sigma_{\mu\nu} = 0$ и $\theta = 0$. Тогда

$$dx_q^\mu(t+dt) = dx_q^\mu(t) + \omega_\nu{}^\mu dt dx_q^\nu(t).$$

Ввиду антисимметрии компоненты $\omega_{\mu\nu}$, при таком движении вектор $dx_q^\mu(t)$ поворачивается, не изменяя своей длины. Такое движение соответствует вращению линий тока.

- Пусть $\omega_{\mu\nu} = 0$ и $\theta = 0$. Рассмотрим три точки q_1 , q_2 и q_3 , которые близки к точке p :

$$q_1 = p + dx_1, \quad q_2 = p + dx_2, \quad q_3 = p + dx_3.$$

Объем параллелепипеда со сторонами dx_1 , dx_2 и dx_3 равен

$$dV = \varepsilon_{\mu\nu\rho} dx_1^\mu dx_2^\nu dx_3^\rho,$$

где $\varepsilon_{\mu\nu\rho}$ — полностью антисимметричный тензор третьего ранга (см. приложение 37.4). Через время dt вершины параллелепипеда сместятся, и его объем в первом приближении станет равным

$$dV' = dV + Ddt,$$

где

$$D := (\varepsilon_{\mu\nu\sigma}\sigma^\sigma{}_\rho + \varepsilon_{\mu\sigma\rho}\sigma^\sigma{}_\nu + \varepsilon_{\sigma\nu\rho}\sigma^\sigma{}_\mu) dx_1^\mu dx_2^\nu dx_3^\rho.$$

Выражение в скобках антисимметрично по индексам ν, ρ . Поэтому его можно свернуть с $\varepsilon^{\lambda\nu\rho}$ без потери информации. В результате получим нуль, что следует из свойств тензора деформаций: $\sigma_{\mu\nu} = \sigma_{\nu\mu}$ и $\sigma^\mu{}_\mu = 0$. Следовательно, $D = 0$. Это означает, что объем параллелепипеда не меняется. Может измениться только его форма. Такое движение называется сдвигом.

Заметим, что при вращении линий тока длина сторон параллелепипеда сохраняется, и поэтому сохраняется его объем. Более того, при вращении сохраняется не только объем, но и форма параллелепипеда, поскольку сохраняются также углы между векторами.

Тензор вращений $\omega_{\mu\nu}$ взаимно однозначно параметризуется вектором вращений

$$\omega = \left\{ \omega^\mu := \frac{1}{2} \varepsilon^{\mu\nu\rho} \omega_{\nu\rho} \right\}. \quad (35.4)$$

Из свойств полностью антисимметричного тензора третьего ранга следует равенство

$$\omega^2 := \omega^\mu \omega_\mu = \frac{1}{2} \omega^{\mu\nu} \omega_{\mu\nu} \geq 0.$$

Вектор вращений направлен вдоль оси поворота, и его длина равна углу поворота.

35.1.2 Сплошная среда в релятивистской механике

В релятивистской механике мы поступаем по аналогии с тем, что было проделано в предыдущем разделе для сплошной среды в классической механике.

Рассмотрим топологически тривиальное псевдориманово многообразие $\mathbb{M} \approx \mathbb{R}^4$, которое покрыто одной картой x^α , $\alpha = 0, 1, 2, 3$. На нем задана метрика $g_{\alpha\beta}$ лоренцевой сигнатуры. Будем считать, что координата x^0 является временем, и все сечения $x^0 = \text{const}$ пространственноподобны.

Основной переменной в релятивистской механике сплошной среды является времениподобное векторное поле скорости $u^\alpha(x)$, $u^2 > 0$. Как и в классической механике, будем считать векторное поле скорости достаточно гладким, и что система уравнений для линий тока

$$\frac{dx^\alpha(s, y)}{ds} = u^\alpha(x)$$

с начальными условиями $x^\alpha(s = 0, y) = y^\alpha$ имеет единственное решение при всех $y \in \mathbb{M}$, определенное для всех значений параметра $s \in \mathbb{R}$. Решение данной системы уравнений $x^\alpha(s, y)$ называется *линией тока* частицы жидкости, проходящей через точку $y \in \mathbb{M}$. Как и в случае нерелятивистской среды линии тока заполняют все многообразие $\mathbb{M}$ без каких-либо лакун и самопересечений. Предположим также, что в качестве параметра s вдоль линий тока выбрана ее длина:

$$ds^2 = dx^\alpha dx^\beta g_{\alpha\beta}.$$

Тогда векторное поле скорости будет иметь единичную длину

$$u^\alpha u^\beta g_{\alpha\beta} = 1. \quad (35.5)$$

Дифференцирование этого равенства приводит к тождеству

$$u^\beta \nabla_\alpha u_\beta = 0, \quad (35.6)$$

где ∇_α – ковариантная производная, построенная по символам Кристоффеля для метрики $g_{\alpha\beta}$. Если уравнение (35.6) продифференцировать еще раз, то получим равенство

$$u^\gamma \nabla_\alpha \nabla_\beta u_\gamma + \nabla_\alpha u^\gamma \nabla_\beta u_\gamma = 0, \quad (35.7)$$

которое позволяет выразить вторые производные от вектора скорости через первые (при суммировании по индексу γ).

Наличие векторного поля скорости позволяет построить проекционные операторы:

$$\Pi_\alpha^\tau := \delta_\alpha^\beta - u_\alpha u^\beta, \quad \Pi_\alpha^\perp := u_\alpha u^\beta, \quad (35.8)$$

где $u_\alpha := u^\beta g_{\beta\alpha}$. Операторы Π^T и Π^L проектируют тензоры на направления, перпендикулярные и параллельные вектору скорости поскольку справедливы следующие равенства:

$$\Pi_\alpha^T u_\beta = 0, \quad \Pi_\alpha^L u_\beta = u_\beta.$$

Отметим также формулы

$$\Pi_{\alpha\beta}^T \Pi^{\alpha\beta} = 3, \quad \Pi_{\alpha\beta}^L \Pi^{\alpha\beta} = 1,$$

которые будут использованы в дальнейшем.

Последующие построения аналогичны тем, которые были проделаны в предыдущем разделе, но с заменой частных производных на ковариантные, $\partial_\alpha \mapsto \nabla_\alpha$.

Рассмотрим тензор второго ранга $\nabla_\alpha u_\beta$. Если метрика $g_{\alpha\beta}$ близка к метрике Лоренца, и система координат мало отличается от декартовой, то этот тензор имеет тот же физический смысл относительной скорости соседних частиц, что и в классической механике, как было описано в предыдущем разделе.

Спроектируем ковариантные производные векторного поля скорости на перпендикулярную гиперповерхность:

$$\Pi_\alpha^T \Pi_\beta^T \nabla_\gamma u_\delta = \Pi_\alpha^T \nabla_\gamma u_\beta = \nabla_\alpha u_\beta - u_\alpha \dot{u}_\beta, \quad (35.9)$$

где мы воспользовались равенством (35.6), и точка обозначает дифференцирование вдоль поля скорости:

$$\dot{u}_\beta := u^\gamma \nabla_\gamma u_\beta. \quad (35.10)$$

Это – компоненты ускорения мировых линий частиц непрерывной среды. Если $\dot{u} = 0$, то частицы непрерывной среды движутся вдоль экстремалей, которые в данном случае совпадают с геодезическими (см. раздел 16.1).

Напомним, что если градиент давления $\mathcal{P}$ в жидкости направлен вдоль линий тока, т.е.

$$\Pi_\alpha^T \partial_\beta \mathcal{P} = 0,$$

то, согласно уравнениям релятивистской гидродинамики (23.135), (23.136), частицы жидкости движутся вдоль экстремалей, хоть и не являются в общем случае свободными, т.к. подвержены давлению.

Как и ранее, выделим из тензора (35.9) след и антисимметричную часть:

$$\Pi_\alpha^T \Pi_\beta^T \nabla_\gamma u_\delta = \sigma_{\alpha\beta} + \omega_{\alpha\beta} + \frac{1}{3} \Pi_{\alpha\beta}^T \theta, \quad (35.11)$$

где

$$\begin{aligned} \theta &:= \Pi^{\alpha\beta} \nabla_\alpha u_\beta = \nabla_\alpha u^\alpha && - \text{скаляр расширения,} \\ \sigma_{\alpha\beta} &:= \frac{1}{2} (\Pi_\alpha^T \nabla_\gamma u_\beta + \Pi_\beta^T \nabla_\gamma u_\alpha) - \frac{1}{3} \Pi_{\alpha\beta}^T \theta && - \text{тензор сдвига,} \\ \omega_{\alpha\beta} &:= \frac{1}{2} (\Pi_\alpha^T \nabla_\gamma u_\beta - \Pi_\beta^T \nabla_\gamma u_\alpha) && - \text{тензор вращения.} \end{aligned} \quad (35.12)$$

Используя производную вдоль векторного поля скорости (35.10), выражения для тензоров сдвига и вращения можно переписать в следующем виде

$$\begin{aligned} \sigma_{\alpha\beta} &:= \frac{1}{2} (\nabla_\alpha u_\beta + \nabla_\beta u_\alpha - u_\alpha \dot{u}_\beta - u_\beta \dot{u}_\alpha) - \frac{1}{3} \Pi_{\alpha\beta}^T \theta, \\ \omega_{\alpha\beta} &:= \frac{1}{2} (\nabla_\alpha u_\beta - \nabla_\beta u_\alpha - u_\alpha \dot{u}_\beta + u_\beta \dot{u}_\alpha). \end{aligned}$$

В дальнейшем нам понадобится выражение для полной ковариантной производной ковекторного поля скорости

$$\nabla_\alpha u_\beta = \Pi_\alpha^\tau \nabla_\gamma u_\beta + \Pi_\alpha^L \nabla_\gamma u_\beta = \sigma_{\alpha\beta} + \omega_{\alpha\beta} + \frac{1}{3} \Pi_{\alpha\beta}^\tau \theta + u_\alpha \dot{u}_\beta. \quad (35.13)$$

Отметим также полезные для вычислений формулы:

$$\sigma_{\alpha\beta} u^\beta = 0, \quad \omega_{\alpha\beta} u^\beta = 0,$$

которые просто проверяются.

Тензор вращения можно параметризовать псевдовекторным полем

$$\omega^\alpha := \frac{1}{2} \varepsilon^{\alpha\beta\gamma\delta} u_\beta \omega_{\gamma\delta}, \quad (35.14)$$

где $\varepsilon^{\alpha\beta\gamma\delta}$ — полностью антисимметричный тензор четвертого ранга (см. приложение 37.4). Квадрат этого псевдовектора имеет вид

$$\omega^2 := \omega^\alpha \omega_\alpha = \frac{1}{2} \omega^{\alpha\beta} \omega_{\alpha\beta}.$$

Псевдовектор вращения времени подобен или нулевой, т.к. правая часть неотрицательна. Это проще всего увидеть в системе координат, сопутствующей векторному полю скорости $u = (1, 0, 0, 0)$, в которой

$$\Pi^{\tau 00} = \Pi^{\tau 0\mu} = \Pi^{\tau\mu 0} = 0, \quad \Pi^{\tau\mu\nu} = g^{\mu\nu}, \quad \mu, \nu = 1, 2, 3.$$

Из равенства

$$\omega^{\alpha\beta} = \Pi^{\tau\alpha\gamma} \Pi^{\tau\beta\delta} \omega_{\gamma\delta}$$

следует, что в сопутствующей системе координат

$$\omega^{00} = \omega^{0\mu} = \omega^{\mu 0} = 0.$$

Поэтому

$$\omega^2 = \frac{1}{2} \omega^{\mu\nu} \omega_{\mu\nu} \geq 0$$

в силу отрицательной определенности пространственной части метрики $g_{\mu\nu}$.

Кроме этого введем скаляр деформации $\sigma > 0$ с помощью равенства

$$\sigma^2 := \frac{1}{2} \sigma^{\alpha\beta} \sigma_{\alpha\beta} \geq 0. \quad (35.15)$$

Он неотрицателен по той же причине, что и скаляр вращения.

Для того, чтобы оправдать предложенные выше построения, рассмотрим нерелятивистский предел. Для этого заменим метрику $g_{\alpha\beta}$ на метрику Лоренца $\eta_{\alpha\beta}$ и введем наблюдаемое время $x^0 = ct$, где c — скорость света. В этом случае ковариантные производные совпадают с частными. В первом приближении можно считать, что длина линии тока совпадает с x^0 . Тогда

$$u^0 = u_0 = 1, \quad v^\mu := \frac{dx^\mu}{dt} = cu^\mu.$$

Для отличия трехмерную скорость частицы среды мы обозначили новой буквой v^μ (в предыдущем разделе она была обозначена через u^μ). Кроме того, в нерелятивистском пределе $v^2/c^2 \rightarrow 0$ имеют место формулы:

$$\partial_\mu u_0 \rightarrow 0, \quad \partial_0 u_\mu = \frac{1}{c^2} \frac{dv_\mu}{dt} \rightarrow 0.$$

Поэтому определение расширения, сдвига и вращения (35.12) в релятивистском случае сводится к соответствующим нерелятивистским компонентам:

$$\begin{aligned} c\theta &\rightarrow \theta_N, & c\sigma_{\alpha\beta} &\rightarrow \sigma_{N\mu\nu}, & c\omega_{\mu\nu} &\rightarrow \omega_{N\mu\nu}, \\ \sigma_{00}, \sigma_{0\mu} &\rightarrow 0, & \omega_{00}, \omega_{0\mu} &\rightarrow 0, \end{aligned}$$

где неприводимые компоненты в нерелятивистском случае помечены индексом N.

35.1.3 Уравнение Райчаудхури

Получим эволюционные уравнения вдоль линий тока для скаляра расширения и тензоров сдвига и вращения. С этой целью два раза продифференцируем вектор скорости и антисимметризуем:

$$\nabla_\epsilon \nabla_\delta u_\gamma - \nabla_\delta \nabla_\epsilon u_\gamma = -R_{\epsilon\delta\gamma\zeta} u^\zeta, \quad (35.16)$$

где мы воспользовались формулой для коммутатора ковариантных производных (6.94). Умножим это уравнение на u^ϵ . Тогда первое и второе слагаемые примут вид

$$\begin{aligned} u^\epsilon \nabla_\epsilon \nabla_\delta u_\gamma &:= (\nabla_\delta u_\gamma)^\cdot, \\ -u^\epsilon \nabla_\delta \nabla_\epsilon u_\gamma &= -\nabla_\delta \dot{u}_\gamma + \nabla_\delta u^\epsilon \nabla_\epsilon u_\gamma, \end{aligned}$$

где $\dot{u}_\gamma$ – компоненты ускорения линий тока (35.10). Теперь спроектируем полученное уравнение на перпендикулярные направления:

$$\Pi_\alpha^\gamma \Pi_\beta^{\tau\delta} (\nabla_\delta u_\gamma)^\cdot - \Pi_\alpha^\gamma \Pi_\beta^{\tau\delta} \nabla_\delta \dot{u}_\gamma + \Pi_\alpha^\gamma \Pi_\beta^{\tau\delta} \nabla_\delta u^\epsilon \nabla_\epsilon u_\gamma - R_{\alpha\epsilon\beta\zeta} u^\epsilon u^\zeta = 0, \quad (35.17)$$

где мы воспользовались антисимметрией тензора кривизны:

$$\Pi_\alpha^\gamma \Pi_\beta^{\tau\delta} R_{\gamma\epsilon\delta\zeta} u^\epsilon u^\zeta = R_{\alpha\epsilon\beta\zeta} u^\epsilon u^\zeta.$$

Эволюционное уравнение (35.17) является кинематическим в том смысле, что при его получении мы не использовали уравнений Эйнштейна.

Чтобы получить эволюционные уравнения, разложим уравнение (35.17) на неприводимые компоненты. Оно представляет собой тензор второго ранга с индексами α и β . Сначала выделим след, свернув его с обратной метрикой $g^{\alpha\beta}$. Тогда первое, второе и третье слагаемые примут вид:

$$\begin{aligned} g^{\alpha\beta} \Pi_\alpha^\gamma \Pi_\beta^{\tau\delta} (\nabla_\delta u_\gamma)^\cdot &= \dot{\theta} + \dot{u}^\gamma \dot{u}_\gamma, \\ -g^{\alpha\beta} \Pi_\alpha^\gamma \Pi_\beta^{\tau\delta} \nabla_\delta \dot{u}_\gamma &= -\nabla_\gamma \dot{u}^\gamma - \dot{u}^\gamma \dot{u}_\gamma, \\ g^{\alpha\beta} \Pi_\alpha^\gamma \Pi_\beta^{\tau\delta} \nabla_\delta u^\epsilon \nabla_\epsilon u_\gamma &= 2\sigma^2 - 2\omega^2 + \frac{1}{3}\theta^2. \end{aligned} \quad (35.18)$$

Для преобразования четвертого слагаемого используем уравнения Эйнштейна, которые запишем в виде

$$R_{\alpha\beta} = -\frac{1}{2\kappa} \left[(\mathcal{E} + \mathcal{P})u_\alpha u_\beta - \frac{1}{2}(\mathcal{E} - \mathcal{P})g_{\alpha\beta} \right].$$

Тогда четвертое слагаемое примет вид

$$R_{\alpha\beta}u^\alpha u^\beta = -\frac{1}{4\kappa}(\mathcal{E} + 3\mathcal{P}).$$

Суммируя все вместе, получаем *уравнение Райчаудхури*

$$\dot{\theta} + \frac{1}{3}\theta^2 - \nabla_\gamma \dot{u}^\gamma + 2(\sigma^2 - \omega^2) + \frac{1}{4\kappa}(\mathcal{E} + 3\mathcal{P}) = 0. \quad (35.19)$$

Это уравнение было получено Райчаудхури [274, 275] для пыли, $\mathcal{P} = 0$, без определения тензоров сдвига и вращений. В форме (35.19) оно было получено позже [276].

Теперь выделим из уравнения (35.17) антисимметричную часть. Простые вычисления дают следующие выражения для первого, второго и третьего слагаемого:

$$\begin{aligned} \Pi_{[\alpha}^T \gamma \Pi_{\beta]}^T \delta (\nabla_\delta u_\gamma)^\cdot &= \Pi_\alpha^T \gamma \Pi_\beta^T \delta \dot{\omega}_{\delta\gamma}, \\ -\Pi_{[\alpha}^T \gamma \Pi_{\beta]}^T \delta \nabla_\delta \dot{u}_\gamma &= -\Pi_\alpha^T \gamma \Pi_\beta^T \delta \nabla_{[\delta} \dot{u}_{\gamma]}, \\ \Pi_{[\alpha}^T \gamma \Pi_{\beta]}^T \delta \nabla_\delta u^\epsilon \nabla_\epsilon u_\gamma &= 2\omega_{[\beta} \gamma \sigma_{\alpha]\gamma} + \frac{2}{3}\theta\omega_{\beta\alpha}, \end{aligned}$$

где квадратные скобки обозначают антисимметризацию по паре индексов. Четвертое слагаемое при антисимметризации дает нуль, поскольку тензор кривизны симметричен относительно перестановки первой и второй пар индексов. В результате получаем эволюционное уравнение для тензора вращений:

$$\Pi_\alpha^T \gamma \Pi_\beta^T \delta \dot{\omega}_{\delta\gamma} - \Pi_\alpha^T \gamma \Pi_\beta^T \delta \nabla_{[\delta} \dot{u}_{\gamma]} + 2\omega_{[\beta} \gamma \sigma_{\alpha]\gamma} + \frac{2}{3}\theta\omega_{\beta\alpha} = 0. \quad (35.20)$$

Сейчас построим бесследовую симметричную часть уравнения (35.17). Первые три слагаемых имеют вид

$$\begin{aligned} \Pi_{(\alpha}^T \gamma \Pi_{\beta)}^T \delta (\nabla_\delta u_\gamma)^\cdot - \frac{1}{3}(\dot{\theta} + \dot{u}^\gamma \dot{u}_\gamma) \Pi_{\alpha\beta}^T &= \Pi_\alpha^T \gamma \Pi_\beta^T \delta \dot{\sigma}_{\delta\gamma} - \frac{1}{3}\dot{u}^\gamma \dot{u}_\gamma \Pi_{\alpha\beta}^T, \\ -\Pi_{(\alpha}^T \gamma \Pi_{\beta)}^T \delta \nabla_\delta \dot{u}_\gamma + \frac{1}{3}(\nabla_\gamma \dot{u}^\gamma + \dot{u}^\gamma \dot{u}_\gamma) \Pi_{\alpha\beta}^T &= -\Pi_\alpha^T \gamma \Pi_\beta^T \delta \nabla_{(\delta} \dot{u}_{\gamma)} + \frac{1}{3}(\nabla_\gamma \dot{u}^\gamma + \dot{u}^\gamma \dot{u}_\gamma) \Pi_{\alpha\beta}^T, \\ \Pi_{(\alpha}^T \gamma \Pi_{\beta)}^T \delta \nabla_\delta u^\epsilon \nabla_\epsilon u_\gamma - \frac{1}{3}\left(2\sigma^2 - 2\omega^2 + \frac{1}{3}\theta^2\right) \Pi_{\alpha\beta}^T &= \sigma_\beta^\epsilon \sigma_{\epsilon\alpha} + \omega_\beta^\epsilon \omega_{\epsilon\alpha} + \\ &\quad + \frac{2}{3}\theta\sigma_{\beta\alpha} - \frac{1}{3}(2\sigma^2 - 2\omega^2) \Pi_{\alpha\beta}^T, \end{aligned}$$

где скобки обозначают симметризацию по паре индексов, и использованы формулы для следов (35.18). Четвертое слагаемое равно

$$\begin{aligned} R_{\alpha\gamma\beta\delta}u^\gamma u^\delta - \frac{1}{3}R_{\gamma\delta}u^\gamma u^\delta \Pi_{\alpha\beta}^T &= \\ = C_{\alpha\gamma\beta\delta}u^\gamma u^\delta + \frac{1}{2}(R_{\alpha\beta} - R_{\alpha\gamma}u^\gamma u_\beta - R_{\beta\gamma}u^\gamma u_\alpha) + \\ + \frac{1}{6}R_{\gamma\delta}u^\gamma u^\delta g_{\alpha\beta} - \frac{1}{6}R(g_{\alpha\beta} - u_\alpha u_\beta) + \frac{1}{3}R_{\gamma\delta}u^\gamma u^\delta u_\alpha u_\beta, \end{aligned} \quad (35.21)$$

где мы разложили тензор кривизны на неприводимые компоненты, выделив тензор Вейля $C_{\alpha\beta\gamma\delta}$, тензор Риччи $R_{\alpha\beta}$ и скалярную кривизну R (см. раздел 17.1). Из уравнений Эйнштейна следует, что тензор Риччи имеет следующую структуру

$$R_{\alpha\beta} = A g_{\alpha\beta} + B u_\alpha u_\beta \quad \Rightarrow \quad R = 4A + B, \quad (35.22)$$

где A и B – некоторые функции. Если подставить это выражение в правую часть равенства (35.21), то все слагаемые сократятся за исключением тензора Вейля. Следовательно, эволюционное уравнение для тензора сдвига принимает вид

$$\begin{aligned} \Pi_{\alpha}^{\tau\gamma}\Pi_{\beta}^{\tau\delta}\dot{\sigma}_{\delta\gamma} - \Pi_{\alpha}^{\tau\gamma}\Pi_{\beta}^{\tau\delta}\nabla_{(\delta}\dot{u}_{\gamma)} + \sigma_{\alpha\gamma}\sigma^{\gamma}_{\beta} + \omega_{\alpha\gamma}\omega^{\gamma}_{\beta} + \frac{2}{3}\theta\sigma_{\alpha\beta} + \\ + \frac{1}{3}[\nabla_{\gamma}\dot{u}^{\gamma} - 2(\sigma^2 - \omega^2)]\Pi_{\alpha\beta} - \frac{1}{4}E_{\alpha\beta} = 0, \end{aligned} \quad (35.23)$$

где

$$E_{\alpha\beta} := 4C_{\alpha\gamma\beta\delta}u^{\gamma}u^{\delta} \quad (35.24)$$

– электрическая компонента тензора Вейля (17.95).

Таким образом, получены эволюционные уравнения (35.19), (35.20) и (35.23) для всех трех неприводимых компонент тензора $\nabla_{\alpha}u_{\beta}$. Эти уравнения эквивалентны системе уравнений (35.17).

Из исходных уравнений (35.16) можно получить дополнительные независимые уравнения. Получим еще три (но не все) алгебраически независимых уравнения. Поскольку антисимметризация тензора кривизны по любым трем индексам дает нуль, то полная антисимметризация исходного уравнения (35.16) приводит к равенству

$$\nabla_{[\alpha}\omega_{\beta\gamma]} + \omega_{[\alpha\beta}\dot{u}_{\gamma]} + u_{[\alpha}\nabla_{\beta}\dot{u}_{\gamma]} = 0, \quad (35.25)$$

где квадратные скобки обозначают антисимметризацию по трем индексам. Полученное уравнение не зависит от тензора кривизны и, следовательно, от полей материи. Поэтому оно имеет кинематический характер.

Теперь свернем уравнение (35.16) с $g^{\gamma\delta}$:

$$\nabla_{\epsilon}\nabla_{\gamma}u^{\gamma} - \nabla_{\gamma}\nabla_{\epsilon}u^{\gamma} - R_{\epsilon\zeta}u^{\zeta} = 0.$$

Если тензор Риччи имеет структуру (35.22), то последнее слагаемое обращается в нуль. После проектирования на перпендикулярное направление, т.е. умножения на $\Pi^{\tau\alpha\epsilon}$, возникает второе уравнение

$$\Pi^{\tau\alpha\epsilon}\left(\frac{2}{3}\nabla_{\epsilon}\theta - \nabla_{\gamma}\sigma_{\epsilon}^{\gamma} - \nabla_{\gamma}\omega_{\epsilon}^{\gamma}\right) - (\sigma_{\gamma}^{\alpha} + \omega_{\gamma}^{\alpha})\dot{u}^{\gamma} = 0. \quad (35.26)$$

Хотя оно не зависит от тензора кривизны, полученное уравнение нельзя назвать кинематическим, т.к. при его получении использована структура тензора Риччи (35.22).

Еще одно уравнение возникает после поднятия индексов ϵ, δ в уравнении (35.16) и умножения его на тензор

$$\frac{1}{2}\varepsilon_{\eta\zeta\epsilon\delta}u^{\zeta}\Pi_{\alpha}^{\tau\eta}\Pi_{\beta}^{\tau\gamma}.$$

После несложных преобразований получим уравнение

$$2\omega_{(\alpha}\dot{u}_{\beta)} - \Pi_{\alpha}^{\tau\gamma}\Pi_{\beta}^{\tau\delta}u^{\epsilon}\varepsilon_{\epsilon\eta\zeta(\gamma}\nabla^{\eta}[\sigma^{\zeta}_{\delta)} + \omega^{\zeta}_{\delta)}] = -\frac{1}{4}H_{\alpha\beta}, \quad (35.27)$$

где

$$H_{\alpha\beta} := 4 * C_{\alpha\beta\gamma\delta}u^{\beta}u^{\delta} = 2\varepsilon_{\alpha\gamma\delta\epsilon}C^{\delta\epsilon}_{\beta\zeta}u^{\gamma}u^{\zeta} \quad (35.28)$$

– магнитная составляющая тензора Вейля (17.96) (звездочка обозначает дуальный тензор).

Значение полученных эволюционных уравнений (35.19), (35.20), (35.23), (35.25)–(35.27) заключается в том, что они приводят к важным следствиям, если на кинематические характеристики среды наложены ограничения.

Пример 35.1.1. Пусть $\sigma_{\alpha\beta} = 0$, $\omega_{\alpha\beta} = 0$ и $\dot{u}^\alpha = 0$. Тогда уравнения (35.20) и (35.25) тождественно удовлетворяются. Из уравнения (35.26) следует, что скаляр расширения может меняться только вдоль линий тока. Уравнения (24.117) и (35.27) приводят к равенству нулю тензора Вейля. $\square$

35.1.4 Большой взрыв и большое сжатие

Рассмотрим вселенную, которая заполнена непрерывной средой. Допустим, что частицы среды движутся по экстремалиям пространства-времени, т.е. ускорение линий тока равно нулю, $\dot{u} = 0$. Предположим также, что тензор вращения при движении равен нулю, $\omega = 0$.

Теперь введем функцию $l(s)$ как решение обыкновенного дифференциального уравнения вдоль линий тока

$$\frac{\dot{l}}{l} := \frac{1}{3}\theta,$$

где $\dot{l} := u^\alpha \partial_\alpha l$. Тогда уравнение Райчаудхури (35.19) примет вид

$$3\frac{\ddot{l}}{l} + 2\sigma^2 + \frac{1}{4\kappa}(\mathcal{E} + 3\mathcal{P}) = 0. \quad (35.29)$$

Для всех известных в настоящее время видов “нормальной” материи $\mathcal{E} > 0$ и $\mathcal{P} > 0$. Поэтому из полученного уравнения следует, что $\ddot{l}/l < 0$.

Рассмотрим случай, когда $l > 0$. Тогда из уравнения (35.29) вытекает, что функция l является вогнутой, т.е. $\ddot{l} < 0$. Это означает, что если к кривой $l(s)$ провести касательную, то вся кривая будет лежать ниже этой касательной. Зафиксируем некоторую точку $s = p$. Тогда в этой точке возможны три случая $\dot{l} < 0$, $\dot{l} = 0$ и $\dot{l} > 0$, которые качественно показаны на рис. 35.1. Если в точке p производная больше нуля,

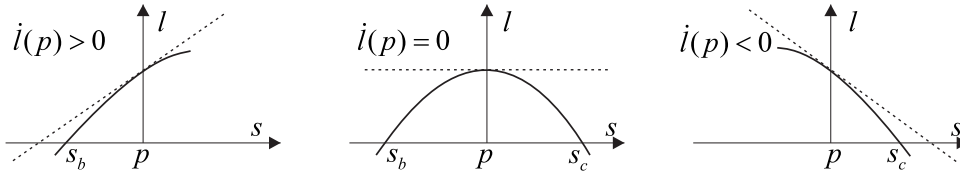


Рис. 35.1: Качественное поведение кривой $l(s)$ в трех случаях: $\dot{l}(p) > 0$, $\dot{l}(p) = 0$ и $\dot{l}(p) < 0$. В любом случае функция $l(s)$ обратится в нуль при конечном значении s .

$\dot{l}(p) > 0$, то функция $l(s)$ должна обратиться в нуль при конечном значении $s < p$, которое мы обозначим через s_b . Если $\dot{l}(p) < 0$, то функция $l(s)$ обратится в нуль при некотором значении $s_c > p$. При $\dot{l}(p) = 0$ функция $l(s)$ обязательно обратится в нуль при некоторых $s_b < p$ и $s_c > p$.

Пусть теперь $l < 0$. В этом случае функция l является выпуклой, т.к. $\ddot{l} > 0$. Как и предыдущем случае функция l обратится в нуль при конечных значениях параметра s . Нужно просто отразить графики, показанные на рис. 35.1, относительно оси абсцисс s .

При переходе через точку $l(s) = 0$ вторая производная $\ddot{l}$ меняет знак. Это значит, что $\ddot{l}(s_{b,c}) = 0$, и точки s_b и s_c являются точками перегиба. Допустим, что функция $l(s)$ при переходе через нуль ведет себя степенным образом. Поскольку она меняет знак, то главный член разложения будет нечетной степени:

$$l \approx C(s - s_{b,c})^{2k+1}, \quad C = \text{const} \neq 0, \quad k = 1, 2, \dots$$

Тогда

$$\lim_{s \rightarrow s_{b,c}} \frac{\ddot{l}}{l} = \lim_{s \rightarrow s_{b,c}} \frac{2k(2k+1)}{(s - s_{b,c})^2} = \infty.$$

Поэтому из уравнения (35.29) вытекает, что в точках s_b и s_c либо $\mathcal{E} + 3\mathcal{P} \rightarrow \infty$, либо $\sigma^2 \rightarrow \infty$. Это означает, что решения уравнений Эйнштейна будут иметь особенность либо в прошлом, либо в будущем. Поскольку параметр вдоль линий тока выбран каноническим, то эта особенность будет лежать при конечном значении собственного времени. Особенность, лежащая в прошлом, называется *большим взрывом*, а лежащая в будущем – *большим сжатием*.

Замечание. В англоязычной литературе употребляются термины *big bang* и *big crunch*. Дословный перевод второго термина – *большой хруст*. Его также иногда употребляют вместо “большого сжатия”. $\square$

Таким образом, при достаточно общих предположениях мы доказали наличие особенности в космологических решениях уравнений Эйнштейна либо в прошлом, либо в будущем. Подчеркнем, что при этом не были использованы какие-либо решения уравнений Эйнштейна. Большинство рассмотренных в настоящей главе космологических решений действительно имеет особенности. В то же время существуют отдельные решения без особенностей. В настоящее время принято считать, что окружающая нас вселенная родилась из большого взрыва, что неплохо согласуется с наблюдательными данными.

35.2 Трехмерные пространства постоянной кривизны

В настоящем разделе мы рассмотрим трехмерные римановы пространства положительной и отрицательной постоянной кривизны, которые являются однородными и изотропными многообразиями (см. раздел 15.3). Эти пространства важны для дальнейшего рассмотрения, поскольку играют роль пространственных сечений постоянного времени в космологических моделях Фридмана.

35.2.1 Трехмерная сфера $\mathbb{S}^3$

Рассмотрим четырехмерное евклидово пространство $\mathbb{R}^4$ с декартовой системой координат w, x, y, z . Это значит, что метрика в $\mathbb{R}^4$ имеет вид

$$ds^2 := dw^2 + dx^2 + dy^2 + dz^2. \quad (35.30)$$

Трехмерная сфера $\mathbb{S}^3$ радиуса a с центром в начале координат задается уравнением

$$w^2 + x^2 + y^2 + z^2 = a^2, \quad a > 0. \quad (35.31)$$

Это уравнение задает множество точек в $\mathbb{R}^4$, равноудаленных от начала координат. Мы предполагаем, что топология и дифференцируемая структура на $\mathbb{S}^3$ заданы вложением $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$. Тем самым множество точек $\mathbb{S}^3$ становится компактным трехмерным гладким многообразием. Кроме этого мы предполагаем, что на $\mathbb{S}^3$ задана метрика, также индуцированная вложением $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$. Тогда сфера $\mathbb{S}^3$ превращается в риманово многообразие, как мы увидим, постоянной положительной кривизны.

Чтобы определить индуцированную метрику, решим уравнение (35.31) относительно w :

$$w = \pm \sqrt{a^2 - x^2 - y^2 - z^2},$$

где знаки $\pm$ соответствуют верхней и нижней полусфере. Тогда

$$dw = \mp \frac{xdx + ydy + zdz}{\sqrt{a^2 - x^2 - y^2 - z^2}}.$$

Подстановка полученного выражения в евклидов интервал (35.30) приводит к следующему выражению

$$ds^2 = \frac{(a^2 - y^2 - z^2)dx^2 + (a^2 - x^2 - z^2)dy^2 + (a^2 - x^2 - y^2)dz^2}{a^2 - x^2 - y^2 - z^2} + 2 \frac{xydx dy + xzdx dz + yzdy dz}{a^2 - x^2 - y^2 - z^2}. \quad (35.32)$$

Это и есть метрика на сфере, индуцированная вложением $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$. Координаты x, y, z и на нижней, и на верхней полусфере определены внутри трехмерного шара $\mathbb{B}_a^3$ радиуса a :

$$x^2 + y^2 + z^2 < a^2.$$

Отсюда, в частности, следует, что сфера $\mathbb{S}^3$ трехмерна.

Индуцированная метрика на $\mathbb{S}^3$ принимает более простой диагональный вид, если ввести сферические координаты r, θ, φ в трехмерном евклидовом подпространстве $\mathbb{R}^3 \subset \mathbb{R}^4$ с координатами x, y, z :

$$\begin{aligned} x &:= r \sin \theta \cos \varphi, \\ y &:= r \sin \theta \sin \varphi, \\ z &:= r \cos \theta. \end{aligned} \quad (35.33)$$

Тогда уравнение сферы (35.31) примет вид

$$w^2 + r^2 = a^2 \quad \Leftrightarrow \quad w = \pm \sqrt{a^2 - r^2}.$$

Отсюда следует равенство дифференциалов

$$dw = \mp \frac{r dr}{\sqrt{a^2 - r^2}}.$$

Подстановка dw в интервал (35.30) приводит к диагональной индуцированной метрике

$$ds^2 = \frac{dr^2}{1 - r^2/a^2} + r^2(d\theta^2 + \sin^2 \theta d\varphi^2). \quad (35.34)$$

Полученная метрика определена при

$$0 < r < a, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi.$$

Эти значения координат покрывают либо верхнюю полусферу, $w > 0$, либо нижнюю, $w < 0$. По аналогии с двумерной сферой, мы считаем, что значение $r = 0$ соответствует либо северному, либо южному полюсу. Экватору сферы $\mathbb{S}^3$ соответствует максимальное значение $r = a$.

Теперь сделаем замену радиальной координаты $r \mapsto \chi$:

$$r := a \sin \chi \quad \Rightarrow \quad dr = a \cos \chi d\chi,$$

где χ – новая безразмерная радиальная координата. Тогда метрика примет вид

$$ds^2 = a^2 [d\chi^2 + \sin^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2)]. \quad (35.35)$$

Новая координата меняется в интервале $\chi \in (0, \pi/2)$ для каждой из полусфер. Однако метрика (35.35) определена в более широком интервале $\chi \in (0, \pi)$. Поэтому можно считать, что индуцированная метрика в виде (35.35) покрывает всю сферу, за исключением двумерного сечения $\varphi = 0$, которое включает северный и южный полюса. При этом две полусферы гладко сшиты по экватору $\chi = \pi/2$. Таким образом, безразмерные сферические координаты определены при

$$0 < \chi < \pi, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi.$$

Фактически, безразмерные сферические координаты χ, θ, φ – это сферические координаты на сфере $\mathbb{S}^3$ единичного радиуса, $a = 1$.

Метрика трехмерной сферы $\mathbb{S}^3$ в безразмерных сферических координатах диагональна. Следовательно, диагональна также обратная метрика. Поэтому координатные векторные поля $\partial_\chi, \partial_\theta$ и ∂_φ ортогональны.

Каждое двумерное сечение $r := a \sin \chi = \text{const}$ трехмерной сферы $\mathbb{S}^3$, которое соответствует постоянному радиусу, представляет двумерную сферу $\mathbb{S}_r^2 \subset \mathbb{S}^3$ с обычной индуцированной метрикой

$$dl^2 = a^2 \sin^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2), \quad \chi = \text{const}.$$

Площадь поверхности этой сферы максимальна на экваторе $\chi = \pi/2$ и равна $4\pi a^2$. На полюсах ($\chi = 0, 2\pi$) площадь поверхности двумерной сферы равна нулю. Для сравнения напомним, что сечения двумерной сферы $\mathbb{S}^2$ при обычном вложении $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ плоскостями $z = \text{const}$ представляют собой окружности $\mathbb{S}^1$, длина которых убывает от экватора к полюсам.

Объем трехмерной сферы просто вычисляется в безразмерных сферических координатах:

$$V = \int_0^\pi d\chi \int_0^\pi d\theta \int_0^{2\pi} d\varphi a^3 \sin^2 \chi \sin \theta = 2\pi^2 a^3. \quad (35.36)$$

Метрику (35.34), индуцированную на сфере, можно записать в конформно плоском виде. Действительно, совершив преобразование радиальной координаты $r \mapsto \rho(r)$ такое, чтобы было выполнено уравнение

$$\frac{1}{1 - r^2/a^2} \left(\frac{dr}{d\rho} \right)^2 = \frac{r^2}{\rho^2}.$$

Это уравнение легко решается. С точностью до несущественной постоянной интегрирования решение имеет вид

$$r = \frac{\rho}{1 + \frac{\rho^2}{4a^2}}. \quad (35.37)$$

После такого преобразования радиальной координаты метрика становится конформно плоской:

$$ds^2 = \frac{d\rho^2 + \rho^2 (d\theta^2 + \sin^2 \theta d\varphi^2)}{\left(1 + \frac{\rho^2}{4a^2} \right)^2}. \quad (35.38)$$

Эта метрика определена при $0 < \rho < \infty$. Северному и южному полюсам ставятся в соответствие точки $\rho = \infty$ и $\rho = 0$. Экватору сферы $\mathbb{S}^3$ соответствует значение $\rho = 2a$. Заметим, что на экваторе метрика сферы невырождена. По аналогии с двумерной сферой будем называть систему координат ρ, θ, φ *стереографической*.

Рассмотрим координаты ρ, θ, φ в качестве сферических координат в $\mathbb{R}^3$ и перейдем к декартовым координатам:

$$\begin{aligned} x^1 &:= \rho \sin \theta \cos \varphi, \\ x^2 &:= \rho \sin \theta \sin \varphi, \\ x^3 &:= \rho \cos \theta. \end{aligned} \quad (35.39)$$

Введем обозначение $\{x^\mu\} = \{x^1, x^2, x^3\}$, $\mu = 1, 2, 3$. Тогда метрика примет вид

$$ds^2 = \frac{dx^2}{\left(1 + \frac{x^2}{4a^2}\right)^2}, \quad (35.40)$$

где

$$dx^2 := \delta_{\mu\nu} dx^\mu dx^\nu, \quad x^2 := \rho^2 = \delta_{\mu\nu} x^\mu x^\nu.$$

Индукцированная метрика на сфере $\mathbb{S}^3$ с стереографических координатах (35.40) определена во всем евклидовом пространстве $\{x^\mu\} \in \mathbb{R}^3$ и является гладкой. Эти координаты покрывают всю сферу за исключением северного полюса, которому соответствует бесконечно удаленная точка $\rho = \infty$. Северная и южная полусферы гладко сшиваются по экватору, которому соответствует двумерная сфера $\rho = 2a$.

Теперь вычислим геометрические характеристики трехмерной сферы $\mathbb{S}^3$ для индуцированной метрики (35.40) в стереографических координатах. Начнем с символов Кристоффеля. Несложные вычисления дают следующее выражение

$$\Gamma_{\mu\nu}^\rho = -\frac{1}{2a^2 \left(1 + \frac{x^2}{4a^2}\right)} (x_\mu \delta_\nu^\rho + x_\nu \delta_\mu^\rho - x^\rho \delta_{\mu\nu}), \quad (35.41)$$

где $x_\mu := \delta_{\mu\nu} x^\nu$. Ему соответствует тензор кривизны

$$R_{\mu\nu\rho\sigma} = -\frac{\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\nu\rho}\delta_{\mu\sigma}}{a^2 \left(1 + \frac{x^2}{4a^2}\right)^4} = -\frac{g_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}g_{\mu\sigma}}{a^2}. \quad (35.42)$$

Отсюда следует, что трехмерная сфера $\mathbb{S}^3$ с индуцированной метрикой является римановым пространством постоянной кривизны. Тензор Риччи и скалярная кривизна имеют вид

$$\begin{aligned} R_{\mu\rho} &= -\frac{2\delta_{\mu\rho}}{a^2 \left(1 + \frac{x^2}{4a^2}\right)^2} = -\frac{2g_{\mu\rho}}{a^2}, \\ R &= -\frac{6}{a^2}. \end{aligned} \quad (35.43)$$

Гауссова кривизна при этом положительна:

$$K := -\frac{1}{6}R = \frac{1}{a^2}.$$

Поэтому трехмерная сфера $\mathbb{S}^3$ с индуцированной метрикой является пространством постоянной положительной кривизны.

Евклидова метрика (35.30) и уравнение (35.31), определяющее вложение $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$, инвариантны относительно группы вращений $\mathbb{O}(4)$, которая действует в исходном евклидовом пространстве $\mathbb{R}^4$ обычным образом. Группа Ли вращений $\mathbb{O}(4)$ шестипараметрическая, и поэтому в $\mathbb{R}^4$ существует шесть линейно независимых векторов Киллинга:

$$\begin{aligned} K_{01} &:= w\partial_x - x\partial_w, & K_{12} &:= x\partial_y - y\partial_x, \\ K_{02} &:= w\partial_y - y\partial_w, & K_{13} &:= x\partial_z - z\partial_x, \\ K_{03} &:= w\partial_z - z\partial_w, & K_{23} &:= y\partial_z - z\partial_y. \end{aligned} \quad (35.44)$$

Если пронумеровать координаты, $\{x^i\} := \{w, x, y, z\}$, $i = 0, 1, 2, 3$, то векторы Киллинга можно записать в компактной форме

$$K_{ij} := x_i\partial_j - x_j\partial_i,$$

где $x_i := x^j\delta_{ji}$. Векторы Киллинга удовлетворяют алгебре Ли $\mathfrak{so}(4)$:

$$[K_{ij}, K_{kl}] = -\delta_{ik}K_{jl} + \delta_{il}K_{jk} + \delta_{jk}K_{il} - \delta_{jl}K_{ik}, \quad (35.45)$$

в чем нетрудно убедиться с помощью прямых вычислений.

Поскольку алгебра Ли векторных полей Киллинга замкнута, то они находятся в инволюции, и, согласно теореме Фробениуса, у распределения векторных полей Киллинга существуют интегральные подмногообразия, которыми, как мы увидим, являются сферы $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$.

Для того, чтобы получить явные выражения для векторных полей Киллинга на сфере $\mathbb{S}^3$ сделаем следующее. Переменные r, χ, θ, φ можно рассматривать, как локальные координаты в $\mathbb{R}^4$:

$$\begin{aligned} w &:= r \cos \chi, & x &:= r \sin \chi \sin \theta \cos \varphi, \\ z &:= r \sin \chi \cos \theta, & y &:= r \sin \chi \sin \theta \sin \varphi. \end{aligned}$$

Отсюда следует, что значения $\chi = 0, \pi$ соответствуют северному и южному полюсу, а $\chi = \pi/2$ – экватору.

В дальнейшем нам понадобятся следующие формулы:

$$\begin{aligned} x^2 + y^2 &= r^2 \sin^2 \chi \sin^2 \theta, \\ x^2 + y^2 + z^2 &= r^2 \sin^2 \chi, \\ w^2 + x^2 + y^2 + z^2 &= r^2. \end{aligned}$$

Обратные преобразования координат имеют вид:

$$\begin{aligned} r &= \sqrt{w^2 + x^2 + y^2 + z^2}, \\ \chi &= \arctg \frac{\sqrt{x^2 + y^2 + z^2}}{w}, \\ \theta &= \arctg \frac{\sqrt{x^2 + y^2}}{z}, \\ \varphi &= \arctg \frac{y}{x}. \end{aligned}$$

Для нахождения явного вида векторов Киллинга будут использованы частные

производные, определяющие дифференциал отображения $(w, x, y, z) \mapsto (r, \chi, \theta, \varphi)$:

$$\begin{aligned}
 \frac{\partial r}{\partial w} &= \cos \chi, & \frac{\partial \theta}{\partial w} &= 0, \\
 \frac{\partial r}{\partial x} &= \sin \chi \sin \theta \cos \varphi, & \frac{\partial \theta}{\partial x} &= \frac{\cos \theta \cos \varphi}{r \sin \chi}, \\
 \frac{\partial r}{\partial y} &= \sin \chi \sin \theta \sin \varphi, & \frac{\partial \theta}{\partial y} &= \frac{\cos \theta \sin \varphi}{r \sin \chi}, \\
 \frac{\partial r}{\partial z} &= \sin \chi \cos \theta, & \frac{\partial \theta}{\partial z} &= -\frac{\sin \theta}{r \sin \chi}, \\
 \frac{\partial \chi}{\partial w} &= -\frac{\sin \chi}{r}, & \frac{\partial \varphi}{\partial w} &= 0, \\
 \frac{\partial \chi}{\partial x} &= \frac{\cos \chi \sin \theta \cos \varphi}{r}, & \frac{\partial \varphi}{\partial x} &= -\frac{\sin \varphi}{r \sin \chi \sin \theta}, \\
 \frac{\partial \chi}{\partial y} &= \frac{\cos \chi \sin \theta \sin \varphi}{r}, & \frac{\partial \varphi}{\partial y} &= \frac{\cos \varphi}{r \sin \chi \sin \theta}, \\
 \frac{\partial \chi}{\partial z} &= \frac{\cos \chi \cos \theta}{r}, & \frac{\partial \varphi}{\partial z} &= 0.
 \end{aligned} \tag{35.46}$$

Тогда несложные вычисления приводят к следующим выражениям для векторов Киллинга в новых координатах:

$$\begin{aligned}
 K_{01} &= \sin \theta \cos \varphi \partial_\chi + \operatorname{ctg} \chi \cos \theta \cos \varphi \partial_\theta - \frac{\operatorname{ctg} \chi \sin \varphi}{\sin \theta} \partial_\varphi, \\
 K_{02} &= \sin \theta \sin \varphi \partial_\chi + \operatorname{ctg} \chi \cos \theta \sin \varphi \partial_\theta + \frac{\operatorname{ctg} \chi \cos \varphi}{\sin \theta} \partial_\varphi, \\
 K_{03} &= \cos \theta \partial_\chi - \operatorname{ctg} \chi \sin \theta \partial_\theta, \\
 K_{12} &= \partial_\varphi, \\
 K_{13} &= -\cos \varphi \partial_\theta + \operatorname{ctg} \theta \sin \varphi \partial_\varphi, \\
 K_{23} &= -\sin \varphi \partial_\theta - \operatorname{ctg} \theta \cos \varphi \partial_\varphi.
 \end{aligned} \tag{35.47}$$

Мы видим, что все шесть векторов Киллинга направлены вдоль касательных векторов ∂_χ , ∂_θ и ∂_φ . Следовательно, гиперповерхности, выделяемые равенством $r = \text{const}$, являются интегральными гиперповерхностями распределения векторных полей Киллинга. Это – в точности трехмерные сферы $\mathbb{S}^3$, и векторные поля Киллинга касательны к ним. Векторных поля Киллинга линейно независимы. Поскольку сфера трехмерна, то между векторными полями Киллинга можно написать линейную зависимость, коэффициентами которой будут не числа, а функции. Ситуация здесь аналогична вложению двумерной сферы $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$, которое было рассмотрено в разделе 29.1.

Таким образом, векторные поля (35.47) являются векторными полями Киллинга на сфере $\mathbb{S}^3$, вложенной в четырехмерное евклидово пространство $\mathbb{R}^4$. Они удовлетворяют алгебре Ли (35.45), т.к. коммутатор векторных полей не зависит от выбора системы координат. При желании векторы Киллинга (35.47) можно записать в любой другой системе координат на сфере $\mathbb{S}^3$.

Согласно теореме 15.3.2 шесть – это максимально возможное число линейно независимых векторов Киллинга на трехмерном многообразии, и сфера $\mathbb{S}^3$ является пространством постоянной кривизны. Последнее утверждение уже было проверено с помощью прямых вычислений (35.42).

Рассмотрим южный полюс сферы $x = y = z = 0$, $w = -a$. Подгруппа группы изометрий сферы, оставляющей южный полюс неподвижным, является группа Ли

трехмерных вращений $\mathbb{SO}(3)$. Это – группа изотропии южного полюса. Ее генераторами являются векторные поля Киллинга K_{12} , K_{13} и K_{23} , которые мы рассмотрим в декартовой системе координат (35.47), т.к. она невырождена в южном полюсе. Они обращаются в нуль в южном полюсе, и поэтому эта точка является неподвижной относительно действия подгруппы $\mathbb{SO}(3) \subset \mathbb{SO}(4)$. В южном полюсе остальные векторные поля Киллинга принимают вид

$$K_{01}^\circ = -a\partial_x, \quad K_{02}^\circ = -a\partial_y, \quad K_{03}^\circ = -a\partial_z,$$

т.е. имеют тот же вид, что и трансляции в трехмерном евклидовом пространстве. Мы видим, что эти векторные поля играют роль трансляций для южного полюса. Поэтому многообразие $\mathbb{S}^3$ однородно и изотропно в окрестности южного полюса. Из явного вида коммутатора векторных полей Киллинга (35.45) следует, что векторы Киллинга K_{01} , K_{02} и K_{03} не образуют подалгебры в алгебре Ли $\mathfrak{so}(4)$ и, следовательно, не генерируют никакой подгруппы в группе Ли $\mathbb{SO}(4)$. Это отличает “трансляции” в группе $\mathbb{SO}(4)$ от трансляций в группе Пуанкаре, где они образуют подгруппу. Одно и то же векторное поле Киллинга может генерировать для одной точки сферы вращение, а для другой – трансляцию.

Поскольку группа Ли изометрий $\mathbb{SO}(4)$ действует на сфере $\mathbb{S}^3$ транзитивно, то все, сказанное выше о южном полюсе, справедливо для любой точки сферы. Поэтому сфера $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$ является однородным и изотропным пространством в смысле определения, данного в разделе 15.3.

В дальнейшем нам понадобятся явные выражения для векторных полей Киллинга в стереографических координатах. Чтобы упростить их получение, введем обозначения для начальных декартовых координат в $\mathbb{R}^4$:

$$(w, x, y, z) \mapsto (w, y^\mu), \quad \mu = 1, 2, 3.$$

Координаты y^μ связаны со стереографическими координатами (35.39) простой формулой

$$y^\mu = \frac{x^\mu}{1 + \frac{\rho^2}{4a^2}}, \quad (35.48)$$

где мы использовали равенство (35.37). Отсюда следует выражение для частных производных

$$\frac{\partial y^\mu}{\partial x^\nu} = \frac{\delta_\nu^\mu}{1 + \frac{\rho^2}{4a^2}} - \frac{x^\mu x_\nu}{2a^2 \left(1 + \frac{\rho^2}{4a^2}\right)^2},$$

где $x_\nu := x^\mu \delta_{\mu\nu}$. Обратная матрица дает матрицу Якоби преобразования координат $y^\mu \mapsto x^\mu$:

$$J_\mu^\nu := \frac{\partial x^\nu}{\partial y^\mu} = \left(1 + \frac{\rho^2}{4a^2}\right) \delta_\mu^\nu + \frac{1}{2a^2} \frac{1 + \frac{\rho^2}{4a^2}}{1 - \frac{\rho^2}{4a^2}} x^\nu x_\mu. \quad (35.49)$$

Уравнение сферы в координатах w, x^μ принимает вид

$$w^2 + x^2 + y^2 + z^2 = w^2 + \frac{\rho^2}{\left(1 + \frac{\rho^2}{4a^2}\right)^2} = a^2.$$

Отсюда следует выражение для w через стереографические координаты:

$$w = \pm a \frac{1 - \frac{\rho^2}{4a^2}}{1 + \frac{\rho^2}{4a^2}}. \quad (35.50)$$

Используя полученное выражение для w и матрицу Якоби (35.49), легко вычислить векторы Киллинга:

$$\begin{aligned} K_{0\mu} &= w \frac{\partial x^\rho}{\partial y^\mu} \frac{\partial}{\partial x^\rho} = \pm a \left[\left(1 - \frac{\rho^2}{4a^2} \right) \frac{\partial}{\partial x^\mu} + \frac{1}{2a^2} x_\mu x^\rho \frac{\partial}{\partial x^\rho} \right], \\ K_{\mu\nu} &= y_\mu \frac{\partial x^\rho}{\partial y^\nu} \frac{\partial}{\partial x^\rho} - y_\nu \frac{\partial x^\rho}{\partial y^\mu} \frac{\partial}{\partial x^\rho} = x_\mu \frac{\partial}{\partial x^\nu} - x_\nu \frac{\partial}{\partial x^\mu}. \end{aligned} \quad (35.51)$$

Мы видим важное отличие векторов Киллинга в различных системах координат. В сферической системе координат векторы Киллинга (35.47) не зависят от радиуса a сферы. Это кажется естественным, т.к. векторы Киллинга определяют симметрию многообразия, а она не зависит от радиуса сферы. В то же время три вектора Киллинга $K_{0\mu}$ (35.51), соответствующие трансляциям в южном полюсе, в стереографических координатах явно зависят от радиуса сферы. Векторы Киллинга $K_{\mu\nu}$ (35.51) соответствуют вращениям в южном полюсе и от радиуса сферы не зависят.

Трехмерная сфера $\mathbb{S}^3$ естественным образом вкладывается в четырехмерное евклидово пространство $\mathbb{R}^4$, как было описано в настоящем разделе. Группа вращений $\mathbb{SO}(4)$ действует на сфере транзитивно, т.е. для любых двух точек найдется такое вращение, которое переводит одну точку в другую (см. главу 9). Следовательно, она является однородным пространством согласно определению, данному в конце раздела 9.1. Подгруппа изотропии каждой точки сферы изоморфна группе трехмерных вращений $\mathbb{SO}(3)$. Поэтому согласно теореме 9.1.2 между точками сферы и точками факторпространства $\mathbb{SO}(4)/\mathbb{SO}(3)$ существует взаимно однозначное соответствие. Более того, из теоремы 9.1.1 следует, что на факторпространстве существует единственная структура вещественно аналитического многообразия. Поэтому сфера, как многообразие, диффеоморфна факторпространству

$$\mathbb{S}^3 \approx \frac{\mathbb{SO}(4)}{\mathbb{SO}(3)}. \quad (35.52)$$

С другой стороны, на трехмерной сфере была определена групповая структура (1.142)

$$\mathbb{S}^3 \approx \mathbb{SU}(2). \quad (35.53)$$

Группа унитарных матриц $\mathbb{SU}(2)$ действует на трехмерной сфере транзитивно и свободно, т.е. любой элемент группы, отличный от единичного, сдвигает все точки многообразия. Поэтому сфера $\mathbb{S}^3$ является главным однородным пространством, которое можно просто отождествить с $\mathbb{SU}(2)$.

35.2.2 Двуполостный гиперболоид $\mathbb{H}^3$

Рассмотрим четырехмерное пространство Минковского $\mathbb{R}^{3,1}$ с декартовой системой координат x, y, z, w и метрикой

$$ds^2 := dx^2 + dy^2 + dz^2 - dw^2. \quad (35.54)$$

Двуполостный гиперболоид $\mathbb{H}^3$ “радиуса” a с центром в начале координат задается уравнением

$$x^2 + y^2 + z^2 - w^2 = -a^2, \quad a > 0. \quad (35.55)$$

При работе с трехмерным гиперболоидом $\mathbb{H}^3$ полезно иметь перед глазами картинку 29.2, где показано вложение двумерного двуполостного гиперболоида $\mathbb{H}^2$ в трехмерное пространство Минковского $\mathbb{R}^{1,2}$.

Будем считать, что топология, дифференцируемая структура и метрика на $\mathbb{H}^3$ индуцированы вложением $\mathbb{H}^3 \hookrightarrow \mathbb{R}^{3,1}$.

Из уравнения (35.55), определяющего вложение, можно исключить координату w :

$$w = \pm \sqrt{a^2 + x^2 + y^2 + z^2},$$

где знак $\pm$ соответствует северной и южной полам гиперboloида. Отсюда следует, что двуполостный гиперboloид состоит из двух компонент связности: северной и южной. Для дифференциалов справедлива формула

$$dw = \pm \frac{xdx + ydy + zdz}{\sqrt{a^2 + x^2 + y^2 + z^2}}.$$

Подстановка полученного выражения в евклидов интервал (35.54) приводит к индуцированной метрике

$$ds^2 = \frac{(a^2 + y^2 + z^2)dx^2 + (a^2 + x^2 + z^2)dy^2 + (a^2 + x^2 + y^2)dz^2}{a^2 + x^2 + y^2 + z^2} - 2 \frac{xydx dy + xzdx dz + yzdy dz}{a^2 + x^2 + y^2 + z^2}. \quad (35.56)$$

Координаты x, y, z можно выбрать в качестве координат на двуполостном гиперboloиде $\mathbb{H}^3$. На каждой полё гиперboloида они пробегают все евклидово пространство $x, y, z \in \mathbb{R}^3$.

Для определенности выберем северную полё гиперboloида. Она диффеоморфна $\mathbb{R}^3$, некомпактна, и система координат x, y, z является глобальной, т.е. покрывает всю полё гиперboloида. Для нее сохраним прежнее обозначение $\mathbb{H}^3$.

Индуцированная метрика на гиперboloиде $\mathbb{H}^3$ принимает более простой диагональный вид, если ввести сферические координаты r, θ, φ в подпространстве $\mathbb{R}^3 \subset \mathbb{R}^{3,1}$ с координатами x, y, z (35.33). Тогда уравнение гиперboloида примет вид

$$r^2 - w^2 = -a^2 \quad \Leftrightarrow \quad w = \pm \sqrt{a^2 + r^2}.$$

Отсюда следует равенство дифференциалов

$$dw = \pm \frac{rdr}{\sqrt{a^2 + r^2}}.$$

Подстановка dw в исходный интервал (35.54) приводит к диагональной индуцированной метрике

$$ds^2 = \frac{dr^2}{1 + r^2/a^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2). \quad (35.57)$$

Сферические координаты определены на всей верхней полё гиперboloида:

$$0 < r < \infty, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi.$$

Отсюда, в частности, следует, что индуцированная метрика на двуполостном гиперboloиде положительно определена. Это утверждение не является очевидным, т.к. исходным было пространство Минковского $\mathbb{R}^{3,1}$.

Теперь сделаем замену радиальной координаты $r \mapsto \chi$:

$$r := a \operatorname{sh} \chi \quad \Rightarrow \quad dr = a \operatorname{ch} \chi d\chi,$$

где $\chi \in (0, \infty)$ – новая безразмерная радиальная координата. Тогда метрика примет вид

$$ds^2 = a^2 [d\chi^2 + \operatorname{sh}^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2)]. \quad (35.58)$$

Каждое двумерное сечение $r = a \operatorname{sh} \chi = \text{const}$ трехмерного гиперboloида $\mathbb{H}^3$ представляет собой двумерную сферу $\mathbb{S}_r^2 \subset \mathbb{H}^3$ с обычной индуцированной метрикой. Площадь этой сферы равна нулю в вершине $\chi = 0$ и бесконечно возрастает при $\chi \rightarrow \infty$.

Легко проверить, что объем каждой полу двуполостного гиперboloида бесконечен.

Теперь запишем метрику двуполостного гиперboloида $\mathbb{H}^3$ в стереографических координатах, которые часто встречаются в приложениях. Как и в случае трехмерной сферы $\mathbb{S}^3$ совершим преобразование радиальной координаты $r \mapsto \rho(r)$ такое, чтобы было выполнено уравнение

$$\frac{1}{1 + r^2/a^2} \left(\frac{dr}{d\rho} \right)^2 = \frac{r^2}{\rho^2}.$$

С точностью до несущественной постоянной интегрирования решение имеет вид

$$r = \frac{\rho}{1 - \frac{\rho^2}{4a^2}}.$$

После этого преобразования радиальной координаты метрика становится конформно плоской:

$$ds^2 = \frac{d\rho^2 + \rho^2(d\theta^2 + \sin^2 \theta d\varphi^2)}{\left(1 - \frac{\rho^2}{4a^2}\right)^2}. \quad (35.59)$$

Метрика двуполостного гиперboloида $\mathbb{H}^3$ определена внутри шара $\rho < \frac{1}{2a}$. Значение $\rho = 0$ соответствует вершине гиперboloида $r = 0$, а $\rho = 2a$ – бесконечно удаленной точке $r = \infty$.

В евклидовом пространстве $(\rho, \theta, \varphi) \in \mathbb{R}^3$ можно ввести декартовы координаты (35.39). Тогда метрика примет вид

$$ds^2 = \frac{dx^2}{\left(1 - \frac{x^2}{4a^2}\right)^2}. \quad (35.60)$$

Она отличается от метрики трехмерной сферы (35.40) только формальной заменой $a^2 \mapsto -a^2$.

Геометрические характеристики двуполостного гиперboloида получаются из формул (35.42)–(35.43) заменой $a^2 \mapsto -a^2$:

$$\begin{aligned} R_{\mu\nu\rho\sigma} &= \frac{g_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}g_{\mu\sigma}}{a^2}, \\ R_{\mu\rho} &= \frac{2g_{\mu\rho}}{a^2}, \\ R &= \frac{6}{a^2}, \quad K = -\frac{1}{a^2}. \end{aligned} \quad (35.61)$$

Отсюда следует, что двуполостный гиперboloид $\mathbb{H}^3$ с индуцированной метрикой является римановым пространством постоянной отрицательной кривизны.

Лоренцева метрика (35.54) и уравнение (35.55), определяющее вложение, инвариантны относительно группы Лоренца $\mathbb{O}(3, 1)$, которая действует в пространстве Минковского $\mathbb{R}^{3,1}$ обычным образом. Группа Ли $\mathbb{O}(3, 1)$ шестипараметрическая, и поэтому в $\mathbb{R}^{3,1}$ существует шесть векторов Киллинга:

$$K_{ij} := x_i \partial_j - x_j \partial_i, \quad i, j = 1, 2, 3, 4, \quad (35.62)$$

где $\{x^i\} := \{x, y, z, w\}$, и $x_i := x^j \eta_{ji}$, $\eta_{ij} := \text{diag}(+ + + -)$. Нетрудно проверить, что векторные поля Киллинга удовлетворяют алгебре Ли $\mathfrak{so}(3.1)$:

$$[K_{ij}, K_{kl}] = -\eta_{ik} K_{jl} + \eta_{il} K_{jk} + \eta_{jk} K_{il} - \eta_{jl} K_{ik}. \quad (35.63)$$

Чтобы получить явные выражения для векторов Киллинга на $\mathbb{H}^3$ поступим так же, как и в случае сферы $\mathbb{S}^3$. А именно, будем рассматривать переменные r, χ, θ, φ в качестве сферических координат в пространстве Минковского $\mathbb{R}^{3,1}$:

$$\begin{aligned} x &:= r \operatorname{sh} \chi \sin \theta \cos \varphi, & z &:= r \operatorname{sh} \chi \cos \theta, \\ y &:= r \operatorname{sh} \chi \sin \theta \sin \varphi, & w &:= r \operatorname{ch} \chi. \end{aligned}$$

При этом вершине гиперboloида соответствует значение $\chi = 0$.

Нетрудно проверить следующие формулы:

$$\begin{aligned} x^2 + y^2 &= r^2 \operatorname{sh}^2 \chi \sin^2 \theta, \\ x^2 + y^2 + z^2 &= r^2 \operatorname{sh}^2 \chi, \\ -x^2 - y^2 - z^2 + w^2 &= r^2. \end{aligned}$$

Обратные преобразования координат имеют вид

$$\begin{aligned} r &= \sqrt{w^2 - x^2 - y^2 - z^2} \\ \chi &= \operatorname{arcth} \frac{\sqrt{x^2 + y^2 + z^2}}{w}, \\ \theta &= \operatorname{arctg} \frac{\sqrt{x^2 + y^2}}{z}, \\ \varphi &= \operatorname{arctg} \frac{y}{x}. \end{aligned}$$

Частные производные координат r, χ, θ, φ по координатам x, y, z, w легко вычисляются и имеют вид (35.46) с заменой

$$\sin \chi \mapsto \operatorname{sh} \chi, \quad \cos \chi \mapsto \operatorname{ch} \chi.$$

Поскольку мы находимся в пространстве Минковского $\mathbb{R}^{3,1}$, то декартовы координаты с опущенным индексом имеют вид $\{x_i\} = \{x, y, z, -w\}$. Теперь нетрудно найти явный вид векторных полей Киллинга в сферических координатах:

$$\begin{aligned} K_{12} &= \partial_\varphi, \\ K_{13} &= -\cos \varphi \partial_\theta + \operatorname{ctg} \theta \sin \varphi \partial_\varphi, \\ K_{14} &= \sin \theta \cos \varphi + \operatorname{cth} \chi \cos \theta \cos \varphi \partial_\theta - \frac{\operatorname{cth} \chi \sin \varphi}{\sin \theta} \partial_\varphi, \\ K_{23} &= -\sin \varphi \partial_\theta - \operatorname{cth} \theta \cos \varphi \partial_\varphi, \\ K_{24} &= \sin \theta \sin \varphi \partial_\chi + \operatorname{cth} \chi \cos \theta \sin \varphi \partial_\theta + \frac{\operatorname{cth} \chi \cos \varphi}{\sin \theta} \partial_\varphi, \\ K_{34} &= \cos \theta \partial_\chi - \operatorname{cth} \chi \sin \theta \partial_\theta. \end{aligned} \quad (35.64)$$

Как видим, векторы Киллинга не зависят от радиуса r и не имеют вдоль него компонент.

Векторные поля Киллинга можно также записать в стереографических координатах аналогично тому, как это было сделано для сферы в предыдущем разделе.

Как и для сферы $\mathbb{S}^3$ распределение векторных полей Киллинга находится в инволюции и, согласно теореме Фробениуса, определяет интегральные подмногообразия в пространстве Минковского $\mathbb{R}^{3,1}$. Этими интегральными подмногообразиями являются двуполостные гиперboloиды $\mathbb{H}^3$, определяемые равенством $r = a$. Таким образом, на двуполостном гиперboloиде определено максимальное число (шесть) линейно независимых векторных полей Киллинга (35.64). Следовательно, согласно теореме 15.3.2, двуполостный гиперboloид $\mathbb{H}^3$ является пространством постоянной кривизны, что уже было проверено нами с помощью прямых вычислений (35.61).

Рассмотрим вершину северной полу гиперboloида $x = y = z = 0$, $w = a$. Группа изометрий этой точки является группа Ли трехмерных вращений $\mathbb{SO}(3)$. Ее генераторами являются векторные поля Киллинга K_{12} , K_{13} и K_{23} , которые мы рассмотрим в декартовой системе координат (35.62). Они обращаются в нуль в вершине, и поэтому эта точка является неподвижной относительно действия подгруппы $\mathbb{SO}(3) \subset \mathbb{SO}(3, 1)$. В вершине северной полу гиперboloида остальные векторные поля Киллинга принимают вид

$$K_{14}^\circ = a\partial_x, \quad K_{24}^\circ = a\partial_y, \quad K_{34}^\circ = a\partial_z,$$

т.е. имеют тот же вид, что и трансляции в трехмерном евклидовом пространстве. Поэтому эти векторные поля играют роль трансляций для вершины северной полу гиперboloида. Следовательно, многообразие $\mathbb{H}^3$ однородно и изотропно в окрестности вершины северной полу гиперboloида. Из явного вида коммутатора векторных полей Киллинга (35.63) вытекает, что векторы Киллинга K_{14} , K_{24} и K_{34} не образуют подалгебры в алгебре Ли $\mathfrak{so}(3, 1)$ и, следовательно, не генерируют никакой подгруппы в группе Ли $\mathbb{SO}(3, 1)$. Это отличает “трансляции” в группе $\mathbb{SO}(3, 1)$ от трансляций в группе Пуанкаре, где они образуют подгруппу.

Поскольку группа Ли изометрий $\mathbb{O}(3, 1)$ действует на двуполостном гиперboloиде $\mathbb{H}^3$ транзитивно, то все, сказанное выше о вершине северной полу гиперboloида, справедливо для любой точки $\mathbb{H}^3$. Поэтому двуполостный гиперboloид $\mathbb{H}^3 \hookrightarrow \mathbb{R}^{3,1}$ является однородным и изотропным пространством в смысле определения, данного в разделе 15.3.

Группа Лоренца $\mathbb{SO}(3, 1)$ действует на верхней полу гиперboloида $\mathbb{H}^3$ транзитивно. Поэтому верхняя полу гиперboloида является однородным пространством. Группа изотропии произвольной точки из $\mathbb{H}^3$ изоморфна группе трехмерных вращений $\mathbb{SO}(3)$. Поэтому, как многообразие, верхняя полу гиперboloида $\mathbb{H}^3$ изоморфна фактор пространству

$$\mathbb{H}^3 \approx \frac{\mathbb{SO}(3, 1)}{\mathbb{SO}(3)}. \quad (35.65)$$

На верхней полу гиперboloида нельзя ввести групповую структуру. Поэтому многообразиие $\mathbb{H}^3$ не является главным однородным пространством в отличие от сферы $\mathbb{S}^3$.

35.3 Вселенная Фридмана

Во вселенной очень много объектов, типичными представителями которых являются звезды, вокруг которых вращаются планеты. Звезды объединены в галактики,

которые, в свою очередь, образуют скопления галактик. Кроме этого во вселенной присутствует электромагнитное излучение, метеориты, кометы и множество других объектов, о которых мы в настоящее время, вполне возможно, и не догадываемся. К настоящему времени человечество накопило довольно много данных, полученных в результате наблюдений за звездным небом.

Согласно современным наблюдательным данным трехмерное пространство, которое нас окружает, в каждый момент времени $t := x^0$ в крупном масштабе является однородным и изотропным. Правильнее сказать, что современные данные не противоречат предположению об однородности и изотропии вселенной, которые являются естественными с физической точки зрения. Физически однородность пространства означает, что свойства вселенной не зависят от того, какая точка выбрана в качестве начала системы координат, а изотропия – что свойства вселенной не зависят от направления, выбранного в каждой точке пространства. Однородность и изотропию вселенной часто называют *космологическим принципом*. Слова “в крупном масштабе” следует понимать по аналогии с однородностью газа: нельзя говорить об однородности в малых областях вселенной; она однородна только после усреднения по достаточно большим областям пространства, включающим много скоплений галактик.

Об изотропии вселенной говорит изотропия реликтового микроволнового излучения, которое было открыто А. Пензиасом и Р. Вилсоном в 1965 году [277]. За это открытие они получили Нобелевскую премию. Впоследствии реликтовое излучение было измерено в широком диапазоне частот. С высокой степенью точности оно является изотропным и неполяризованным, а его спектр соответствует излучению абсолютно черного тела с температурой около $2,7^\circ K$.

Для построения модели вселенной в рамках общей теории относительности слов однородная и изотропная вселенная недостаточно. Определение и математический смысл этим словам был дан в разделе 15.3. Там было доказано, что однородное и изотропное многообразие является пространством постоянной кривизны. Верно также обратное утверждение: если (псевдо)риманово многообразие является пространством постоянной кривизны в смысле определения (15.25) при $R = \text{const}$, то оно является однородным и изотропным. Для построения космологической модели мы требуем, чтобы все сечения пространства-времени, соответствующие постоянному времени, и которые мы предполагаем пространственноподобными, были однородными и изотропными, т.е. пространствами постоянной кривизны, которые имеют максимально возможное число (шесть) векторов Киллинга. В разделе 35.3.2 мы покажем, что это требование является необходимым, но не достаточным. При этом в каждой точке пространства три векторных поля Киллинга генерируют группу вращений $\mathbb{O}(3)$ (изотропия вселенной) и три векторных поля Киллинга – симметрию относительно “трансляций” (однородность вселенной), как было показано в разделах 35.2.1 и 35.2.2.

35.3.1 Метрика однородной и изотропной вселенной

При рассмотрении моделей вселенных Фридмана мы предполагаем, что пространство-время с топологической точки зрения представляет собой топологическое произведение $\mathbb{M} = \mathbb{R} \times \mathbb{S}$, где $t \in \mathbb{R}$ – временная ось и $x = \{x^\mu\} \in \mathbb{S}$, $\mu = 1, 2, 3$, – трехмерное пространство постоянной кривизны с отрицательно определенной метрикой. Это требование эквивалентно условию, что все пространственные сечения $\mathbb{M}$, соответствующие постоянному времени $t = \text{const}$, являются пространствами постоянной кривизны $\mathbb{S}$. Мы также предполагаем, что пространство-время является геодезически

полным, т.е. любую геодезическую (экстремаль) можно продолжить до бесконечного значения канонического параметра. Это требование необходимо для того, чтобы, например, рассматривать в качестве пространства постоянной кривизны всю сферу целиком, а не какую-либо ее часть.

В дальнейшем мы увидим, что космологические решения уравнений Эйнштейна часто имеют сингулярности, и решение определено не на всей оси времени, а на полуоси или на конечном интервале. Для таких решений пространство-время не является геодезически полным. В настоящем разделе мы рассматриваем только кинематические свойства метрики, т.е. без учета уравнений движения, и, для простоты, будем считать, что метрика определена на всей вещественной оси $t \in \mathbb{R}$.

Прежде чем дать определение однородной и изотропной вселенной, напомним некоторые свойства пространств постоянной кривизны. В рассматриваемом случае под пространством постоянной кривизны мы понимаем пару $(\mathbb{S}, \overset{\circ}{g})$, где $\mathbb{S}$ – трехмерное многообразие и $\overset{\circ}{g}$ – заданная на нем метрика, которая удовлетворяет равенству

$$\overset{\circ}{R}_{\mu\nu\rho\sigma} = \overset{\circ}{K} \left(\overset{\circ}{g}_{\mu\rho}\overset{\circ}{g}_{\nu\sigma} - \overset{\circ}{g}_{\mu\sigma}\overset{\circ}{g}_{\nu\rho} \right), \quad (35.66)$$

где $\overset{\circ}{R}_{\mu\nu\rho\sigma}$ – тензор кривизны для метрики $\overset{\circ}{g}_{\mu\nu}$ и $\overset{\circ}{K} \in \mathbb{R}$ – некоторая постоянная (гауссова кривизна). Напомним, что на пространственных сечениях мы выбираем отрицательно определенную метрику, и поэтому для пространств положительной кривизны $\overset{\circ}{K} < 0$. Соответствующий тензор Риччи и скалярная кривизна имеют вид

$$\overset{\circ}{R}_{\mu\nu} = 2\overset{\circ}{K}\overset{\circ}{g}_{\mu\nu}, \quad \overset{\circ}{R} = 6\overset{\circ}{K}. \quad (35.67)$$

Замечание. В разделах 35.2.1 и 35.2.2 было показано, что метрика пространств постоянной кривизны в стереографических координатах является конформно плоской. Это значит, что тензор Коттона для метрики $\overset{\circ}{g}_{\mu\nu}$ равен нулю. В этом так же трудно убедиться с помощью прямых вычислений, используя явные выражения (35.66) и (35.67). Напомним, что тензор Вейля для трехмерных (псевдо)римановых многообразий тождественно равен нулю. $\square$

Координаты на $\mathbb{S}$ можно масштабировать таким образом, что гауссова кривизна $\overset{\circ}{K}$ будет принимать только три значения: $\overset{\circ}{K} = 1, 0, -1$. Следовательно, возможны три случая: $\overset{\circ}{K} = -1$, $\mathbb{S} \approx \mathbb{S}^3$ – трехмерная сфера (пространство положительной кривизны, замкнутая вселенная); $\overset{\circ}{K} = 0$, $\mathbb{S} \approx \mathbb{R}^3$ – трехмерное евклидово пространство (пространство нулевой кривизны, открытая вселенная) и $\overset{\circ}{K} = 1$, $\mathbb{S} \approx \mathbb{H}^3$ – трехмерное гиперболическое пространство или псевдосфера (пространство неотрицательной кривизны, открытая вселенная).

Трехмерное евклидово пространство $\mathbb{R}^3$ рассмотрено в разделе 1.8, трехмерная сфера $\mathbb{S}^3$ – в разделе 35.2.1, а трехмерное гиперболическое пространство $\mathbb{H}^3$ – в разделе 35.2.2. При этом под пространством $\mathbb{H}^3$ мы понимаем верхнюю полу́ двуполостного гиперболоида. Все три пространства постоянной кривизны $\mathbb{S} = \{\mathbb{S}^3, \mathbb{R}^3, \mathbb{H}^3\}$ являются связными и односвязными.

Отметим, что любое геодезически полное пространство постоянной кривизны является фактор пространством универсальной накрывающей $\mathbb{S} = \{\mathbb{S}^3, \mathbb{R}^3, \mathbb{H}^3\}$ по некоторой группе преобразований $\mathbb{G}$, действующей на $\mathbb{S}$ свободно и собственно разрывно (см., например, [54], следствие 2.4.10).

Если кривизна пространственных сечений равна нулю или отрицательна, то вселенная также может быть замкнута, если евклидово $\mathbb{R}^3$ или гиперболическое $\mathbb{H}^3$ пространства факторизовать по группе преобразований, действующей свободно и собственноразрывно. Например, для евклидова пространства $\mathbb{R}^3$ после факторизации мы получим трехмерный цилиндр или тор. Дальнейшее рассмотрение мы, для определенности, ограничим только универсальными накрывающими $\mathbb{S} = \{\mathbb{S}^3, \mathbb{R}^3, \mathbb{H}^3\}$.

Любое пространство постоянной кривизны является однородным и изотропным (см. раздел 15.3). Это значит, что на пространственных сечениях определена шести-параметрическая группа преобразований

$$\mathbb{G} = \begin{cases} \mathrm{SO}(4), & \mathbb{S} = \mathbb{S}^3 & \text{– сфера,} \\ \mathrm{ISO}(3), & \mathbb{S} = \mathbb{R}^3 & \text{– евклидово пространство,} \\ \mathrm{SO}(3, 1), & \mathbb{S} = \mathbb{H}^3 & \text{– гиперboloид.} \end{cases} \quad (35.68)$$

Следовательно, на каждом трехмерном пространстве постоянной кривизны $\mathbb{S}$ определены шесть линейно независимых векторных полей Киллинга $K_i = K_i^\mu(x)\partial_\mu$, $i = 1, \dots, 6$. В каждой точке $x \in \mathbb{S}$ три векторных поля Киллинга соответствуют трансляциям (однородность), а три – вращениям (изотропия).

Поскольку пространство-время четырехмерно, а группа преобразований определена только на пространственных сечениях $\mathbb{S} \hookrightarrow \mathbb{M}$, то ее действие надо доопределить на всем $\mathbb{M}$. Будем считать, что группа преобразований не действует на время (см. раздел 15.5):

$$\begin{aligned} t &\mapsto t' = t, \\ x^\mu &\mapsto x'^\mu = x^\mu + \epsilon K^\mu + o(\epsilon), \quad \epsilon \ll 1. \end{aligned} \quad (35.69)$$

Тогда соответствующие векторные поля Киллинга на $\mathbb{M}$ имеют равную нулю временную компоненту, $\{K_i^\alpha\} = \{0, K_i^\mu\}$, где пространственные компоненты $K_i^\mu(t, x)$ в общем случае зависят от времени. Ясно, что алгебра Ли группы преобразований осталась прежней, т.к. время t входит в компоненты векторных полей Киллинга K_i как параметр.

Теперь у нас есть пространство-время $(\mathbb{M}, g)$, на котором определена группа преобразований, порожденная инфинитезимальными преобразованиями (35.69), и мы в состоянии дать

Определение. Вселенная $(\mathbb{M}, g)$ называется *однородной и изотропной*, если

- 1) многообразие $\mathbb{M}$ представляет собой топологическое произведение $\mathbb{M} = \mathbb{R} \times \mathbb{S}$, где $t \in \mathbb{R}$ – время и $\mathbb{S}$ – трехмерное пространство постоянной кривизны с отрицательно определенной метрикой;
- 2) метрика g на $\mathbb{M}$ инвариантна относительно преобразований (35.69), порожденных преобразованиями симметрии для $\mathbb{S}$. $\square$

Из теорем 15.5.1 и 15.5.2 вытекает, что с точностью до преобразования координат наиболее общий вид метрики однородной и изотропной вселенной следующий

$$ds^2 = dt^2 + a^2 \overset{\circ}{g}_{\mu\nu} dx^\mu dx^\nu, \quad (35.70)$$

где $\overset{\circ}{g}_{\mu\nu}(x)$ – отрицательно определенная метрика на $\mathbb{S}$, не зависящая от времени. Функция времени $a(t) > 0$ предполагается достаточно гладкой и называется *масштабным множителем*. При этом векторные поля Киллинга не зависят от времени.

Как видим, требование однородности и изотропии вселенной приводит к тому, что вся зависимость метрики от времени содержится в масштабном множителе, который определяется уравнениями Эйнштейна.

Впервые метрику вида (35.70) для космологических моделей вселенных рассмотрел А. Фридман [278, 279], и она называется *метрикой Фридмана*.

Замечание. Напомним, что из изотропии вселенной следует ее однородность (теорема 15.3.1), поэтому можно было бы просто говорить об изотропной вселенной. Тем не менее мы будем употреблять термин “однородная и изотропная вселенная”, поскольку эпитет “однородная” отражает важные физические свойства вселенной, и это также принято в современной литературе. $\square$

Из вида метрики (35.70) следует, что однородная и изотропная вселенная является сплетенным произведением вещественной прямой $\mathbb{R}$ и пространства постоянной кривизны $\mathbb{S}$, которое было определено в разделе 33.1.

Метрика (35.70) имеет блочно диагональный вид, и, в соответствии с результатами раздела 23.13.2, часы во всей вселенной можно синхронизировать. Поэтому система координат $\{t, x^\mu\}$ является синхронной, и координата t называется *космологическим временем*. Ситуация здесь аналогична тому, что имеет место в механике Ньютона: наблюдаемое время t едино для всех точек вселенной. Сравнение наблюдательных космологических данных с теорией проводится, как правило, с использованием метрики в виде (35.70). К сожалению, эта процедура не является инвариантной, т.к. вид метрики зависит от выбора системы отсчета.

Метрика Фридмана (35.70) допускает по крайней мере шестипараметрическую группу симметрии (35.68), которая действует на пространственных сечениях $t = \text{const}$. Масштабный множитель находится из уравнений Эйнштейна, и в некоторых случаях группа симметрии пространства-времени может быть шире.

Дадим эквивалентное

Определение. Вселенная называется *однородной и изотропной*, если

- 1) все сечения постоянного времени, $t = \text{const}$, являются пространствами постоянной кривизны $\mathbb{S}$;
- 2) внешняя кривизна гиперповерхностей $\mathbb{S} \hookrightarrow \mathbb{M}$ является однородной и изотропной. $\square$

Внешняя кривизна вложенных гиперповерхностей определена в разделе 24.3. Для блочно диагональной метрики (15.31) внешняя кривизна (24.44) пропорциональна производной по времени от метрики

$$K_{\mu\nu} = -\frac{1}{2}\dot{h}_{\mu\nu},$$

где $h_{\mu\nu}(t, x)$ – пространственный блок метрики. Для метрики Фридмана $h_{\mu\nu} = a^2 g_{\mu\nu}$.

Последнее определение однородной и изотропной вселенной эквивалентно первому определению. Действительно, первое требование эквивалентно тому, что пространство-время является топологическим произведением $\mathbb{M} = \mathbb{R} \times \mathbb{S}$. После этого метрику можно привести к блочно диагональному виду (15.31). Затем второе условие в определении ведет к уравнению (15.36), и можно следовать доказательству теоремы 15.5.2.

Заметим, что второе условие в определении является необходимым, т.к. метрика (35.83), рассмотренная в следующем разделе дает контрпример.

Вычислим геометрические характеристики метрики (35.70), не фиксируя координат на пространствах постоянной кривизны. Простые вычисления показывают, что только четыре символа Кристоффеля отличны от нуля:

$$\begin{aligned}\Gamma_{0\mu}{}^\nu &= \Gamma_{\mu 0}{}^\nu = \frac{\dot{a}}{a}\delta_\mu^\nu, \\ \Gamma_{\mu\nu}{}^0 &= -a\dot{a}\overset{\circ}{g}_{\mu\nu}, \\ \Gamma_{\mu\nu}{}^\rho &= \overset{\circ}{\Gamma}_{\mu\nu}{}^\rho,\end{aligned}\tag{35.71}$$

где $\overset{\circ}{\Gamma}_{\mu\nu}{}^\rho$ – символы Кристоффеля для метрики $\overset{\circ}{g}_{\mu\nu}$ и точка обозначает дифференцирование по t . Тензор кривизны имеет только две отличные от нуля независимые компоненты:

$$\begin{aligned}R_{0\mu 0\nu} &= a\ddot{a}\overset{\circ}{g}_{\mu\nu}, \\ R_{\mu\nu\rho\sigma} &= a^2\overset{\circ}{R}_{\mu\nu\rho\sigma} + a^2\dot{a}^2(\overset{\circ}{g}_{\mu\rho}\overset{\circ}{g}_{\nu\sigma} - \overset{\circ}{g}_{\mu\sigma}\overset{\circ}{g}_{\nu\rho}),\end{aligned}\tag{35.72}$$

где $\overset{\circ}{R}_{\mu\nu\rho\sigma}$ – тензор кривизны (35.66) трехмерных пространственных сечений для метрики $\overset{\circ}{g}_{\mu\nu}$. Теперь вычислим тензор Риччи и скалярную кривизну:

$$\begin{aligned}R_{00} &= 3\frac{\ddot{a}}{a}, & R_{0\mu} &= R_{\mu 0} = 0, \\ R_{\mu\nu} &= \left(a\ddot{a} + 2\dot{a}^2 + 2\overset{\circ}{K}\right)\overset{\circ}{g}_{\mu\nu}, \\ R &= \frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + \overset{\circ}{K}),\end{aligned}\tag{35.73}$$

где мы воспользовались выражением для тензора Риччи пространства постоянной кривизны (35.67).

Зная тензор кривизны (35.72), нетрудно вычислить тензор Вейля. Оказывается, что он тождественно равен нулю. Следовательно, метрика Фридмана относится к типу 0 по классификации Петрова (см. раздел 17.8). Согласно теореме 17.5.1 вселенная Фридмана является конформно плоской.

Заметим, что скалярная кривизна (35.73), которая является геометрическим инвариантом, не зависит от пространственных координат. Это находится в согласии с требованием симметрии скалярных полей относительно действия транзитивных групп преобразований.

Конкретный вид метрики пространства постоянной кривизны $\overset{\circ}{g}_{\mu\nu}$ зависит от системы координат, выбранной на $\mathbb{S}$. Часто используются безразмерные сферические координаты χ, θ, φ , в которых метрика (положительно определенная) имеет вид

$$\overset{\circ}{g}_{\mu\nu}dx^\mu dx^\nu = \begin{cases} d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta d\varphi^2), & \overset{\circ}{K} = 1, \\ d\chi^2 + \chi^2(d\theta^2 + \sin^2\theta d\varphi^2), & \overset{\circ}{K} = 0, \\ d\chi^2 + \text{sh}^2\chi(d\theta^2 + \sin^2\theta d\varphi^2), & \overset{\circ}{K} = -1. \end{cases}\tag{35.74}$$

Безразмерные сферические координаты для сферы $\mathbb{S}^3$ и двулопастного гиперболоида $\mathbb{H}^3$ были рассмотрены в разделах 35.2.1 и 35.2.2 соответственно. В этом случае масштабный множитель имеет размерность длины, $[a] = l$.

В стереографических координатах на сечениях $t = \text{const}$ метрика Фридмана примет вид

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{a^2 \eta_{\mu\nu}}{(1 + b_0 x^2)^2} \end{pmatrix}, \quad (35.75)$$

где $\eta_{\mu\nu} := \text{diag}(- - -)$ – отрицательно определенная евклидова метрика, и постоянная b_0 связана с гауссовой кривизной простым соотношением $\overset{\circ}{K} = 4b_0$ (напомним, что, поскольку метрика $\eta_{\mu\nu}$ на пространственных сечениях отрицательно определена, то $x^2 := x^\mu x^\nu \eta_{\mu\nu} \leq 0$). Метрике (35.75) соответствует интервал

$$ds^2 = dt^2 + a^2 \frac{\eta_{\mu\nu} dx^\mu dx^\nu}{(1 + b_0 x^2)^2}.$$

В такой записи размерности следующие:

$$[t] = [x^\mu] = l, \quad [a] = 1, \quad [b_0] = l^{-2}.$$

Для пространств положительной ($b_0 < 0$) и нулевой ($b_0 = 0$) кривизны стереографические координаты определены во всем евклидовом пространстве $x \in \mathbb{R}^3$. Для пространств отрицательной кривизны ($b_0 > 0$) координаты определены внутри шара $|x^2| < 1/b_0$.

Стереографические координаты удобны для проведения вычислений. Поэтому вычислим геометрические характеристики для метрики Фридмана в виде (35.75). Отличные от нуля символы Кристоффеля равны

$$\begin{aligned} \Gamma_{0\mu}{}^\nu &= \Gamma_{\mu 0}{}^\nu = \frac{\dot{a}}{a} \delta_\mu^\nu, \\ \Gamma_{\mu\nu}{}^0 &= -\frac{\dot{a}a}{(1 + b_0 x^2)^2} \eta_{\mu\nu}, \\ \Gamma_{\mu\nu}{}^\rho &= \overset{\circ}{\Gamma}_{\mu\nu}{}^\rho. \end{aligned} \quad (35.76)$$

Прямые вычисления показывают, что только две независимые компоненты тензора кривизны отличны от нуля:

$$\begin{aligned} R_{0\mu 0\nu} &= \frac{\ddot{a}a}{(1 + b_0 x^2)^2} \eta_{\mu\nu}, \\ R_{\mu\nu\rho\sigma} &= a^2 \overset{\circ}{R}_{\mu\nu\rho\sigma} + \frac{\dot{a}^2 a^2}{(1 + b_0 x^2)^2} (\eta_{\mu\rho} \eta_{\nu\sigma} - \eta_{\mu\sigma} \eta_{\nu\rho}), \end{aligned} \quad (35.77)$$

где $\overset{\circ}{R}_{\mu\nu\rho\sigma}$ – тензор кривизны (35.66), построенный по метрике $\overset{\circ}{g}_{\mu\nu}$.

Выпишем отличные от нуля компоненты тензора Риччи и скалярную кривизну в стереографических координатах:

$$\begin{aligned} R_{00} &= 3\frac{\ddot{a}}{a}, & R_{0\mu} &= R_{\mu 0} = 0, \\ R_{\mu\nu} &= \frac{\ddot{a}a + 2\dot{a}^2 - 8b_0}{(1 + b_0 x^2)^2} \eta_{\mu\nu}, \\ R &= \frac{6\ddot{a}a + 6\dot{a}^2 - 24b_0}{a^2}. \end{aligned} \quad (35.78)$$

В дальнейшем метрика Фридмана, как правило, будет использоваться в общем виде (35.70) без использования конкретной системы координат на пространственных сечениях. Этого оказывается достаточно для анализа уравнений Эйнштейна.

Для исследования свойств метрики Фридмана часто используется другая система координат в пространстве-времени, в которой метрика Фридмана конформно эквивалентна некоторой статической метрике. А именно, совершим преобразование временной координаты $t \mapsto \eta$, где монотонная функция $t = t(\eta)$ определяется дифференциальным уравнением

$$dt = a(t)d\eta. \quad (35.79)$$

Это уравнение при достаточно общих предположениях имеет единственное решение с точностью до сдвига новой временной координаты: $\eta \mapsto \eta + \text{const}$. После преобразования временной координаты интервал принимает вид

$$ds^2 = a^2 \left(d\eta^2 + \overset{\circ}{g}_{\mu\nu} dx^\mu dx^\nu \right), \quad (35.80)$$

где $a = a(t(\eta))$ – функция нового времени η . Метрика, стоящая в скобках, статична, и вся зависимость от времени вынесена в общий конформный множитель. В таком виде удобно исследовать причинную структуру пространства-времени, т.к. световые конусы для метрики $g_{\alpha\beta}$ те же, что и для статической метрики, стоящей в скобках.

35.3.2 Пример

Рассмотрим поучительный пример метрики, для которой пространственные сечения для всех моментов времени являются пространствами постоянной кривизны, а метрика в целом, тем не менее, не является однородной и изотропной.

Пусть метрика Фридмана задана в стереографических координатах (35.75). Совершим преобразование координат $x^\mu \mapsto x^\mu/a$. Тогда метрика вселенной перестанет быть диагональной, а масштабный множитель исчезнет:

$$g = \begin{pmatrix} 1 + \frac{\dot{b}^2 x^2}{4b^2(1+bx^2)^2} & \frac{\dot{b}x_\nu}{2b(1+bx^2)^2} \\ \frac{\dot{b}x_\mu}{2b(1+bx^2)^2} & \frac{\eta_{\mu\nu}}{(1+bx^2)^2} \end{pmatrix}, \quad (35.81)$$

где

$$b(t) := \frac{b_0}{a^2(t)}, \quad (35.82)$$

и точка обозначает дифференцирование по времени. Мы видим, что метрика однородной и изотропной вселенной может быть недиагональной и не содержать масштабного множителя. При этом скалярная кривизна пространственных сечений, которая пропорциональна $b(t)$, явно зависит от времени.

Теперь просто отбросим недиагональные члены, положим $g_{00} = 1$ и восстановим масштабный множитель. Тогда метрика примет вид

$$g = \begin{pmatrix} 1 & 0 \\ 0 & \frac{a^2 \eta_{\mu\nu}}{(1+bx^2)^2} \end{pmatrix}. \quad (35.83)$$

Эта метрика содержит две произвольные функции времени: $a(t) > 0$ и $b(t)$. Она невырождена для всех значений b , включая нуль. При этом все пространственные

сечения $t = \text{const}$ являются пространствами постоянной кривизны и, следовательно, однородны и изотропны. Метрика (35.83) представляет самостоятельный интерес, т.к. появляется возможность исследовать такие решения, если они существуют, для которых функция $b(t)$ проходит через нуль. Это означает, что соответствующие решения уравнений Эйнштейна описывают изменение топологии пространственных сечений в процессе эволюции. Таких решений в настоящее время не найдено.

Отметим, что произвольную функцию $b(t)$ нельзя обратить в константу путем преобразования координат без появления недиагональных членов в метрике. Поэтому возникает парадокс: с одной стороны мы доказали, что наиболее общий вид метрики для однородной и изотропной вселенной содержит одну произвольную функцию – масштабный множитель (теорема 15.5.2). С другой стороны, метрика (35.83) содержит две произвольные и независимые функции. При этом все пространственные сечения $t = \text{const}$ однородны и изотропны.

Решение этого парадокса следующее. Пространственные компоненты (μ, ν) уравнений Киллинга (15.34) удовлетворяются, поскольку все пространственные сечения – пространства постоянной кривизны. Однако смешанные компоненты $(0, \mu)$ уравнений Киллинга (15.33) не удовлетворены. Действительно, в стереографических координатах три векторных поля Киллинга зависят от функции $b(t)$ и, следовательно, от времени t (35.51) (в этих формулах a – это радиус сферы, определяющий кривизну, а не масштабный множитель). Это означает, что второе утверждение теоремы (15.5.1) не выполнено. Поэтому метрика (35.83) не описывает однородную и изотропную вселенную.

Существует и другой способ убедиться в том, что метрика (35.83) не является однородной и изотропной. С этой целью вычислим скалярную кривизну

$$R = -\frac{24b}{a^2} + 6 \left[\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} - \frac{1}{1+bx^2} \left(4\frac{\dot{a}\dot{b}x^2}{a} + \ddot{b}x^2 \right) + 3\frac{\dot{b}^2x^4}{(1+bx^2)^2} \right].$$

Она явно зависит от x , и, следовательно, не является однородной и изотропной функцией.

35.3.3 Космологическое красное смещение

Если масштабный множитель меняется со временем, $\dot{a} \neq 0$, то расстояние между двумя наблюдателями, имеющими фиксированные пространственные координаты, скажем, $\mathbf{x}_1$ и $\mathbf{x}_2$, также меняется со временем. Это приводит к смещению частот световых сигналов. Если вселенная расширяется, $\dot{a} > 0$, то, как мы увидим, возникает красное смещение спектральных линий.

Свет распространяется вдоль нулевых экстремалей $\{t(\tau), x^\mu(\tau)\}$, $\tau \in \mathbb{R}$, при этом 4-вектор скорости $\{t', x^{\mu'}\}$, где штрих обозначает дифференцирование по каноническому параметру τ (точка зарезервирована для дальнейших обозначений), представляет собой 4-импульс фотона. Нулевая компонента импульса является энергией фотона и, следовательно, пропорциональна частоте фотона ω :

$$t' = \hbar\omega,$$

где $\hbar$ – постоянная Планка. При этом длина волны фотона λ связана с частотой равенством

$$\lambda\omega = 2\pi c,$$

где c – скорость света.

Допустим, что мы находимся в точке x_0 и наблюдаем световой сигнал, испущенный кем-то в точке x_1 . Тогда величина красного смещения z определяется относительным смещением длин волн

$$z := \frac{\lambda_0 - \lambda_1}{\lambda_1},$$

что можно выразить через нулевую компоненту скорости:

$$z = \frac{\omega_1 - \omega_0}{\omega_0} = \frac{t'_1 - t'_0}{t'_0}. \quad (35.84)$$

Фотоны распространяются в пространстве-времени вдоль экстремалей $t(\tau), x^\mu(\tau)$. Уравнения для экстремалей определяются символами Кристоффеля (35.71) и имеют вид

$$\begin{aligned} t'' &= a \dot{a} g_{\mu\nu} x^{\mu'} x^{\nu'}, \\ x^{\mu''} &= -2 \frac{\dot{a}}{a} x^{\mu'} t' - \dot{\Gamma}_{\nu\rho}^{\mu} x^{\nu'} x^{\rho'}, \end{aligned} \quad (35.85)$$

где точка обозначает дифференцирование по t . Поскольку фотоны распространяются вдоль нулевых экстремалей, то для них выполнено равенство

$$t'^2 + a^2 g_{\mu\nu} x^{\mu'} x^{\nu'} = 0.$$

Подстановка этого выражения в правую часть первого уравнения (35.85) приводит к равенству

$$t'' + \frac{\dot{a}}{a} t'^2 = 0.$$

Разделим полученное соотношение на t' и проинтегрируем вдоль экстремали. В результате получим условие на нулевую компоненту 4-скорости фотона:

$$at' = \text{const}.$$

Отсюда следует, что красное смещение можно выразить через масштабный множитель:

$$z = \frac{a_0 - a_1}{a_1}. \quad (35.86)$$

Это – универсальная и точная формула для красного смещения. Она определяется масштабным множителем a_1 в момент испускания света и a_0 в момент наблюдения, и не зависит от истории сигнала в промежуточные моменты времени. Из полученной формулы вытекает, что относительное смещение длин волн положительно, если вселенная расширяется, и отрицательно – если сжимается.

Другое важное следствие из вида метрики Фридмана – это связь красного смещения с расстоянием до объекта. Для света, испущенного и полученного в близких пространственно-временных точках, справедливо равенство

$$dt^2 = dl^2, \quad c = 1,$$

где dl – пространственное расстояние между событиями:

$$dl^2 := -a^2 g_{\mu\nu} dx^\mu dx^\nu.$$

Поэтому из формулы (35.86) следуют равенства:

$$z := \frac{d\lambda}{\lambda} = \frac{da}{a} = \frac{\dot{a}dt}{a} = \frac{\dot{a}}{a}dl.$$

Введем *параметр Хаббла*

$$H := \left. \frac{\dot{a}}{a} \right|_{t=t_0}. \quad (35.87)$$

Тогда красное смещение спектральных линий в настоящее время t_0 определяется *законом Хаббла*

$$\frac{d\lambda}{\lambda} = Hdl. \quad (35.88)$$

Содержание этого закона состоит в том, что относительное красное смещение спектральных линий прямо пропорционально расстоянию до объекта. Этот закон получен для бесконечно близких событий. На практике он приближенно справедлив для конечных расстояний, $dl \mapsto \Delta l$. Ранее точность измерений была невысока, и считалось, что параметр Хаббла является постоянным. Поэтому до сих пор его часто называют *постоянной Хаббла*. Современные наблюдательные данные показывают, что параметр Хаббла зависит от времени и “постоянная Хаббла” не является постоянной.

Если параметр Хаббла является постоянным, то из уравнения (35.87) следует экспоненциальное поведение масштабного множителя

$$a = C e^{Ht}, \quad C > 0.$$

Если $H > 0$, то происходит экспоненциальный рост. Для решения некоторых космологических проблем, которые мы обсуждать не будем, предполагают, что масштабный множитель, по крайней мере на ранних этапах эволюции вселенной, экспоненциально растет. Такой экспоненциальный рост масштабного множителя называется *инфляцией*. Предположение об инфляционной стадии развития вселенной в настоящее время широко распространено. Моделей, приводящих к инфляции, не так много. Позже мы увидим, что это происходит, например, в решении де Ситтера (35.129).

Наша интуиция основана на механике Ньютона. Поэтому посмотрим на наблюдаемое красное смещение спектральных линий с другой точки зрения. Пусть галактики движутся в плоском пространстве-времени Минковского $\mathbb{R}^{1,3}$. Если галактика удаляется, то наблюдаемая длина волны светового сигнала будет увеличиваться, а частота – уменьшаться (красное смещение). Это явление известно как эффект Доплера и описывается формулой (1.221), полученной ранее. В рассматриваемом случае

$$\omega_1 = \omega', \quad \omega_0 = \omega \quad \text{и} \quad \alpha = 0.$$

Тогда в нерелятивистском пределе $V/c \ll 1$ частота в первом порядке меняется по правилу

$$\omega_1 = \omega_0 \left(1 + \frac{V}{c} \right),$$

где V – скорость галактики, и c – скорость света. Подстановка этого выражения в формулу для красного смещения (35.84) приводит к равенству

$$z = \frac{V}{c}. \quad (35.89)$$

Дальнейшее сравнение с законом Хаббла позволяет найти скорость галактики

$$V = cH\Delta l,$$

где Δl – расстояние между галактиками. Мы видим, что наблюдаемая скорость разбегания галактик прямо пропорциональна расстоянию.

В заключение вычислим наблюдаемую скорость разбегания галактик из других соображений. Согласно предложению 35.3.1, которое будет доказано в следующем разделе, в однородной и изотропной вселенной частицы среды (галактики) движутся вдоль экстремалей, совпадающих с координатными линиями времени. То есть пространственные координаты галактик x^μ остаются неизменными. Поэтому наблюдаемое расстояние до галактики равно

$$\Delta l = a(t) \sqrt{-\overset{\circ}{g}_{\mu\nu} \Delta x^\mu \Delta x^\nu}.$$

Это расстояние меняется только благодаря наличию масштабного множителя. Вычислим наблюдаемую скорость

$$V := \frac{d}{dt} \Delta l = c \frac{\dot{a}}{a} \Delta l = cH \Delta l, \quad (35.90)$$

где мы восстановили скорость света, $x^0 := ct$, и $\dot{a} := da/dx^0$. Как видим, полученная формула совпадает с той, что была получена из закона Хаббла.

35.3.4 Уравнения движения

Масштабный множитель, входящий в метрику Фридмана, находится из уравнений Эйнштейна и зависит от выбора модели для полей материи. В дальнейшем мы рассмотрим несколько моделей вселенной, для которых масштабный множитель имеет явное аналитическое выражение.

Прежде всего получим уравнения Эйнштейна. Из компонент тензора Риччи и скалярной кривизны (35.73) составляем тензор Эйнштейна. Он имеет следующие компоненты:

$$\begin{aligned} G_0^0 &= -3 \frac{\epsilon + \dot{a}^2}{a^2}, \\ G_0^\mu &= 0, \quad G_\mu^0 = 0, \\ G_\mu^\nu &= -\frac{2a\ddot{a} + \dot{a}^2 + \epsilon}{a^2} \delta_\mu^\nu, \end{aligned} \quad (35.91)$$

где постоянная ϵ , которая введена для удобства, равна гауссовой кривизне с обратным знаком, что соответствует положительно, а не отрицательно определенной метрике на пространственных сечениях:

$$\epsilon := -\overset{\circ}{K} = -4b_0. \quad (35.92)$$

Постоянная b_0 была определена стереографическими координатами (35.75). Пространственные сечения $t = \text{const}$ при этом являются сферой $\mathbb{S}^3$ ($\epsilon > 0$), евклидовым пространством $\mathbb{R}^3$ ($\epsilon = 0$) или гиперboloидом $\mathbb{H}^3$ ($\epsilon < 0$).

Мы записываем компоненты тензора Эйнштейна (35.91) с одним ковариантным и одним контравариантным индексом. Это удобно, т.к. в этом случае тензор энергии-импульса, как мы увидим, является диагональным.

Следующий шаг состоит в выборе тензора энергии-импульса материи. При построении моделей вселенной мы используем космологический принцип, который состоит из двух независимых требований:

- Метрика пространства-времени должна быть однородна и изотропна.
- Тензор энергии-импульса материи должен быть однороден и изотропен.

Первому требованию мы уже удовлетворили, выбрав метрику Фридмана (35.70).

Наиболее общий вид однородных и изотропных функций, векторов и тензоров второго ранга был установлен в разделе 15.4. В общей теории относительности тензор энергии-импульса $T_{\alpha\beta}$ является четырехмерным симметричным тензором второго ранга. Группа изометрий действует только на пространственных сечениях, и по отношению к ее действию тензор энергии-импульса разлагается на три неприводимые компоненты:

$$\begin{aligned} T_{00} & - \text{скаляр,} \\ T_{\mu 0} = T_{0\mu} & - \text{вектор,} \\ T_{\mu\nu} & - \text{симметричный тензор второго ранга.} \end{aligned} \quad (35.93)$$

Из результатов, полученных в разделе 15.4, следует, что наиболее общий вид инвариантных компонент следующий

$$T_{00} = \mathcal{E}(t), \quad T_{0\mu} = T_{\mu 0} = 0, \quad T_{\mu\nu} = -\mathcal{P}(t)\overset{\circ}{g}_{\mu\nu},$$

где $\mathcal{E}$ и $\mathcal{P}$ – некоторые функции (скалярные поля). Сравнение этих компонент с тензором энергии-импульса сплошной среды

$$T^{\alpha\beta} = (\mathcal{E} + \mathcal{P})u^\alpha u^\beta - \mathcal{P}g^{\alpha\beta}, \quad (35.94)$$

показывает, что $\mathcal{E}$ и $\mathcal{P}$ есть ни что иное как плотность энергии и давление материи соответственно, и тензор энергии-импульса записан в сопутствующей системе координат $u = (1, 0, 0, 0)$.

Предложение 35.3.1. *Если u метрика, u тензор энергии-импульса среды однородны и изотропны, то линии тока непрерывной среды являются геодезическими (экстремалами).*

Доказательство. Зафиксируем систему координат так, чтобы метрика была метрикой Фридмана (35.70) и тензор энергии-импульса имел вид (35.93). Символы Кристоффеля были вычислены ранее (35.71), и уравнения для экстремалей имеют вид (35.85). Ясно, что координатные линии $x^0 = \tau, x^\mu = \text{const}$ являются экстремалами, где $\tau \in \mathbb{R}$ – канонический параметр. С другой стороны, эти координатные линии представляют собой линии тока для векторного поля скорости $u = (1, 0, 0, 0)$. $\square$

Таким образом, в рассматриваемой космологической модели выбрана синхронная система координат, в которой метрика имеет вид (35.70), и каждая точка сплошной среды движется вдоль экстремали, которая совпадает с координатной линией $x^0 := t$. Точки среды покоятся в том смысле, что пространственные координаты имеют постоянные значения $x^\mu = \text{const}$. При этом наблюдаемое расстояние между частицами среды может увеличиваться или уменьшаться в зависимости от поведения масштабного множителя.

Здесь возникает вопрос. Если присутствует давление, то в общем случае частицы среды не будут двигаться по экстремалам, т.к. присутствуют не только гравитационные силы. Тем не менее рассматриваемая модель является самосогласованной. Действительно, поскольку давление зависит только от времени, то градиент давления

параллелен вектору скорости. Следовательно, согласно уравнениям релятивистской гидродинамики (23.135), (23.136), движение частиц происходит вдоль экстремалей. Причиной этого является требование однородности и изотропии тензора энергии-импульса полей материи.

Приведенные выше аргументы однозначно фиксируют однородный и изотропный тензор энергии-импульса материи. С одним контра- и одним ковариантным индексом он всегда диагонален:

$$T_{\alpha}^{\beta} = \begin{pmatrix} \mathcal{E} & 0 & 0 & 0 \\ 0 & -\mathcal{P} & 0 & 0 \\ 0 & 0 & -\mathcal{P} & 0 \\ 0 & 0 & 0 & -\mathcal{P} \end{pmatrix}. \quad (35.95)$$

Уравнения релятивистской гидродинамики $\nabla_{\beta} T_{\alpha}^{\beta} = 0$, выполнение которых необходимо для разрешимости уравнений Эйнштейна, для метрики Фридмана (35.75) принимают вид

$$\dot{\mathcal{E}} + \frac{3\dot{a}}{a}(\mathcal{E} + \mathcal{P}) = 0. \quad (35.96)$$

Если подставить выражение для тензора энергии-импульса (35.95) в уравнения Эйнштейна с космологической постоянной Λ ,

$$\Phi_{\alpha}^{\beta} := G_{\alpha}^{\beta} + \delta_{\alpha}^{\beta} \Lambda + \frac{1}{2} T_{\alpha}^{\beta} = 0,$$

где мы для простоты положили гравитационную постоянную равной единице, $\kappa = 1$, и учесть явный вид тензора Эйнштейна (35.91), то получим два уравнения:

$$-3 \frac{\dot{a}^2 + \epsilon}{a^2} + \Lambda + \frac{1}{2} \mathcal{E} = 0, \quad (35.97)$$

$$-\frac{2a\ddot{a} + \dot{a}^2 + \epsilon}{a^2} + \Lambda - \frac{1}{2} \mathcal{P} = 0. \quad (35.98)$$

Продифференцируем уравнение (35.97) по времени t и учтем уравнение совместности (35.96). В результате получим, что уравнение (35.98) тождественно удовлетворяется. Можно проверить, что это свойство является следствием линейной зависимости уравнений Эйнштейна (вторая теорема Нетер)

$$\nabla_{\beta} \Phi_{\alpha}^{\beta} = 0,$$

вытекающей из ковариантности (см. раздел 23.5).

Таким образом, у нас есть два независимых обыкновенных дифференциальных уравнения (35.97) и (35.96) на три неизвестные функции $a(t)$, $\mathcal{E}(t)$ и $\mathcal{P}(t)$. Чтобы замкнуть систему уравнений будем считать, что давление среды связано с плотностью энергии уравнением состояния. Для баротропной жидкости уравнение состояния имеет вид

$$\mathcal{P} = \mathcal{P}(\mathcal{E}), \quad (35.99)$$

где $\mathcal{P}(\mathcal{E})$ – некоторая достаточно гладкая функция. Это равенство корректно, т.к. плотность энергии и давление – скалярные поля (функции).

Следовательно, для вселенной Фридмана мы имеем два независимых обыкновенных дифференциальных уравнения на две независимые функции: масштабный

множитель a и плотность энергии $\mathcal{E}$, которые мы запишем в следующем виде

$$\dot{a}^2 = \frac{a^2}{3} \left(\frac{1}{2} \mathcal{E} + \Lambda \right) - \epsilon, \quad (35.100)$$

$$\frac{d\mathcal{E}}{\mathcal{E} + \mathcal{P}(\mathcal{E})} = -3 \frac{da}{a}. \quad (35.101)$$

Это и есть полная система уравнений для космологических моделей Фридмана.

Таким образом, при заданном уравнении состояния (35.99), решая уравнение (35.101), мы находим зависимость плотности энергии от масштабного множителя: $\mathcal{E} = \mathcal{E}(a)$. Подставляя эту функцию в уравнение (35.100), мы получаем одно обыкновенное дифференциальное уравнение первого порядка на масштабный множитель. Это уравнение является основным в стандартных моделях однородной и изотропной вселенной.

Для интегрирования уравнения (35.100) удобно ввести временную координату $\eta = \eta(t)$, определенную равенством (35.79). Тогда уравнение (35.100) можно формально проинтегрировать

$$\eta = \pm \int \frac{da}{a \sqrt{\frac{1}{6} \mathcal{E} a^2 + \frac{1}{3} \Lambda a^2 - \epsilon}}. \quad (35.102)$$

Постоянная интегрирования соответствует сдвигу временной координаты $\eta \mapsto \eta + \text{const}$ и является несущественной. Знак $\pm$ соответствует выбору направления течения времени: если функция $a(t)$ является решением уравнений Фридмана, то $a(-t)$ – также решение. Зависимость плотности энергии от масштабного множителя, $\mathcal{E} = \mathcal{E}(a)$, находится из уравнения (35.101). Это решает поставленную задачу в общем виде. На практике, уравнения (35.101) и (35.102) решаются в явном виде далеко не всегда.

Не смотря на то, что одно из уравнений Эйнштейна (35.98) было отброшено как следствие остальных, из него можно сделать важные выводы. Перепишем уравнение (35.98) в виде

$$\ddot{a} = -\frac{a}{12}(\mathcal{E} + 3\mathcal{P}) + \frac{a\Lambda}{3}, \quad (35.103)$$

где мы использовали уравнение (35.97) для исключения $\dot{a}^2 + \epsilon$. Пусть $\Lambda = 0$ и $\mathcal{E} > 0$ и $\mathcal{P} \geq 0$, что соответствует обычной материи. По предположению, масштабный множитель положителен, $a > 0$, и, следовательно, $\ddot{a} < 0$. Это означает, что функция $a(t)$ вогнутая. Допустим, что в настоящее время t_0 вселенная расширяется, $\dot{a}(t_0) > 0$, что соответствует современным наблюдательным данным. Качественное поведение масштабного множителя в этом случае показано на рис. 35.2. Мы видим, что при

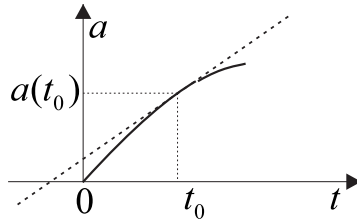


Рис. 35.2: Качественное поведение масштабного множителя $a(t)$ в предположении, что вселенная в настоящее время расширяется, $\dot{a}(t_0) > 0$.

конечном значении времени t масштабный множитель должен обратиться в нуль,

что соответствует большому взрыву. Этот вывод является довольно общим, т.к. не зависит от уравнения состояния материи и не использует какое-либо из решений уравнений Эйнштейна. В космологии принято отсчитывать наблюдаемое время t от большого взрыва при $t = 0$.

Из уравнений Эйнштейна (35.97), (35.98) следует, что наличие космологической постоянной можно трактовать как поля материи с диагональным тензором энергии-импульса

$$T_{\Lambda\alpha}{}^{\beta} := 2\Lambda \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Эта материя необычна, т.к. при $\Lambda > 0$ плотность энергии положительна, а давление отрицательно. Материя, обладающая таким свойством, называется *темной энергией*.

Понятие темной энергии играет важную роль в космологии. В 1998 году Сол Перлмуттер [280], Брайан П. Шмидт и Адам Рисс [281] открыли, что в настоящее время вселенная не просто расширяется, а расширяется с ускорением. За это открытие они получили Нобелевскую премию в 2011 году. Если принять модель вселенной Фридмана за основу, то из уравнения (35.103) вытекает, что обычная материя, $\mathcal{E} > 0$ и $\mathcal{P} > 0$, приводит к замедлению, а положительная космологическая постоянная – к ускорению. Таким образом, к ускорению приводит либо положительная космологическая постоянная, либо какая-то другая материя с отрицательным давлением (плотность энергии всегда считается положительной), которая была названа темной энергией. В настоящее время к темной энергии относят положительную космологическую постоянную и (или) другие модели материи, обсуждение которых выходит за рамки тех ограничений, которые поставил перед собой автор этих строк. Наличие темной энергии приводит к антигравитации: отталкиванию вместо притяжения. Это отталкивание и приводит к ускоренному расширению вселенной.

Существующие в настоящее время наблюдательные данные не противоречат представлению о темной энергии, как о положительной космологической постоянной. Поэтому данная точка зрения считается стандартной в современной космологии.

Замечание. Помимо темной энергии в космологии используется понятие *темной материи*. Это – гипотетическая материя, которая не взаимодействует с электромагнитным полем и, следовательно, не видна, т.к. не излучает фотонов. Понятие темной материи стало обсуждаться с 1922 года, когда Джеймс Джинс [282] и Якобус Каптейн [283] пришли к выводу, что большая часть вещества в нашей галактике (Млечный Путь) невидима, т.к. гравитационного поля всех видимых звезд явно недостаточно для удержания звезд, находящихся на периферии. Впоследствии нехватка видимого вещества была подтверждена для многих других галактик. Кандидатов на роль темной материи много, и мы их обсуждать не будем.

В настоящее время наблюдательные данные указывают на то, что темная энергия составляет более 70% всего вещества вселенной, а темная материя – более 20%. Если учесть еще межгалактический газ, то на нашу долю (звезды, планеты и др.) остается меньше половины процента всего вещества вселенной. $\square$

35.3.5 Вселенная Эйнштейна

Одну из первых космологических моделей в рамках общей теории относительности предложил А. Эйнштейн [284] в 1917 году задолго до статей А. Фридмана. Он нашел точное статическое решение для однородной и изотропной замкнутой вселенной, все пространственные сечения которой являются сферами, $\mathbb{S} = \mathbb{S}^3$. Это решение укладывается в общую схему однородной и изотропной вселенной с метрикой Фридмана.

Космологические уравнения Фридмана (35.97), (35.98) допускают статические решения. Положим

$$a = \text{const}, \quad \mathcal{E} = \text{const}, \quad \mathcal{P} = \text{const}.$$

Тогда уравнения (35.97) и (35.98) определяют равновесное значение плотности энергии и давления

$$\begin{aligned} \mathcal{E} &= 2 \left(3 \frac{\epsilon}{a^2} - \Lambda \right), \\ \mathcal{P} &= 2 \left(-\frac{\epsilon}{a^2} + \Lambda \right). \end{aligned}$$

Эти уравнения имеют много решений в зависимости от значений постоянных ϵ и Λ (для замкнутых вселенных $\epsilon > 0$). Для обычной материи $\mathcal{P} \leq \mathcal{E}/3$. Отсюда вытекает ограничение сверху на космологическую постоянную

$$\Lambda \leq \frac{3\epsilon}{2a^2}.$$

С другой стороны, давление обычной среды положительно, $\mathcal{P} \geq 0$. Это дает оценку снизу

$$\frac{\epsilon}{a^2} \leq \Lambda.$$

Вместе получаем ограничение на значение космологической постоянной при заданных кривизне ϵ и масштабном множителе a :

$$\frac{\epsilon}{a^2} \leq \Lambda \leq \frac{3}{2} \frac{\epsilon}{a^2}. \quad (35.104)$$

Как видим, для обычной материи это неравенство имеет нетривиальные решения только при $\epsilon > 0$ (замкнутая вселенная) и положительной космологической постоянной Λ .

Метрику вселенной Эйнштейна можно записать в виде

$$ds^2 = dt^2 - a^2 \left[\frac{dr^2}{1 - \epsilon r^2} + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \right]. \quad (35.105)$$

Из формул (35.72) и (35.66) вытекает, что у полного тензора кривизны для модели Эйнштейна отличны от нуля только пространственные компоненты

$$R_{\mu\nu\rho\sigma} = a^2 \overset{\circ}{R}_{\mu\nu\rho\sigma} = a^2 \epsilon \left(\overset{\circ}{g}_{\mu\rho} \overset{\circ}{g}_{\nu\sigma} - \overset{\circ}{g}_{\mu\sigma} \overset{\circ}{g}_{\nu\rho} \right).$$

В своей статье [284] А. Эйнштейн нашел решение для пыли, $\mathcal{P} = 0$, в замкнутой вселенной, $\epsilon > 0$. В этом случае $\epsilon = a^2 \Lambda$, и была отмечена необходимость введения положительной космологической постоянной.

Наблюдательные данные последних лет говорят о том, что вселенная расширяется. Это не укладывается в статическую модель, которая рассмотрена выше. Поэтому статическая модель вселенной Эйнштейна в настоящее время носит теоретический характер: необходимо знать, что уравнения общей теории относительности допускают такие решения.

35.3.6 Линейное уравнение состояния

Согласно современным наблюдательным данным вселенная расширяется. Это приводит к необходимости построения нестационарных моделей вселенных.

Для того, чтобы построить конкретную космологическую модель и решить уравнения Фридмана (35.100), (35.101), необходимо задать уравнение состояния $\mathcal{P} = \mathcal{P}(\mathcal{E})$. Важный класс космологических моделей описывается линейным уравнением состояния. Например, когда вселенная заполнена идеальным газом с уравнением состояния (23.132), пылью или излучением (23.128). Для упрощения некоторых последующих формул запишем линейное уравнение состояния в виде

$$\mathcal{P} = (\gamma - 1)\mathcal{E}, \quad \gamma = \text{const.} \quad (35.106)$$

Для обычной материи $1 \leq \gamma \leq 4/3$. Граничное значение $\gamma = 1$ соответствует пыли ($\mathcal{P} = 0$), а $\gamma = 4/3$ – излучению ($\mathcal{P} = \mathcal{E}/3$).

Часто говорят, что в случае пыли вселенная заполнена *холодной материей*. Это значит, что вселенная состоит из массивных частиц, которые движутся с нерелятивистскими скоростями, а безмассовыми частицами (излучением) можно пренебречь. В противоположном случае излучения, $\gamma = 4/3$, мы говорим, что вселенная заполнена *горячей материей*. Частицы горячей материи движутся с около световыми скоростями и ведут себя так же, как и безмассовые. В настоящее время в качестве стандартной космологической модели рассматривается так называемая Λ CDM-модель. Эта модель содержит космологическую постоянную Λ и холодную материю (от английского сокращения CDM = Cold Dark Matter). Λ CDM-модель экономна, и ее следствия хорошо согласуются с современными наблюдательными данными.

Можно доказать, что скорость звука в непрерывной среде превышает скорость света при $\mathcal{P} > \mathcal{E}$. Поэтому на давление накладывают ограничение сверху $\mathcal{P} < \mathcal{E}$, что немного увеличивает интервал значений постоянной γ в уравнении состояния (35.106).

Для линейного уравнения состояния (35.106) уравнение (35.101) легко интегрируется

$$\mathcal{E} = \frac{C}{a^{3\gamma}}, \quad (35.107)$$

где $C > 0$ – постоянная интегрирования. Полученное решение имеет простой физический смысл: плотность энергии обратно пропорциональна объему вселенной в данный момент времени, a^3 , в степени γ . Для пыли ($\mathcal{P} = 0$, $\gamma = 1$) плотность энергии просто пропорциональна плотности числа частиц.

Для сопоставления предсказаний теории с наблюдательными данными, необходимо фиксировать постоянные интегрирования в решениях уравнений Фридмана. Обычно полагают, что в настоящее время $t = t_0$ масштабный множитель равен единице, $a(t_0) = 1$. Тогда постоянная интегрирования в (35.107) равна плотности энергии (плотности числа частиц) в настоящий момент времени, $C = \mathcal{E}_0$.

Теперь запишем уравнение Фридмана (35.100) в виде

$$\dot{a}^2 + V(a) = -\epsilon, \quad (35.108)$$

где

$$V(a) := -Ma^{-3\gamma+2} - La, \quad (35.109)$$

и

$$M := \frac{\mathcal{E}_0}{6} = \text{const}, \quad L := \frac{\Lambda}{3} = \text{const.} \quad (35.110)$$

Напомним, что $\epsilon := -\overset{\circ}{K}$, где $\overset{\circ}{K}$ – гауссова кривизна пространственных сечений $t = \text{const}$ с отрицательно определенной метрикой. Уравнение (35.108) имеет тот же вид, что и закон сохранения энергии в механике Ньютона при одномерном движении точечной частицы в потенциальном поле $V(a)$ с полной энергией $-\epsilon$. Поэтому, нарисовав потенциал для фиксированных значений постоянных M , L и γ , можно определить точки поворота и понять качественное поведение решений.

Продemonстрируем интегрирование уравнений модели Фридмана в простейшем случае пыли, когда давление равно нулю, $\mathcal{P} = 0$. В этом случае плотность энергии пропорциональна плотности частиц

$$\mathcal{E} = \frac{\mathcal{E}_0}{a^3}. \quad (35.111)$$

Положим, для простоты, космологическую постоянную равной нулю, $\Lambda = 0$. Тогда подстановка решения (35.111) в уравнение (35.108) дает простое уравнение на масштабный множитель

$$\dot{a}^2 - \frac{M}{a} = -\epsilon. \quad (35.112)$$

Решение данного уравнение удобно записать, используя вместо времени t координату η , определенную уравнением (35.79). Это уравнение имеет различные решения в зависимости от знака ϵ , т.е. в зависимости от знака кривизны пространственных сечений:

$$a = \begin{cases} \frac{M}{2\epsilon} [1 - \cos(\sqrt{\epsilon}\eta)], & \epsilon > 0, \\ \frac{M\eta^2}{4}, & \epsilon = 0, \\ \frac{M}{2|\epsilon|} [\text{ch}(\sqrt{|\epsilon|}\eta) - 1], & \epsilon < 0. \end{cases} \quad (35.113)$$

Постоянная интегрирования соответствует сдвигу $\eta \mapsto \eta + \text{const}$ и положена равной нулю. Параметр η связан с временем следующими соотношениями:

$$t = \begin{cases} \frac{M}{2\epsilon^{3/2}} [\sqrt{\epsilon}\eta - \sin(\sqrt{\epsilon}\eta)], & \epsilon > 0, \\ \frac{M\eta^3}{12}, & \epsilon = 0, \\ \frac{M}{2|\epsilon|^{3/2}} [\text{sh}(\sqrt{|\epsilon|}\eta) - \sqrt{|\epsilon|}\eta], & \epsilon < 0, \end{cases} \quad (35.114)$$

где постоянная интегрирования, соответствующая сдвигу времени, положена равной нулю. Поскольку уравнение (35.112) инвариантно относительно инверсии времени, $t \mapsto -t$, то каждому решению $a(t)$ соответствует также решение $a(-t)$. Таким образом, равенства (35.113) и (35.114) определяют масштабный множитель $a(t)$ в параметрическом виде. Качественное поведение масштабного множителя для пространственных сечений положительной, нулевой и отрицательной кривизны для пыли показано на рис. 35.3. Во всех случаях масштабный множитель равен нулю в момент большого взрыва при $t = 0$. В настоящее время при $t = t_0$ он положителен. Дальнейшая эволюция масштабного множителя зависит от кривизны пространственных сечений. Для открытых моделей вселенных $\epsilon \leq 0$ и масштабный множитель монотонно возрастает. Для замкнутых моделей вселенных $\epsilon > 0$, масштабный множитель возрастает, достигает своего максимального значения и затем начинает убывать до нулевого значения, которое соответствует большому сжатию.

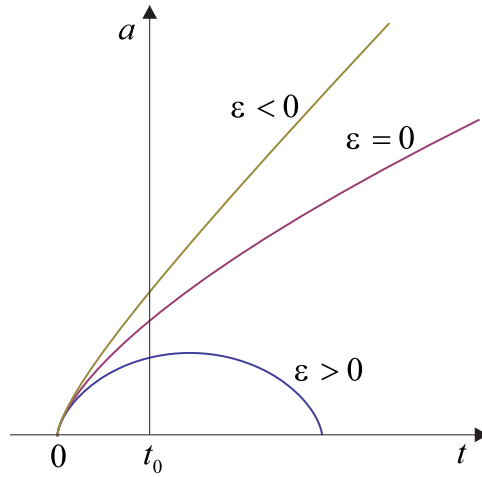


Рис. 35.3: Качественное поведение масштабного множителя для пространственных сечений положительной, нулевой и отрицательной кривизны для пыли и излучения.

Для излучения $\gamma = 4/3$ при нулевой космологической постоянной уравнение Фридмана (35.108) принимает вид

$$\dot{a}^2 - \frac{M}{a^2} = -\epsilon.$$

Это уравнение легко решается

$$a = \sqrt{t \left(2\sqrt{M} - \epsilon t \right)}. \quad (35.115)$$

Качественное поведение масштабного множителя для излучения такое же, как и для пыли, и показано на рис. 35.3 для трех случаев: $\epsilon < 0$, $\epsilon = 0$ и $\epsilon > 0$. Если вселенная замкнута, $\epsilon > 0$, то большой взрыв заканчивается большим сжатием. При этом время меняется в конечном интервале $0 < t < 2\sqrt{M}/\epsilon$. Для открытой вселенной, $\epsilon < 0$ или $\epsilon = 0$, после большого взрыва происходит бесконечное расширение вселенной.

Как уже упоминалось, качественное поведение решений можно понять, анализируя “потенциал” (35.109). При нулевой космологической постоянной, $L = 0$, для пыли, $\gamma = 1$, и излучения, $\gamma = 4/3$, он изображен на рис. 35.4. Если “энергия” $-\epsilon$ отрицательна, т.е. $\epsilon > 0$, масштабный множитель меняется в конечных интервалах $0 < a < a_{\max}$. Максимальное значение масштабного множителя определяется из условия

$$\dot{a} = 0 \quad \Leftrightarrow \quad V(a) = \epsilon.$$

Для пыли и излучения получаем следующие значения

$$a_{\max} = \begin{cases} M/\epsilon, & \gamma = 1, \\ \sqrt{M/\epsilon}, & \gamma = 4/3. \end{cases}$$

Для открытой вселенной $\epsilon \leq 0$, и масштабный множитель меняется в бесконечном полуинтервале $0 < a < \infty$, что соответствует бесконечному расширению.

Для уравнения Фридмана (35.108) аналитические решения известны 1) при $\epsilon = 0$ для всех γ ; 2) при $L = 0$ для $\gamma = 1$ и $\gamma = 4/3$ и 3) при $\gamma = 4/3$ для всех L и ϵ . Детальный анализ приведен в [285], глава 5. Некоторые решения не имеют космологических особенностей, но они не приемлемы с других точек зрения.

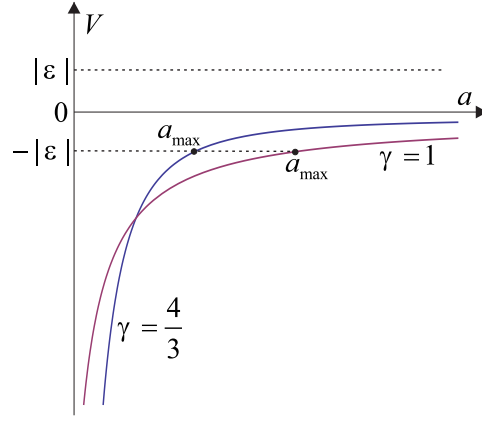


Рис. 35.4: Качественное поведение “потенциала” для пыли, $\gamma = 1$, и излучения, $\gamma = 3/4$.

В заключение проанализируем более детально космологическую особенность для пыли и излучения. Уравнение Фридмана для пыли в общем случае имеет вид

$$\dot{a}^2 = \frac{M}{a} + La^2 - \epsilon. \quad (35.116)$$

Космологическая особенность возникает в решениях этого уравнения при $a \rightarrow 0$. Вблизи особенности уравнение принимает вид

$$\dot{a}^2 = \frac{M}{a},$$

независимо от значений космологической постоянной Λ и ϵ . Поэтому асимптотика масштабного множителя вблизи космологической особенности следующая

$$a \approx \left(\frac{9M}{4} \right)^{\frac{1}{3}} t^{\frac{2}{3}}, \quad (35.117)$$

где мы положили постоянную интегрирования равной нулю. Подстановка полученной асимптотики в выражение для скалярной кривизны (35.78) приводит к следующей асимптотике при $t \rightarrow 0$

$$R \approx \frac{2}{t^{\frac{4}{3}}} \left(\frac{2}{3} t^{-\frac{2}{3}} + K_0 \right) \approx \frac{4}{3t^2}.$$

Мы видим, что при конечном значении наблюдаемого времени $t = 0$ масштабный множитель обращается в нуль, метрика вырождается, и кривизна обращается в бесконечность. Отметим, что асимптотика скалярной кривизны не зависит от постоянных M , L и ϵ , входящих в уравнение Фридмана.

Для излучения уравнение Фридмана вблизи особенности имеет вид

$$\dot{a}^2 = \frac{M}{a^2},$$

также независимо от значения космологической постоянной Λ и кривизны пространственных сечений ϵ . Его решение имеет вид

$$a \approx M^{\frac{1}{4}} \sqrt{2t}. \quad (35.118)$$

Для скалярной кривизны (35.78) получаем следующую асимптотику

$$R \approx \frac{\overset{\circ}{K}}{\sqrt{M}t}.$$

Мы видим, что для излучения возникает сингулярность скалярной кривизны только при отличной от нуля кривизне пространственных сечений. Тем не менее кривизна имеет особенность и при $\overset{\circ}{K} = 0$. Для этого достаточно вычислить квадрат тензора Риччи, который является геометрическим инвариантом,

$$R^{\alpha\beta} R_{\alpha\beta} = \frac{9\ddot{a}^2}{a^2} + \frac{3}{a^4} \left(\ddot{a}a + 2\dot{a}^2 + \frac{2}{3}\overset{\circ}{K} \right)^2,$$

где мы использовали явный вид компонент тензора Риччи (35.78). Подстановка в это выражение асимптотики масштабного множителя (35.118) приводит к следующему выражению

$$R^{\alpha\beta} R_{\alpha\beta} \approx \frac{3}{4t^4}.$$

Следовательно, квадрат тензора Риччи имеет особенность при $t \rightarrow 0$, которая не зависит от кривизны пространственных сечений. Интересно отметить, что асимптотика квадрата тензора Риччи не зависит также от M и L .

Таким образом, кривизна имеет неустранимую особенность при $a(0) = 0$, и, следовательно, соответствующие космологические решения непродолжаемы. Поскольку координатные линии времени являются геодезическими (экстремалами) и время t является каноническим параметром, то построенные космологические решения геодезически неполны.

35.4 Вакуумные решения

Вакуумные решения уравнений Фридмана, хотя, возможно, и менее реалистичны, но представляют значительный интерес. В настоящем разделе мы покажем, что все вакуумные космологические решения для однородной и изотропной вселенной – это плоское пространство Минковского $\mathbb{R}^{1,3}$ ($\Lambda = 0$), пространство де Ситтера dS ($\Lambda > 0$) и анти-де Ситтера AdS ($\Lambda < 0$).

35.4.1 Пространство-время Минковского

Докажем, что плоское пространство-время Минковского является единственным решением вакуумных уравнений Фридмана для однородной и изотропной вселенной при нулевой космологической постоянной.

Положим $\Lambda = 0$ и предположим, что материя отсутствует, $\mathcal{E} = \mathcal{P} = 0$. Тогда уравнение (35.116) принимает вид

$$\dot{a}^2 = -\epsilon.$$

Оно имеет решение только при $\epsilon \leq 0$. Следовательно, в отсутствие материи (вакуумное решение) и космологической постоянной пространственные сечения должны иметь неположительную кривизну, т.е. представляют собой либо трехмерное гиперболическое пространство $\mathbb{H}^3$, либо трехмерное евклидово пространство $\mathbb{R}^3$. Это соответствует открытой модели вселенной.

Если $\epsilon = 0$, то $a = \text{const}$ и метрика Фридмана переходит в метрику Лоренца. Поэтому пространство-время становится плоским пространством Минковского $\mathbb{R}^{1,3}$.

При $\epsilon < 0$ масштабный множитель линеен по времени:

$$a = \sqrt{-\epsilon} t, \quad (35.119)$$

где мы отбросили несущественную постоянную интегрирования и выбрали знак плюс у квадратного корня. Это вакуумное решение имеет особенность, поскольку при $t = 0$ масштабный множитель обращается в нуль, и метрика вырождается. Если вычислить компоненты тензора кривизны (35.72), то они окажутся тождественно равными нулю. Поэтому линейный масштабный множитель (35.119) в метрике Фридмана описывает плоское пространство Минковского $\mathbb{R}^{1,3}$.

Полученный вывод можно подтвердить явным преобразованием координат. Интервал для вакуумного решения (35.119) после растяжки временной координаты можно записать в виде

$$ds^2 = dt^2 - t^2 [d\chi^2 + \text{sh}^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2)],$$

где мы выбрали сферические координаты (35.74) на пространственных сечениях $t = \text{const}$. Если теперь совершить преобразование координат $t, \chi \mapsto \tau, r$, где

$$\tau := t \text{ch } \chi, \quad r := t \text{sh } \chi,$$

то интервал станет евклидовым,

$$ds^2 = d\tau^2 - dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2).$$

Якобиан преобразования координат $t, \chi \mapsto \tau, r$ равен $J = t$ и вырожден при $t = 0$. Если масштабный множитель $a(t)$ определен при положительных временах $t > 0$, то временная координата τ в пространстве Минковского также положительна $\tau > 0$.

Таким образом, вакуумное решение при $\epsilon < 0$, сводится к плоскому пространству Минковского $\mathbb{R}^{1,3}$, которое никаких особенностей не имеет. Отсюда вытекает, что особенность в решении (35.119) при $t = 0$ является координатной.

То, что мы получили одно и то же пространство Минковского $\mathbb{R}^{1,3}$ и при $\epsilon = 0$, и при $\epsilon < 0$ связано с различным выбором пространственных сечений в пространстве Минковского.

Полученное космологическое решение для вакуума следовало ожидать. Действительно, метрика (35.75) является блочно диагональной и сферически симметричной. Это очевидно в безразмерных сферических координатах. Следовательно, должна быть выполнена теорема Бирхгоффа, утверждающая, что единственным сферически симметричным решением вакуумных уравнений Эйнштейна является решение Шварцшильда, которое при нулевой массе дает пространство Минковского. При этом решение Шварцшильда с отличной от нуля массой возникнуть не может, т.к. оно не является однородным.

35.4.2 Пространство-время де Ситтера dS

Одну из первых моделей вселенных предложил де Ситтер [99, 100]. Эта модель описывает пустую вселенную и является максимально симметричной. Она представляет самостоятельный интерес, хотя и не является реалистичной с современной точки зрения.

Пусть $\Lambda > 0 \Leftrightarrow L > 0$ и материя отсутствует, $\mathcal{E} = \mathcal{P} = 0$. Тогда уравнения Фридмана (35.103) и (35.108) принимают вид

$$\ddot{a} = La, \quad (35.120)$$

$$\dot{a}^2 = La^2 - \epsilon. \quad (35.121)$$

Конечно, первое уравнение (35.120) есть следствие второго уравнения (35.121). Однако в данном случае проще их совместное рассмотрение. Первое уравнение имеет решение, зависящее от двух постоянных интегрирования:

$$a = c_+ e^{t/R} + c_- e^{-t/R}, \quad c_{\pm} = \text{const}, \quad (35.122)$$

где введено обозначение

$$R := \frac{1}{\sqrt{L}}.$$

Подстановка этого решения в уравнение (35.121) приводит к связи между постоянными интегрирования

$$c_+ c_- = \frac{\epsilon}{4L} = \frac{R^2 \epsilon}{4}. \quad (35.123)$$

Таким образом, решение уравнений Фридмана зависит только от одной постоянной интегрирования. Эту постоянную интегрирования можно фиксировать, используя инвариантность уравнения (35.121) относительно сдвига времени $t \mapsto t + \text{const}$.

При растяжке пространственных координат постоянная ϵ умножается на положительный множитель. Поэтому, не ограничивая общности, рассмотрим последовательно три случая $\epsilon = 1$, $\epsilon = 0$ и $\epsilon = -1$.

Сфера $\mathbb{S}^3$

Если пространственные сечения являются сферами, $\epsilon = 1$, положим

$$c_+ = c_- = \frac{R}{2}.$$

Тогда масштабный множитель будет равен

$$a = R \operatorname{ch}(t/R),$$

и метрика примет вид

$$ds^2 = dt^2 - R^2 \operatorname{ch}^2(t/R) \left[\frac{dr^2}{1-r^2} + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \right], \quad (35.124)$$

где мы использовали метрику на пространственных сечениях в виде (35.34). Эта метрика называется *метрикой де Ситтера* и определена при

$$-\infty < t < \infty, \quad 0 < r < 1, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi. \quad (35.125)$$

Как видим, она определена при всех временах и никакой космологической особенности не имеет.

Из формулы (35.124) не видно, что метрика де Ситтера описывает пространство-время постоянной кривизны и инвариантна относительно группы Лоренца $\mathbb{SO}(1, 4)$.

Тем не менее это так. Для доказательства построим пространство постоянной кривизны и укажем необходимое преобразование координат.

Рассмотрим пятимерное пространство Минковского $\mathbb{R}^{1,4}$ с декартовой системой координат v, w, x, y, z . По определению, метрика пространства-времени имеет вид

$$ds^2 := dv^2 - dw^2 - dx^2 - dy^2 - dz^2. \quad (35.126)$$

Пусть в него вложен однополостный гиперboloид $dS \hookrightarrow \mathbb{R}^{1,4}$:

$$v^2 - w^2 - x^2 - y^2 - z^2 = -R^2, \quad R > 0. \quad (35.127)$$

Будем считать, что топология, дифференцируемая структура и метрика на гиперboloиде индуцированы вложением. Тогда уравнение (35.127) определяет связное четырехмерное псевдориманово многообразие, вложенное в пространство Минковского $\mathbb{R}^{1,4}$ (см. рис. 35.5, на котором показаны сечения гиперboloида плоскостями $y = z = 0$). В выбранной системе координат пространство-время де Ситтера является топологическим произведением гиперboloида, изображенного на рисунке, на двумерную сферу, соответствующую координатам y, z . Это и есть *пространство де Ситтера* dS .

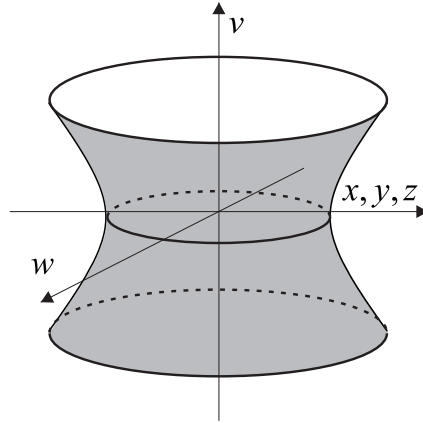


Рис. 35.5: Двумерные сечения пространства-времени де Ситтера dS .

И метрика (35.126), и уравнение гиперboloида (35.127) инвариантны относительно группы Лоренца $S\mathbb{O}(1, 4)$. Поэтому она является группой изометрий построенного гиперboloида. Для четырехмерного многообразия эта группа максимальна, и, следовательно, однополостный гиперboloид (35.127) является пространством постоянной кривизны (теорема 15.3.2).

Каждое сечение гиперboloида $v = \text{const}$ определяет трехмерную сферу $S^3 \hookrightarrow \mathbb{R}^4$, определенную уравнением

$$w^2 + x^2 + y^2 + z^2 = R^2 + v^2.$$

При $v = 0$ радиус сферы минимален и равен R . Условимся считать, что значения $0 < w < \sqrt{R^2 + v^2}$ покрывают верхнюю (северную) полусферу, а $-\sqrt{R^2 + v^2} < w < 0$ – нижнюю (южную).

Теперь совершим преобразование координат $v, w, x, y, z \mapsto R, r, t, \theta, \varphi$:

$$\begin{aligned} v &:= R \operatorname{sh}(t/R), \\ w &:= \pm R \sqrt{1 - r^2} \operatorname{ch}(t/R), \\ x &:= Rr \operatorname{ch}(t/R) \sin \theta \cos \varphi, \\ y &:= Rr \operatorname{ch}(t/R) \sin \theta \sin \varphi, \\ z &:= Rr \operatorname{ch}(t/R) \cos \theta. \end{aligned} \quad (35.128)$$

Знак $\pm$ в преобразовании координаты w соответствует выбору либо верхней полусферы $\mathbb{S}^3$, либо нижней. Тогда уравнение гиперboloида (35.127) сведется к тождеству $R^2 = R^2$. Таким образом, в новой системе координат гиперboloид задается равенством $R = \text{const}$, и координаты t, r, θ, φ можно выбрать в качестве координат на гиперboloиде. Для получения явного вида метрики, индуцированной на гиперboloиде, необходимо просто выразить дифференциалы dv, dw, dx, dy, dz через $dt, dr, d\theta, d\varphi$, используя формулы (35.128), и подставить их в исходную метрику (35.126). Для упрощения вычислений, вспомним, что для сферической системы координат в подпространстве $\mathbb{R}^3 \subset \mathbb{R}^{1,4}$, натянутом на координатные оси x, y, z , справедлива формула

$$dx^2 + dy^2 + dz^2 = [d(Rr \operatorname{ch}(t/R))]^2 + R^2 r^2 \operatorname{ch}^2(t/R)(d\theta^2 + \sin^2 \theta d\varphi^2).$$

С учетом этого равенства после несложных вычислений получим в точности метрику де Ситтера (35.124).

Мы видим, что область значений координат (35.125) при $0 < t < \infty$ покрывают либо верхнюю полусферу (знак $+$ в уравнениях (35.128)), либо нижнюю (знак $-$). Если зафиксировать знак в преобразовании координат (35.128) и разрешить времени меняться на всей вещественной прямой $-\infty < t < \infty$, то метрика де Ситтера (35.124) покроем дважды либо верхнюю, либо нижнюю полусферы $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$.

Пространственные сечения метрики де Ситтера (35.124) задаются сечениями $t = \text{const}$. В исходном пространстве Минковского $\mathbb{R}^{1,4}$ время параметризует гиперплоскости $\mathbb{R}^4$, которые задаются уравнением

$$v = R \operatorname{sh}(t/R) = \text{const}.$$

Это значит, что пространственные сечения метрики де Ситтера соответствуют сечениям гиперboloида гиперплоскостями $\mathbb{R}^4$, которые определяются условием $v = \text{const}$.

Для решения де Ситтера в виде (35.124) параметр Хаббла равен

$$H = \frac{\dot{a}}{a} = \frac{\operatorname{th}(t/R)}{R}$$

и не является постоянным.

Евклидово пространство $\mathbb{R}^3$

Зафиксируем $\epsilon = 0$ и положим $c_- = 0$ и $c_+ = R$. Тогда масштабный множитель равен

$$a = R e^{t/R}. \quad (35.129)$$

Метрика де Ситтера теперь примет вид

$$ds^2 = dt^2 - R^2 e^{2t/R} [dr^2 + r^2(d\theta^2 + \sin^2 \theta d\varphi^2)]. \quad (35.130)$$

Она соответствует слоению однополостного гиперboloида евклидовыми гиперплоскостями $\mathbb{R}^3$. Полученная метрика определена при

$$-\infty < t < \infty, \quad 0 < r < \infty, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi. \quad (35.131)$$

Для того, чтобы доказать, что метрика (35.130) действительно является метрикой де Ситтера, вернемся к вложению $d\mathbb{S} \hookrightarrow \mathbb{R}^{1,4}$, рассмотренному в предыдущем разделе. Совершим преобразование координат $v, w, x, y, z \mapsto R, r, t, \theta, \varphi$, которое задано следующими формулами:

$$\begin{aligned} v &:= \pm R \left[\frac{1}{2} r^2 e^{t/R} + \operatorname{sh}(t/R) \right], \\ w &:= \pm R \left[\frac{1}{2} r^2 e^{t/R} - \operatorname{ch}(t/R) \right], \\ x &:= Rr e^{t/R} \sin \theta \cos \varphi, \\ y &:= Rr e^{t/R} \sin \theta \sin \varphi, \\ z &:= Rr e^{t/R} \cos \theta. \end{aligned} \quad (35.132)$$

Нетрудно проверить, что уравнение (35.127), задающее вложение гиперboloида, сводится к тождеству $R^2 = R^2$. Это означает, что в новых координатах уравнение гиперboloида задается уравнением $R = \text{const}$. Чтобы получить метрику, индуцированную на гиперboloиде, в координатах t, r, θ, φ , необходимо подставить дифференциалы dv, dw, dx, dy и dz из уравнений (35.132) в исходную метрику Лоренца (35.126). Чтобы упростить вычисления, заметим что

$$dx^2 + dy^2 + dz^2 = [d(Rr e^{t/R})]^2 + R^2 r^2 e^{2t/R} (d\theta^2 + \sin^2 \theta d\varphi^2).$$

Теперь нетрудно проверить, что индуцированная метрика равна метрике де Ситтера в форме (35.130).

Пространственные сечения $\mathbb{R}^3$ для метрики (35.130) задаются сечениями $t = \text{const}$. Эти сечения в объемлющем пространстве $\mathbb{R}^{1,4}$ задают нулевые гиперплоскости

$$v - w = \pm R e^{t/R}.$$

Таким образом, пространственные сечения в метрике де Ситтера (35.130) получаются при сечении гиперboloида нулевыми гиперплоскостями.

Нетрудно вычислить параметр Хаббла для решения де Ситтера в виде (35.130)

$$\frac{\dot{a}}{a} = \frac{1}{R}.$$

Как видим, при таком выборе координат параметр Хаббла является постоянным. Тем самым пространство-время де Ситтера дает один из примеров инфляционного развития вселенной.

Гиперboloид $\mathbb{H}^3$

При $\epsilon = -1$ пространственными сечениями пространства-времени являются двуполостные гиперboloиды $\mathbb{H}^3$, описанные в разделе 35.2.2. В этом случае выберем

$$c_+ = -c_- = \frac{R}{2}.$$

Тогда масштабный множитель равен

$$a = R \operatorname{sh}(t/R),$$

и метрика де Ситтера принимает вид

$$ds^2 = dt^2 - R^2 \operatorname{sh}^2(t/R) \left[\frac{dr^2}{1+r^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2) \right]. \quad (35.133)$$

Эта метрика определена при следующих значениях координат

$$-\infty < t < \infty, \quad 0 < r < \infty, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi, \quad (35.134)$$

и вырождена при $t = 0$.

Для того, чтобы доказать, что это действительно метрика де Ситтера, снова вернемся к вложению $dS \hookrightarrow \mathbb{R}^{1,4}$. Совершим преобразование координат $w, v, x, y, z \mapsto R, r, t, \theta, \varphi$:

$$\begin{aligned} v &:= R\sqrt{1+r^2} \operatorname{sh}(t/R), \\ w &:= R \operatorname{ch}(t/R), \\ x &:= Rr \operatorname{sh}(t/R) \sin\theta \cos\varphi, \\ y &:= Rr \operatorname{sh}(t/R) \sin\theta \sin\varphi, \\ z &:= Rr \operatorname{sh}(t/R) \cos\varphi. \end{aligned} \quad (35.135)$$

Если подставить эти выражения в определяющее уравнение гиперboloида (35.127), то получим тождество $R^2 = R^2$. Следовательно, в исходном пространстве Минковского $\mathbb{R}^{1,4}$ гиперboloид задается равенством $R = \text{const}$.

Для получения явного вида метрики, индуцированной на гиперboloиде, необходимо подставить дифференциалы dv, dw, dx, dy и dz из (35.135) в исходную метрику (35.126). Учтем равенство

$$dx^2 + dy^2 + dz^2 = [d(Rr \operatorname{sh}(t/R))]^2 + R^2 r^2 \operatorname{sh}^2(t/R) (d\theta^2 + \sin^2\theta d\varphi^2).$$

Теперь нетрудно проверить, что в новых координатах метрика действительно принимает вид (35.133).

Пространственные сечения метрики де Ситтера (35.133) задаются равенством $t = \text{const}$. В исходном пространстве Минковского $\mathbb{R}^{1,4}$ время t параметризует гиперплоскости $\mathbb{R}^{1,3}$, которые задаются уравнением

$$w = R \operatorname{ch}(t/R) = \text{const}.$$

Это значит, что пространственные сечения метрики де Ситтера задаются сечениями гиперboloида гиперплоскостями $w = \text{const}$.

Параметр Хаббла в данной системе координат равен

$$H = \frac{\operatorname{cth}(t/R)}{R}$$

и вовсе не является постоянным.

Координаты Шварцшильда

Снова рассмотрим вложение пространства де Ситтера $dS \hookrightarrow \mathbb{R}^{1,4}$. Еще одна важная система координат $R, \rho, \tau, \theta, \varphi$ для пространства де Ситтера задается следующими уравнениями:

$$\begin{aligned} v &:= R\sqrt{1-\rho^2} \operatorname{sh} \tau, \\ w &:= R\sqrt{1-\rho^2} \operatorname{ch} \tau, \\ x &:= R\rho \sin \theta \cos \varphi, \\ y &:= R\rho \sin \theta \sin \varphi, \\ z &:= R\rho \cos \theta. \end{aligned} \quad (35.136)$$

Как и ранее, гиперboloид (35.127) задается уравнением $R = \text{const}$. Индуцированная метрика на гиперboloиде в координатах Шварцшильда $\tau, \rho, \theta, \varphi$ принимает вид

$$ds^2 = R^2 \left[(1-\rho^2)d\tau^2 - \frac{d\rho^2}{1-\rho^2} - \rho^2(d\theta^2 + \sin^2\theta d\varphi^2) \right]. \quad (35.137)$$

Данная метрика определена при

$$-\infty < \tau < \infty, \quad 0 < \rho < 1, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi. \quad (35.138)$$

В этих координатах метрика де Ситтера статична.

Таким образом, при положительной космологической постоянной Λ и в отсутствие полей материи, для всех значений кривизны пространственных сечений: $\epsilon = 1$, $\epsilon = 0$ и $\epsilon = -1$ мы получили пространство-время де Ситтера. Это космологическое решение не имеет никаких особенностей и после максимального продолжения вдоль экстремалей (геодезических) описывает пространство постоянной кривизны (симметрическое пространство). В целом пространство-время де Ситтера представляет собой фактор пространство

$$dS \approx \frac{\mathbb{SO}_0(1, 4)}{\mathbb{SO}_0(1, 3)},$$

что показано в главе 18. Его группа изометрий $\mathbb{SO}(1, 4)$ максимальна. Мы видим, что метрика пространства де Ситтера может быть записана в форме метрики Фридмана со всеми возможными пространственными сечениями: $\mathbb{S}^3$, $\mathbb{R}^3$ или $\mathbb{H}^3$. Кроме этого, постоянная Хаббла в различных системах координат разная. Пример метрики де Ситтера показывает насколько сильно космологические выводы зависят от выбора системы координат. Даже ответ на вопрос о том, является ли вселенная замкнутой или открытой, может зависеть от выбора системы координат. Это является существенным недостатком всех космологических моделей.

35.4.3 Пространство-время анти-де Ситтера AdS

Рассмотрим случай отрицательной космологической постоянной, $\Lambda < 0$, в пустом пространстве-времени, $\mathcal{E} = \mathcal{P} = 0$. Уравнения Фридмана по-прежнему имеют вид (35.120), (35.121), где $L = \Lambda/3 < 0$. Из второго уравнения следует, что решения существуют только при $\epsilon = -1$, т.е. пространственными сечениями метрики Фридмана являются двуполостные гиперboloиды $\mathbb{H}^3$. При отрицательной космологической постоянной решение (35.122) является вещественным только если $c_- = c_+^*$. Положим теперь

$$R := \frac{1}{\sqrt{-L}}$$

и выберем $c_+ = 1/2$. Тогда масштабный множитель (35.122) примет вид

$$a = R \cos(t/R) \quad (35.139)$$

и метрика вселенной равна

$$ds^2 = dt^2 - R^2 \cos^2(t/R) \left[\frac{dr^2}{1+r^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2) \right], \quad (35.140)$$

где мы использовали метрику двуполого пространственного гиперboloида $\mathbb{H}^3$ в форме (35.57). Эта метрика называется *метрикой анти-де Ситтера* и определена при

$$-\pi R/2 < t < \pi R/2, \quad 0 < r < \infty, \quad 0 < \theta < \pi, \quad 0 < \varphi < \pi. \quad (35.141)$$

При $t = \pm\pi R/2$ метрика вырождается.

Опять, из вида метрики анти-де Ситтера (35.140) совершенно не видно, что эта метрика описывает пространство-время постоянной кривизны. Чтобы доказать это, совершим следующее построение.

Рассмотрим плоское пятимерное пространство $\mathbb{R}^{2,3}$ с декартовыми координатами v, w, x, y, z . По определению, его метрика имеет вид

$$ds^2 := dv^2 + dw^2 - dx^2 - dy^2 - dz^2. \quad (35.142)$$

Рассмотрим однополостный вложенный гиперboloид $\text{AdS} \hookrightarrow \mathbb{R}^{2,3}$, который определен уравнением

$$v^2 + w^2 - x^2 - y^2 - z^2 = R^2. \quad (35.143)$$

На рис. 35.6 изображены двумерные сечения пространства анти-де Ситтера, соответствующие $y = z = 0$. В выбранной системе координат пространство-время анти-де Ситтера является топологическим произведением гиперboloида, изображенного на рисунке, на двумерную сферу, соответствующую координатам y, z .

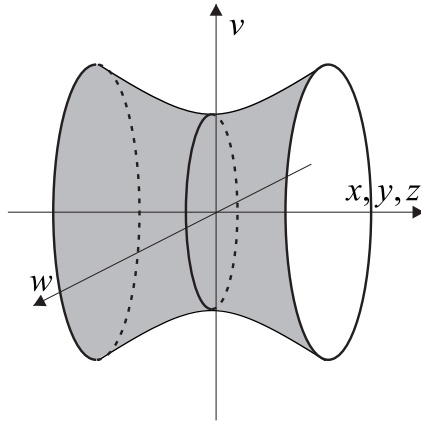


Рис. 35.6: Двумерные сечения пространства-времени анти-де Ситтера AdS .

Поскольку уравнения (35.142) и (35.143) инвариантны относительно действия группы $\text{SO}(2,3)$, то эта группа является группой изометрий пространства анти-де Ситтера. Для четырехмерного пространства-времени она максимальна, и ее алгебра Ли содержит 10 независимых векторных полей Киллинга. Согласно теореме 15.3.2 пространство-время анти-де Ситтера является однородным и изотропным и представляет собой пространство постоянной кривизны. Группа $\text{SO}(2,3)$ называется *группой анти-де Ситтера*.

Чтобы доказать, что метрика (35.140) описывает пространство постоянной кривизны, совершим преобразование координат $v, w, x, y, z \mapsto R, r, t, \theta, \varphi$, определяемое формулами:

$$\begin{aligned} v &:= R \sin(t/R), \\ w &:= \pm R \sqrt{1+r^2} \cos(t/R), \\ x &:= Rr \cos(t/R) \sin \theta \cos \varphi, \\ y &:= Rr \cos(t/R) \sin \theta \sin \varphi, \\ z &:= Rr \cos(t/R) \cos \theta. \end{aligned} \quad (35.144)$$

В новой системе координат гиперboloид задается равенством $R = \text{const}$.

Для получения индуцированной метрики используем равенство

$$dx^2 + dy^2 + dz^2 = [d(Rr \cos(t/R))]^2 + R^2 r^2 \cos^2(t/R) (d\theta^2 + \sin^2 \theta d\varphi^2).$$

Прямые вычисления показывают, что индуцированная метрика принимает вид (35.140), что и следовало доказать.

Пространственные сечения метрики анти-де Ситтера соответствуют постоянному времени $t = \text{const}$. В исходном пространстве $\mathbb{R}^{2,3}$ время t параметризует гиперплоскости $\mathbb{R}^{1,3}$, которые задаются уравнением

$$v = R \sin(t/R) = \text{const}.$$

Это значит, что пространственные сечения метрики анти-де Ситтера (35.140) задаются сечениями гиперboloида гиперплоскостями $v = \text{const}$.

Параметр Хаббла для метрики анти-де Ситтера в рассматриваемой системе координат равен

$$H = -\frac{\text{tg}(t/R)}{R}$$

и не является постоянным.

Таким образом, метрика анти-де Ситтера (35.140) представляет собой метрику пространства постоянной кривизны. Однако однополостный гиперboloид (35.143) не является пространством-временем анти-де Ситтера. Дело в том, что сечения гиперboloида двумерными плоскостями, определяемыми постоянными значениями пространственных координат $x, y, z = \text{const}$, представляют собой окружности

$$v^2 + w^2 = R^2 + x^2 + y^2 + z^2.$$

Эти окружности являются замкнутыми времениподобными кривыми и приводят к нарушению причинности (машина времени). Чтобы исправить ситуацию необходимо от однополостного гиперboloида (35.143) перейти к его универсальной накрывающей. Для этого гиперboloид нужно “развернуть” вдоль времениподобной координаты v . Это достигается путем перехода к новым координатам $t, r \mapsto \tau, \chi$, оставляя координаты θ, φ прежними. Зададим преобразование координат неявными формулами:

$$v := R \frac{\sin \tau}{\cos \chi}, \quad w := R \frac{\cos \tau}{\sin \chi},$$

где

$$0 < \tau < 2\pi, \quad -\frac{\pi}{2} < \chi < \frac{\pi}{2}.$$

Тогда

$$v^2 + w^2 = \frac{R^2}{\cos^2 \chi},$$

и из уравнения гиперболоида следует равенство

$$r \cos(t/R) = \operatorname{tg} \chi.$$

В новых координатах

$$dx^2 + dy^2 + dz^2 = R^2 \frac{d\chi^2}{\cos^4 \chi} + R^2 \operatorname{tg}^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2).$$

Теперь нетрудно вычислить метрику пространства-времени анти-де Ситтера

$$ds^2 = \frac{R^2}{\cos^2 \chi} [d\tau^2 - d\chi^2 - \sin^2 \chi (d\theta^2 + \sin^2 \theta d\varphi^2)]. \quad (35.145)$$

Мы видим, что координата τ является временем. Эта метрика статична и вейлевски (конформно) эквивалентна метрике трехмерной сферы $\mathbb{S}^3$. Важно, что она определена при всех $-\infty < \tau < \infty$. При этом гиперболоид возникает после отождествления $\tau \sim \tau + 2\pi$. Под пространством анти-де Ситтера AdS понимается именно универсальная накрывающая однополостного гиперболоида с метрикой (35.145). В этом пространстве-времени замкнутые времениподобные кривые отсутствуют.

Координаты Шварцшильда

Рассмотрим вложение $\operatorname{AdS} \hookrightarrow \mathbb{R}^{2,3}$ из предыдущего раздела. Метрику анти-де Ситтера можно записать в координатах Шварцшильда. Для этого в объемлющем пространстве введем координаты

$$\begin{aligned} v &:= R\sqrt{1 + \rho^2} \sin \tau, \\ w &:= R\sqrt{1 + \rho^2} \cos \tau, \\ x &:= R\rho \sin \theta \cos \varphi, \\ y &:= R\rho \sin \theta \sin \varphi, \\ z &:= R\rho \sin \theta \sin \varphi. \end{aligned} \quad (35.146)$$

Нетрудно проверить, что гиперболоид задается равенством $R = \text{const}$. Метрика, индуцированная на гиперболоиде, в координатах Шварцшильда $\tau, \rho, \theta, \varphi$ принимает вид

$$ds^2 = R^2 \left[(1 + \rho^2) d\tau^2 - \frac{d\rho^2}{1 + \rho^2} - \rho^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \right]. \quad (35.147)$$

Эта метрика определена при

$$-\infty < \tau < \infty, \quad 0 < \rho < \infty, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi. \quad (35.148)$$

В координатах Шварцшильда метрика де Ситтера статична.

Таким образом, при отрицательной космологической постоянной Λ и в отсутствие полей материи подходит только одно значение $\epsilon = -1$, и мы получили пространство-время анти-де Ситтера. Это космологическое решение не имеет никаких особенностей и после максимального продолжения вдоль экстремалей (геодезических) описывает пространство постоянной кривизны (симметрическое пространство). В целом пространство-время анти-де Ситтера представляет собой фактор пространство

$$\operatorname{AdS} \approx \frac{\operatorname{SO}_0(2, 3)}{\operatorname{SO}_0(1, 3)},$$

что показано в главе 18. Его группа изометрий $\mathbb{SO}(2, 3)$ максимальна.

35.4.4 Вложения вселенной Фридмана

Для того, чтобы лучше представить себе, как устроена вселенная Фридмана, полезно вложить однородное и изотропное пространство-время $\mathbb{M} = \mathbb{R} \times \mathbb{S}$ в плоское (псевдо)евклидово пространство. Оказывается, что указанное вложение можно осуществить в плоское пятимерное пространство [286].

Сфера $\mathbb{S}^3$

Замкнутые вселенные Фридмана, когда пространственными сечениями пространства-времени являются трехмерные сферы $\mathbb{S}^3$, можно вложить глобально в плоское пятимерное пространство Минковского.

Рассмотрим пятимерное пространство Минковского $\mathbb{R}^{1,4}$ с декартовой системой координат $Z^1, \dots, Z^5$. По определению, метрика такого пространства имеет вид

$$ds^2 := (dZ^1)^2 - (dZ^2)^2 - (dZ^3)^2 - (dZ^4)^2 - (dZ^5)^2. \quad (35.149)$$

Рассмотрим в $\mathbb{R}^{1,4}$ гиперповерхность вращения $\mathbb{M} \hookrightarrow \mathbb{R}^{1,4}$, которая определяется следующим уравнением

$$(Z^2)^2 + (Z^3)^2 + (Z^4)^2 + (Z^5)^2 = a^2, \quad (35.150)$$

при этом $Z^1 \in \mathbb{R}$. Эта гиперповерхность представляет собой цилиндр с направляющей $Z^1 \in \mathbb{R}$, в основании которого лежит трехмерная сфера радиуса a , определяемая уравнением (35.150). Теперь положим

$$Z^1 := \int dt \sqrt{1 + \dot{a}^2}, \quad Z^2 = aw, \quad Z^3 := ax, \quad Z^4 := ay, \quad Z^5 := az, \quad (35.151)$$

где $a = a(t)$ – масштабный множитель. Тем самым мы предполагаем, что радиус сферы a зависит от первой координаты: $a = a(t(Z^1))$. Теперь уравнение гиперповерхности (35.150) примет вид

$$w^2 + x^2 + y^2 + z^2 = 1. \quad (35.152)$$

То есть в координатах w, x, y, z мы имеем трехмерную сферу единичного радиуса, обычным образом вложенную в четырехмерное евклидово пространство $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$. Из уравнения вложения (35.152) вытекает соотношение между дифференциалами:

$$wdw + xdx + ydy + zdz = 0, \quad (35.153)$$

которое нам понадобится в дальнейшем.

Найдем метрику, индуцированную на гиперповерхности (35.150). В качестве независимых координат на гиперповерхности можно выбрать, например, t, x, y, z . Однако проще рассматривать координаты t, w, x, y, z , на которые наложено условие (35.152). Тогда из определения (35.151) следуют выражения для дифференциалов:

$$\begin{aligned} dZ^1 &= \sqrt{1 + \dot{a}^2} dt, & dZ^2 &= w\dot{a}dt + adw, & dZ^3 &= x\dot{a}dt + adx, \\ dZ^4 &= y\dot{a}dt + ady, & dZ^5 &= z\dot{a}dt + adz. \end{aligned} \quad (35.154)$$

Подстановка этих выражений в метрику (35.149) приводит к индуцированной метрике

$$\begin{aligned} ds^2 &= (1 + \dot{a}^2)dt^2 - (w^2 + x^2 + y^2 + z^2)\dot{a}^2 dt^2 - \\ &\quad - 2(wdw + xdx + ydy + zdz)a\dot{a}dt - a^2(dw^2 + dx^2 + dy^2 + dz^2) = \\ &= dt^2 - a^2(dw^2 + dx^2 + dy^2 + dz^2), \end{aligned} \quad (35.155)$$

где учтены уравнения (35.152) и (35.153). Мы видим, что индуцированная метрика совпадает с метрикой Фридмана

$$ds^2 = dt^2 + a^2 \overset{\circ}{g}_{\mu\nu} dx^\mu dx^\nu,$$

где $\overset{\circ}{g}_{\mu\nu}$ — отрицательно определенная метрика на сфере (35.152) единичного радиуса.

Для вселенной Эйнштейна (см. раздел 35.3.5) масштабный множитель равен константе, $a = \text{const}$. Поэтому, как многообразие, вселенная Эйнштейна представляет собой цилиндр в пятимерном пространстве Минковского $\mathbb{R}^{1,4}$ с постоянным сечением (35.150), которым является трехмерная сфера $\mathbb{S}^3$. Поэтому часто говорят о *цилиндрическом мире Эйнштейна*.

В общем случае замкнутая вселенная Фридмана представляет собой цилиндр переменного сечения, который вложен в пространство Минковского $\mathbb{R}^{1,4}$. Это согласуется с представлением $\mathbb{M} = \mathbb{R} \times \mathbb{S}^3$.

Гиперболоид $\mathbb{H}^3$

Открытая вселенная Фридмана, пространственными сечениями которой являются гиперболоиды $\mathbb{H}^3$, вкладывается в плоское псевдоевклидово пространство $\mathbb{R}^{2,3}$. Вложение глобально и определяется аналогично тому, как это было сделано в предыдущем случае замкнутой вселенной.

Метрика плоского псевдоевклидова пространства $\mathbb{R}^{2,3}$ в декартовой системе координат имеет вид

$$ds^2 := (dZ^1)^2 + (dZ^2)^2 - (dZ^3)^2 - (dZ^4)^2 - (dZ^5)^2. \quad (35.156)$$

Рассмотрим вложение четырехмерного цилиндра в это пространство, которое задано следующим уравнением

$$(Z^2)^2 - (Z^3)^2 - (Z^4)^2 - (Z^5)^2 = a^2, \quad (35.157)$$

при этом $Z^1 \in \mathbb{R}$. В основании этого цилиндра лежат гиперболоиды (35.157) (как и ранее, мы рассматриваем верхнюю полу́ двуполостного гиперболоида, рассмотренного в разделе 35.2.2).

Теперь введем в псевдоевклидовом пространстве $\mathbb{R}^{2,3}$ координаты t, w, x, y, z с помощью соотношений

$$Z^1 := \int dt \sqrt{1 - \dot{a}^2}, \quad Z^2 = aw, \quad Z^3 := ax, \quad Z^4 := ay, \quad Z^5 := az, \quad (35.158)$$

которые отличаются только знаком в определении Z^1 от формул (35.158). Тогда гиперболоид, лежащий в основании цилиндра, будет задан уравнением

$$w^2 - x^2 - y^2 - z^2 = 1, \quad (35.159)$$

откуда следует равенство дифференциалов

$$wdw - xdx - ydy - zdz = 0. \quad (35.160)$$

Подстановка дифференциалов dZ в метрику (35.156) приводит к интервалу

$$\begin{aligned} ds^2 &= (1 - \dot{a}^2)dt^2 + (w^2 - x^2 - y^2 - z^2)\dot{a}^2 dt^2 + \\ &\quad + 2(wdw - xdx - ydy - zdz)a\dot{a}dt + a^2(dw^2 - dx^2 - dy^2 - dz^2) = \\ &= dt^2 + a^2(dw^2 - dx^2 - dy^2 - dz^2), \end{aligned} \quad (35.161)$$

где учтены равенства (35.159) и (35.160). Таким образом, мы получили метрику Фридмана для открытой вселенной.

Евклидово пространство $\mathbb{R}^3$

Открытая вселенная Фридмана, пространственными сечениями которой являются евклидовы пространства $\mathbb{R}^3$, вкладывается в пространство Минковского $\mathbb{R}^{1,4}$. Однако это вложение несколько сложнее, чем в двух предыдущих случаях.

Если в рассматриваемом случае $a = \text{const}$, то метрика Фридмана растяжкой координат приводится к метрике Лоренца, и вкладывать нечего. Поэтому рассмотрим случай $\dot{a} \neq 0$.

Метрика пространства Минковского $\mathbb{R}^{1,4}$ в декартовых координатах имеет вид (35.149). Положим

$$\begin{aligned} Z^1 &:= \frac{a}{2b}(b^2 + r^2) + f(t), \\ Z^2 &:= \frac{a}{2b}(b^2 - r^2) - f(t) = ab - Z^1, \\ Z^3 &:= ax, \quad Z^4 := ay, \quad Z^5 := az, \end{aligned} \quad (35.162)$$

где

$$b = \text{const} > 0, \quad 2bf := \int \frac{dt}{\dot{a}}, \quad r^2 := x^2 + y^2 + z^2.$$

Если рассматривать переменные t, x, y, z в качестве координат на вложенной гиперповерхности $\mathbb{M} \hookrightarrow \mathbb{R}^{1,4}$, то уравнение вложения можно записать в следующем виде

$$(Z^1)^2 - (Z^2)^2 - (Z^3)^2 - (Z^4)^2 - (Z^5)^2 = 2abf, \quad (35.163)$$

где в правой части надо подставить $f = f(t(Z^1))$.

Чтобы получить выражение для индуцированной метрики, вычислим дифференциалы

$$\begin{aligned} dZ^1 &= \frac{\dot{a}dt}{2b}(b^2 + r^2) + \frac{a}{b}(xdx + ydy + zdz) + f'dt, \\ dZ^2 &= b\dot{a}dt - dZ^1, \end{aligned}$$

при этом дифференциалы dZ^3 , dZ^4 и dZ^5 имеют тот же вид (35.154), что и раньше. Простые вычисления приводят к равенствам:

$$\begin{aligned} (dZ^1)^2 - (dZ^2)^2 &= dt^2 + \dot{a}^2 r^2 dt^2 + 2(xdx + ydy + zdz)a\dot{a}dt, \\ (dZ^3)^2 + (dZ^4)^2 + (dZ^5)^2 &= r^2 \dot{a}^2 dt^2 + 2(xdx + ydy + zdz)a\dot{a}dt + a^2(dx^2 + dy^2 + dz^2). \end{aligned}$$

Разность полученных выражений приводит к индуцированной метрике

$$ds^2 = dt^2 - a^2(dx^2 + dy^2 + dz^2).$$

Мы видим, что индуцированная метрика совпадает с метрикой Фридмана для $\overset{\circ}{K} = 0$.

Заметим, что из определения (35.162) следует неравенство

$$Z^1 + Z^2 = ab > 0.$$

Это значит, что пространство-время представлено только половиной вложенной гиперповерхности (35.163).

В заключение рассмотрим вложение пространства де Ситтера, для которого масштабный множитель имеет вид (35.129):

$$a = R e^{t/R}, \quad R = \text{const} > 0.$$

В этом случае уравнение вложения (35.163) принимает вид

$$(Z^1)^2 - (Z^2)^2 - (Z^3)^2 - (Z^4)^2 - (Z^5)^2 = -R^2,$$

что совпадает с вложением (35.127), описанным ранее.

35.4.5 Экстремали во вселенной Фридмана

Рассмотрим однородное и изотропное пространство-время $\mathbb{M} = \mathbb{R} \times \mathbb{S}$ с метрикой Фридмана (35.70). Эта метрика определяет символы Кристоффеля (35.71), которые задают уравнения экстремалей (геодезических). Оказывается, что, если масштабный множитель известен, то уравнения экстремалей во вселенной Фридмана интегрируются в квадратурах в общем виде [286].

Сначала рассмотрим экстремали на пространственных сечениях $\mathbb{S}$, по которым будут строиться экстремали во вселенной Фридмана $\mathbb{M} = \mathbb{R} \times \mathbb{S}$. На пространстве постоянной кривизны $\mathbb{S}$ уравнения экстремалей $\{x^\mu(u)\}$, $\mu = 1, 2, 3$, имеют вид

$$\frac{d^2 x^\mu}{du^2} = -\overset{\circ}{\Gamma}_{\nu\rho}^\mu \frac{dx^\nu}{du} \frac{dx^\rho}{du}, \quad (35.164)$$

где u – канонический параметр вдоль экстремали. Для определенности, выберем длину экстремали в качестве канонического параметра:

$$du^2 := -\overset{\circ}{g}_{\mu\nu} dx^\mu dx^\nu > 0. \quad (35.165)$$

Поскольку $\mathbb{S}$ – пространство постоянной кривизны, то все экстремали просто описываются. Для евклидова пространства $\mathbb{R}^3$ все прямые и только они являются экстремалими. Для сферы, вложенной в четырехмерное евклидово пространство $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$ (раздел 35.2.1), все экстремали – это сечения сферы трехмерными плоскостями, проходящими через начало координат. Аналогично, сечения гиперboloида $\mathbb{H}^3 \hookrightarrow \mathbb{R}^{3,1}$ (раздел 35.2.2) трехмерными времениподобными плоскостями, проходящими через начало координат, определяют все экстремали. Таким образом, все экстремали $\{x^\mu(u)\}$ на пространственных сечениях $\mathbb{S}$ можно считать известными.

Теперь исследуем экстремали в пространстве-времени Фридмана. Уравнения (35.85) для экстремалей $\{x^\alpha(s)\}$, $\alpha = 0, 1, 2, 3$, вселенной Фридмана, где s – канонический

параметр, уже были выписаны при рассмотрении красного смещения спектральных линий.

Покажем, что знание экстремалей на пространствах постоянной кривизны $\mathbb{S}$, позволяет описать все экстремали во вселенной Фридмана. Для этой цели запишем уравнение экстремалей на $\mathbb{M} = \mathbb{R} \times \mathbb{S}$ не в канонической, а в произвольной параметризации. Пусть u – произвольный параметр вдоль экстремали $x^\alpha(u)$ и s – длина экстремали (канонический параметр). Тогда параметры связаны равенством

$$\left(\frac{ds}{du}\right)^2 = g_{\alpha\beta} \frac{dx^\alpha}{du} \frac{dx^\beta}{du}, \quad (35.166)$$

и уравнение экстремалей принимает вид (16.23)

$$\frac{d^2 x^\alpha}{du^2} = -\Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{du} \frac{dx^\gamma}{du} + \frac{dx^\alpha}{du} \frac{d^2 s}{du^2} \frac{du}{ds}. \quad (35.167)$$

Отметим, что это уравнение инвариантно относительно линейного преобразования канонического параметра $s \mapsto ks$, $k \neq 0$. Важно также, что при получении уравнения (35.167) было использовано условие $ds/du \neq 0$.

Теперь отождествим параметр u вдоль экстремали на $\mathbb{M}$ с каноническим параметром вдоль экстремали на $\mathbb{S}$ (35.165). Тогда для метрики Фридмана справедливы равенства:

$$ds^2 = dt^2 + a^2 g_{\mu\nu} dx^\mu dx^\nu = dt^2 - a^2 du^2 \quad \Leftrightarrow \quad \left(\frac{ds}{du}\right)^2 = \left(\frac{dt}{du}\right)^2 - a^2. \quad (35.168)$$

Последнее равенство справедливо только для времениподобных экстремалей.

Уравнения для экстремали (35.167) разбиваются на временную и пространственные компоненты:

$$\frac{d^2 t}{du^2} = -a \frac{da}{dt} + \frac{dt}{du} \frac{d^2 s}{du^2} \frac{du}{ds}, \quad (35.169)$$

$$\frac{d^2 x^\mu}{du^2} = -\overset{\circ}{\Gamma}_{\nu\rho}^\mu \frac{dx^\nu}{du} \frac{dx^\rho}{du} + \frac{dx^\mu}{du} \left(\frac{d^2 s}{du^2} \frac{du}{ds} - \frac{2}{a} \frac{da}{du} \right), \quad (35.170)$$

где мы учли явный вид символов Кристоффеля (35.71).

По построению системы координат, в которой метрика имеет блочно диагональный вид (35.70), координатные линии $x^\mu = \text{const}$, перпендикулярные пространственным сечениям $\mathbb{S}$, являются экстремалими. Для них уравнение (35.170) тождественно удовлетворяется. Поскольку для этих экстремалей $du = 0$, то вместо уравнения (35.169) следует использовать первое из уравнений (35.85), которое сводится к равенству

$$\frac{d^2 t}{ds^2} = 0 \quad \Leftrightarrow \quad t = as + b, \quad a, b = \text{const}. \quad (35.171)$$

Предложение 35.4.1. *Экстремали на пространственных сечениях $\mathbb{S}$ являются экстремалими вселенной Фридмана $\mathbb{M} = \mathbb{R} \times \mathbb{S}$ тогда и только тогда, когда масштабный множитель постоянен, $a = \text{const}$.*

Доказательство. Экстремали, целиком лежащие в $\mathbb{S}$, определяются условием $t = \text{const}$. В этом случае уравнение (35.169) приводит к равенству $\dot{a} = 0$. $\square$

Теперь проанализируем времениподобные экстремали общего положения. Продифференцируем по u последнее равенство (35.168) и подставим в него вторую производную d^2t/du^2 из уравнения (35.169):

$$2 \frac{ds}{du} \frac{d^2s}{du^2} = 2 \left(\frac{dt}{du} \right)^2 \frac{d^2s}{du^2} \frac{du}{ds} - 4a \frac{da}{du}.$$

Отсюда вытекает, что

$$\frac{d^2s}{du^2} \frac{du}{ds} = \frac{2}{a} \frac{da}{du}, \quad (35.172)$$

где учтено равенство (35.168). Полученное соотношение сводит уравнение (35.170) к уравнению (35.164). Таким образом, уравнения (35.164) для времениподобных экстремалей $x^\mu(u)$ на $\mathbb{S}$ вытекают из уравнений (35.169), (35.170) для экстремалей $x^\alpha(u)$ на $\mathbb{M}$. При доказательстве мы использовали одну и ту же параметризацию экстремалей, которая является канонической только для экстремалей на $\mathbb{S}$.

В обратную сторону утверждение также верно. Действительно, вместо уравнения (35.169) в системе уравнений для экстремалей (35.169), (35.170) можно использовать уравнение (35.172), которое легко интегрируется:

$$\frac{ds}{du} = Ca^2, \quad C = \text{const} \neq 0. \quad (35.173)$$

Знак постоянной интегрирования C определяет взаимную ориентацию параметров s и u . Поэтому, не ограничивая общности, положим $C > 0$. С учетом полученного решения равенство (35.168) для времениподобных экстремалей принимает вид

$$\left(\frac{dt}{du} \right)^2 = a^2(1 + C^2a^2) \quad \Leftrightarrow \quad u = \pm \int \frac{dt}{a\sqrt{1 + C^2a^2}}. \quad (35.174)$$

Последний интеграл определяет функцию $t(u)$ в квадратурах. Знак $\pm$ определяет взаимную ориентацию оси времени t и параметра u , и является несущественным.

При желании можно получить экстремаль, как функцию канонического параметра. Из уравнений (35.173) и (35.174) следуют равенства:

$$ds = Ca^2 du = \pm \frac{Cad t}{\sqrt{1 + C^2a^2}} \quad \Leftrightarrow \quad s = \pm C \int \frac{dt a}{\sqrt{1 + C^2a^2}}. \quad (35.175)$$

Тем самым функция $t(s)$ задана в квадратурах.

Длина касательного вектора к экстремалиям равна

$$g_{\alpha\beta} \frac{dx^\alpha}{du} \frac{dx^\beta}{du} = \left(\frac{dt}{du} \right)^2 - a^2 = C^2a^4 > 0.$$

Если параметр канонический, то длина касательного вектора постоянна:

$$g_{\alpha\beta} \frac{dx^\alpha}{du} \frac{dx^\beta}{du} = \text{const}.$$

Таким образом, проекции времениподобных экстремалей общего положения на пространственные сечения являются экстремалиями на $\mathbb{S}$, а зависимость временной координаты $t(u)$ задается интегралом (35.174), где $C > 0$. В пределе $C \rightarrow \infty$ возникают экстремали, перпендикулярные $\mathbb{S}$.

Светоподобные (нулевые) экстремали определяются теми же формулами, что и времениподобные экстремали общего положения при $C = 0$.

Для пространственноподобных экстремалей уравнение (35.168) необходимо заменить на следующее равенство

$$\left(\frac{ds}{du}\right)^2 = a^2 - \left(\frac{dt}{du}\right)^2.$$

Дальнейший анализ проводится так же, как и в предыдущем случае. При этом зависимости $t(u)$ и $t(s)$ даются интегралами:

$$u = \pm \int \frac{dt}{a\sqrt{1 - C^2 a^2}}, \quad s = \pm C \int \frac{dt a}{\sqrt{1 - C^2 a^2}} \quad C > 0. \quad (35.176)$$

При этом надо следить, чтобы подкоренное выражение было неотрицательно.

Таким образом, мы нашли все экстремали вселенной Фридмана.

Теорема 35.4.1. *Экстремали вселенной Фридмана $\mathbb{M} = \mathbb{R} \times \mathbb{S}$ делятся на два класса.*

- 1) *Экстремали (35.171), перпендикулярные всем пространственным сечениям и проходящие через каждую точку $\mathbb{S}$.*
- 2) *Экстремали $\{t(u), x^\mu(u)\}$, где $x^\mu(u)$ – экстремали на $\mathbb{S}$ и функция $t(u)$ определена либо интегралом (35.174) для времениподобных ($C > 0$) и светоподобных ($C = 0$) экстремалей, либо интегралом (35.176) для пространственноподобных экстремалей. Параметр u является каноническим параметром только для экстремалей на $\mathbb{S}$.*

Пример 35.4.1. Рассмотрим вселенную де Ситтера (см. раздел 35.4.2), представленную в виде топологического произведения $\mathbb{M} = \mathbb{R} \times \mathbb{R}^3$. В этом случае экстремали на $\mathbb{S} = \mathbb{R}^3$ – это все возможные прямые:

$$x^\mu = n^\mu u + x_0^\mu,$$

где x_0^μ – координаты точки, через которую проходит прямая, и n^μ – компоненты единичного направляющего вектора.

Масштабный множитель имеет вид (35.129)

$$a = R e^{t/R}, \quad R = \text{const},$$

и соответствует инфляционной вселенной. Зависимость временной координаты t от параметра u для времениподобных экстремалей задается интегралом (35.174)

$$u = \pm \int \frac{dt}{R e^{t/R} \sqrt{1 + C^2 R^2 e^{2t/R}}} = \mp \sqrt{C^2 R^2 + e^{-2t/R}}.$$

Канонический параметр вдоль экстремалей определяется интегралом (35.176), который также легко берется:

$$s = \pm C \int \frac{dt R e^{t/R}}{\sqrt{1 + C^2 R^2 e^{2t/R}}} = \pm \left[t + R \ln \left(1 + \sqrt{1 + \frac{e^{-2t/R}}{C^2 R^2}} \right) \right]. \quad \square$$

Формулы, полученные в настоящем разделе, дают общую схему анализа экстремалей для вселенной Фридмана. При построении конкретной космологической модели сначала нужно найти масштабный множитель из уравнений Эйнштейна, а затем провести анализ экстремалей.

35.4.6 Космологические горизонты

Пространство-время в космологических моделях, построенных в рамках общей теории относительности, может иметь глобальные отличия от пространства Минковского $\mathbb{R}^{1,3}$, которые важны с физической точки зрения. Ниже мы продемонстрируем эти свойства на примере вселенной Фридмана.

Рассмотрим вселенную Фридмана, метрику которой запишем в виде

$$ds^2 = dt^2 - a^2 \left[\frac{dr^2}{1 - \epsilon r^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2) \right], \quad (35.177)$$

где $\epsilon = 1$ для замкнутой вселенной ($\mathbb{S} = \mathbb{S}^3$), $\epsilon = 0$ для открытой вселенной ($\mathbb{S} = \mathbb{R}^3$) и $\epsilon = -1$ также для открытой вселенной ($\mathbb{S} = \mathbb{H}^3$). Выше мы использовали метрику сферы в виде (35.34) и гиперboloида в виде (35.57). Координатные линии времени являются экстремальными по построению, и их длина, которая является каноническим параметром, равна разности $t_2 - t_1$, где $t_{1,2}$ – начало и конец экстремали. Если масштабный множитель определен при всех $t \in (-\infty, \infty)$, то эти экстремали полны. При наличии особенности, $a(t_b) = 0, \infty$, при конечных значениях t_b пространство-время будет геодезически неполным.

Предположим, что наблюдатель расположен в начале координат $r = 0$. Это не ограничивает общности, т.к. вселенная однородна. Нулевые радиальные экстремали, $ds^2 = 0$, для него определяются уравнением

$$dt^2 = a^2 dr^2.$$

Поэтому световой конус с вершиной в точке t_0 задается следующими соотношениями:

$$\int_{t_0}^t \frac{d\tau}{a(\tau)} = \pm r, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi, \quad (35.178)$$

где знак $\pm$ в правой части интеграла соответствует световому конусу будущего, $t > t_0$, и прошлого, $t < t_0$. В момент времени t_0 наблюдатель может видеть только те события, которые находятся в световом конусе прошлого. Допустим, что вселенная родилась в момент времени $t_b < t_0$. В этот момент масштабный множитель равен нулю, $a(t_b) = 0$, и допустим, что интеграл (35.178) при $t = t_b$ сходится:

$$\int_{t_0}^{t_b} \frac{dt}{a} = -r_b > -\infty.$$

Это значит, что наблюдатель в момент времени t_0 может наблюдать только те события, для которых $r < r_b$. Теперь допустим, что интеграл (35.178) сходится в бесконечности, $t \rightarrow \infty$:

$$\int_{t_0}^{\infty} \frac{dt}{a} = r_{\infty}.$$

В такой ситуации наблюдатель никогда не сможет увидеть те точки вселенной, для которых

$$r > r_b + r_{\infty} = \int_{t_b}^{\infty} \frac{dt}{a}.$$

Поверхность светового конуса прошлого с вершиной в точке $t_0 \rightarrow \infty$ называется *горизонтом событий* наблюдателя.

Для замкнутой вселенной на радиус наложено дополнительное условие $r < 1$. Поэтому для существования горизонта событий необходимо, чтобы $r_b + r_\infty < 1$.

Таким образом, в рассматриваемом случае наблюдатель в принципе никогда не может наблюдать события, происходящие в некоторых областях вселенной (за горизонтом событий). И эти области вселенной могут оказаться большими.

С другой стороны, наблюдатель может повлиять только на те события, которые находятся в световом конусе будущего. Максимальный конус будущего для наблюдателя, находящегося в начале системы координат, $r = 0$, соответствует значению $t \rightarrow t_b$. Поэтому наблюдатель никак не сможет повлиять на события с $r > r_b + r_\infty$. Поверхность светового конуса будущего с вершиной в точке $t \rightarrow t_b$ называется *горизонтом частицы*. Таким образом, наблюдатель никак не может влиять на события, находящиеся за горизонтом частицы.

Анализ других космологических моделей на наличие горизонтов производится аналогично в каждом конкретном случае.

Если в пространстве-времени есть космологические горизонты, то это означает, что существуют такие области вселенной, которые между собой никак причинно не связаны. Наличие горизонтов представляет собой большую космологическую проблему: каким образом во вселенной может установиться постоянная плотность материи, температура и т.д., если некоторые части вселенной никак между собой не взаимодействуют? Ответ на этот вопрос в настоящее время не ясен.

35.5 Однородные вселенные

В предыдущем разделе мы рассмотрели однородные и изотропные модели вселенной. При этом из изотропии вселенной вытекает ее однородность (предложение 15.3.1). Обратное утверждение неверно: однородных многообразий намного больше, чем изотропных. Трехмерных связных и односвязных изотропных многообразий всего три (с точностью до диффеоморфизма): это сфера $\mathbb{S}^3$ (положительная кривизна), евклидово пространство $\mathbb{R}^3$ (отрицательная кривизна) и верхняя полá двуполостного гиперboloида $\mathbb{H}^3$ (отрицательная кривизна). Классификация всех трехмерных однородных многообразий в настоящее время неизвестна. Поэтому мы упростим задачу. Будем требовать, чтобы пространственные сечения постоянного времени $t = \text{const}$ были не просто однородными, а главными однородными пространствами (см. раздел 9.1). Напомним, что главным однородным пространством называется группа преобразований $(\mathbb{M}, \mathbb{G})$, которая действует на многообразии $\mathbb{M}$ транзитивно и свободно. В этом случае размерности многообразия и группы Ли $\mathbb{G}$ равны, и многообразие $\mathbb{M}$ можно отождествить с самой группой Ли. Поэтому классификация главных однородных пространств сводится к классификации трехмерных групп Ли. Согласно теореме 8.11.4 каждой алгебре Ли соответствует единственная, с точностью до изоморфизма, связная и односвязная группа Ли (универсальная накрывающая). Все остальные связные группы Ли с той же алгеброй являются фактор группами универсальной накрывающей по некоторой нормальной подгруппе, действующей свободно и собственнo разрывно. Таким образом, если пространственными сечениями вселенной являются связные и односвязные главные однородные пространства, то их классификация сводится к классификации трехмерных алгебр Ли. Эта классификация хорошо известна.

35.5.1 Классификация Бианки

Трехмерные алгебры Ли были классифицированы Луиджи Бианки [62].

Пусть задана трехмерная алгебра Ли $\mathfrak{g}$ с базисом L_i , $i = 1, 2, 3$. Коммутатор базисных левоинвариантных векторных полей,

$$[L_i, L_j] = f_{ij}^k L_k, \quad (35.179)$$

определяется структурными константами f_{ij}^k , которые, по построению, антисимметричны по нижним индексам

$$f_{ij}^k = -f_{ji}^k. \quad (35.180)$$

Кроме того, структурные константы должны удовлетворять тождествам Якоби (см. раздел 8.1)

$$f_{ij}^l f_{kl}^m + f_{jk}^l f_{il}^m + f_{ki}^l f_{jl}^m = 0. \quad (35.181)$$

Верно также обратное утверждение. Произвольный набор структурных констант, удовлетворяющих условиям (35.180) и (35.181), однозначно определяет некоторую алгебру Ли.

Таким образом, классификация трехмерных алгебр Ли сводится к перечислению всех возможных наборов структурных констант с точностью до некоторого отношения эквивалентности. В данном случае отношение эквивалентности – это выбор базиса в алгебре Ли. Мы говорим, что два базиса эквивалентны, $L_i \sim L'_i$, если они связаны некоторым невырожденным линейным преобразованием:

$$L_i \sim L'_i \quad \Leftrightarrow \quad L'_i = S_i^j L_j, \quad S \in \mathrm{GL}(3, \mathbb{R}).$$

Из определения структурных констант (35.179) вытекает, что при таких преобразованиях структурные константы ведут себя как тензор третьего ранга с одним контравариантным и двумя ковариантными индексами. Поскольку базис алгебры Ли L_i , как базис произвольного векторного пространства, определен с точностью до произвольных невырожденных линейных преобразований, то мы используем этот произвол для приведения структурных констант к некоторому каноническому виду.

Изначально в трехмерном векторном пространстве алгебры Ли $\mathfrak{g}$ никакой метрики нет, и мы не можем поднимать и опускать индексы у структурных констант. Зато есть два полностью антисимметричных тензора третьего ранга: один с ковариантными, ε_{ijk} , а другой с контравариантными индексами, ε^{ijk} . Мы принимаем соглашение, что $\varepsilon_{123} = 1$ и $\varepsilon^{123} = 1$. Их свойства описаны в приложении 37.4.

Поскольку структурные константы антисимметричны по паре индексов, то их можно взаимно однозначно задать тензором второго ранга:

$$f_{ij}^k =: \varepsilon_{ijl} t^{lk} \quad \Leftrightarrow \quad t^{lk} = \frac{1}{2} \varepsilon^{ijl} f_{ij}^k.$$

Тензор второго ранга t^{lk} в общем случае никакой симметрией не обладает, и его можно разложить на симметричную и антисимметричную части:

$$b^{lk} := \frac{1}{2}(t^{lk} + t^{kl}), \quad d^{lk} := \frac{1}{2}(t^{lk} - t^{kl}).$$

В свою очередь, антисимметричный тензор второго ранга можно параметризовать ковектором:

$$d^{ij} =: \varepsilon^{ijk} a_k \quad \Leftrightarrow \quad a_k = \frac{1}{2} d^{ij} \varepsilon_{ijk}.$$

Таким образом, в трехмерном случае структурные константы взаимно однозначно параметризуются симметричным тензором второго ранга b^{ij} и ковектором a_i :

$$f_{ij}^k = \varepsilon_{ijl} b^{lk} + \delta_i^k a_j - \delta_j^k a_i. \quad (35.182)$$

Отметим, что при получении данного представления никакая метрика на алгебре Ли $\mathfrak{g}$ не использовалась.

Теперь подставим структурные константы (35.182) в тождества Якоби (35.181). Для упрощения вычислений заметим, что левая сторона равенства (35.181) антисимметрична по трем индексам i, j, k и поэтому эквивалентна равенству

$$\varepsilon^{ijk} f_{ij}^l f_{kl}^m = 0.$$

Подставляя сюда выражение для структурных констант (35.182) приходим к равенству

$$b^{ij} a_j = 0. \quad (35.183)$$

Таким образом, для классификации алгебр Ли необходимо перечислить все симметричные тензоры $b^{ij} = b^{ji}$ и ковекторы a_i , на которые наложено условие (35.183) и отношение эквивалентности, которое связано со свободой выбора базиса в алгебре Ли $\mathfrak{g}$. Из определения следует, что b^{ij} и a_i при замене базиса в алгебре Ли ведут себя как тензоры. Чтобы привести структурные константы к каноническому виду совершим следующее преобразование.

Уравнение (35.183) означает, что ковектор a , если он отличен от нуля, лежит в одном из главных направлений тензора b^{ij} , в том, которое отвечает нулевому собственному значению b^{ij} . Согласно теореме 37.1.8, с помощью ортогонального преобразования базиса алгебры Ли симметричную матрицу b всегда можно преобразовать к диагональному виду

$$b^{ij} = b^{(i)} \delta^{ij}, \quad \text{суммирование по } i \text{ нет,}$$

где на диагонали стоят собственные числа. Не ограничивая общности, можно считать, что ковектор a лежит в первом главном направлении, т.е. $a = (a, 0, 0)$. Тогда уравнение (35.183) примет вид $b^{(1)} a$, откуда следует, что либо $b^{(1)} = 0$, либо $a = 0$.

После таких преобразований правила коммутации базисных векторных полей алгебры Ли примут вид

$$[L_1, L_2] = b^{(3)} L_3 - a L_2, \quad [L_2, L_3] = b^{(1)} L_1, \quad [L_3, L_1] = b^{(2)} L_2 + a L_3. \quad (35.184)$$

У нас имеется также свобода в масштабировании базиса (умножению на отличные от нуля постоянные). Совершим масштабное преобразование

$$L_1 \mapsto L'_1 := k_1 L_1, \quad L_2 \mapsto L'_2 := k_2 L_2, \quad L_3 \mapsto L'_3 := k_3 L_3, \quad k_1 k_2 k_3 \neq 0.$$

Тогда для нового базиса правила коммутации (35.184) останутся прежними, но с заменой

$$b'^{(1)} = \frac{k_1}{k_2 k_3} b^{(1)}, \quad b'^{(2)} = \frac{k_2}{k_3 k_1} b^{(2)}, \quad b'^{(3)} = \frac{k_3}{k_1 k_2} b^{(3)}, \quad a' = \frac{a}{k_1}.$$

Допустим, что все собственные значения $b^{(i)}$ отличны от нуля, и положим

$$k_1 = \frac{k_2 k_3}{|b^{(1)}|}, \quad k_2 = \frac{k_1 k_3}{|b^{(2)}|}, \quad k_3 = \frac{k_1 k_2}{|b^{(3)}|}.$$

Эти уравнения имеют решения

$$k_1^2 = |b^{(2)}b^{(3)}|, \quad k_2^2 = |b^{(3)}b^{(1)}|, \quad k_3^2 = |b^{(1)}b^{(2)}|. \quad (35.185)$$

Тогда новые собственные значения будут равны по модулю единице $b'^{(i)} = \pm 1$, при этом $a = 0$, как следует из уравнения (35.183).

Заметим также, что изменение знака одной постоянной, например, k_1 при извлечении корня в (35.185) приводит к одновременному изменению всех знаков собственных значений $b^{(i)}$.

Аналогично можно доказать, что при $a = 0$ все отличные от нуля собственные значения $b^{(i)}$ можно по модулю приравнять единице. При этом еще остается свобода одновременного изменения знаков всех собственных значений. В таблице 35.1 приведена классификация Бианки трехмерных алгебр Ли. Первые шесть строк отведены случаю $a = 0$, который мы только что обсудили.

Тип	a	$b^{(1)}$	$b^{(2)}$	$b^{(3)}$	η_{ij}
I	0	0	0	0	diag (0, 0, 0)
II	0	1	0	0	diag (0, 0, 0)
VII ₀	0	1	1	0	diag (0, 0, 2)
VI ₀	0	1	-1	0	diag (0, 0, -2)
IX	0	1	1	1	diag (2, 2, 2)
VIII	0	1	1	-1	diag (-2, -2, 2)
V	1	0	0	0	diag (-2, 0, 0)
IV	1	0	0	1	diag (-2, 0, 0)
VII _a , $a > 0$	a	0	1	1	diag (2 - 2a ² , 0, 0)
III	1	0	1	-1	diag (-4, 0, 0)
VI _a , $a > 0$	a	0	1	-1	diag (-2 - 2a ² , 0, 0)

Таблица 35.1: Классификация трехмерных алгебр Ли по Бианки.

Если $a \neq 0$, то с необходимостью $b^{(1)} = 0$. При $b^{(2)} = b^{(3)} = 0$ и $b^{(2)} = 0$, $b^{(3)} \neq 0$, выбрав $k_1 = a$, получим $a' = 1$. Одновременно можно преобразовать $b^{(3)} \neq 0$ в единицу. Если $b^{(2)} \neq 0$ и $b^{(3)} \neq 0$, то возможны два неэквивалентных случая: $b^{(2)} > 0$, $b^{(3)} > 0$ и $b^{(2)} > 0$, $b^{(3)} < 0$.

Если $b^{(2)} > 0$, $b^{(3)} > 0$, то выберем

$$k_2 = \frac{k_1 k_3}{b^{(2)}}, \quad k_1 = \pm \sqrt{b^{(2)}b^{(3)}}. \quad (35.186)$$

Тогда $b'^{(2)} = 1$ и $b'^{(3)} = 1$, а параметр a преобразуется по правилу

$$a' = \pm \frac{a}{\sqrt{b^{(2)}b^{(3)}}}.$$

Мы видим, что с помощью выбора знака k_1 в масштабном преобразовании (35.186) можно изменить знак постоянной a , но не ее модуль. Поэтому, не ограничивая общности, будем считать $a > 0$. Таким образом, в типе VII_a мы имеем однопараметрическое семейство алгебр, которые параметризуются положительными числами $a > 0$.

В случае $b^{(2)} > 0$, $b^{(3)} < 0$ можно проделать аналогичное построение.

Таким образом, по классификации Бианки мы имеем девять типов трехмерных алгебр Ли I–IX, две из которых, VII_a и VI_a параметризуются положительным числом a .

В последней колонке таблицы 35.1 указана форма Киллинга–Картана

$$\eta_{ij} := -f_{ik}^l f_{jl}^k = \varepsilon_{ikl} \varepsilon_{jmn} b^{km} b^{ln} - 2a_i a_j.$$

В выбранном базисе матрица b^{ij} диагональна и $a = (a, 0, 0)$. Поэтому форма Киллинга–Картана тоже диагональна. Ее диагональные элементы равны:

$$\eta_{11} = 2b^{(2)}b^{(3)} - 2a^2, \quad \eta_{22} = 2b^{(1)}b^{(3)} \quad \eta_{33} = 2b^{(1)}b^{(2)}.$$

Они и приведены в таблице.

Согласно теореме Картана 8.13.7 алгебра Ли полупроста (см. раздел 8.13.2) тогда и только тогда, когда ее форма Киллинга–Картана невырождена. Из таблицы следует, что среди трехмерных алгебр Ли только две являются полупростыми: это алгебры VIII и IX.

Алгебры Ли типа I–VII можно реализовать в виде полупрямого произведения алгебры двумерных матриц $\mathbb{R}(2)$ на одномерную алгебру $\mathbb{R}$, которая действует в $\mathbb{R}(2)$ с помощью некоторой матрицы M . Различные типы алгебр соответствуют различным матрицам M , как описано ниже.

Тип I. Коммутационные соотношения имеют вид

$$[L_i, L_j] = 0. \quad (35.187)$$

Это – коммутативная унимодулярная алгебра Ли $\mathbb{R}(3)$ сдвигов в $\mathbb{R}^3$. Соответствующая универсальная накрывающая группа Ли представляет собой коммутативную группу сдвигов в трехмерном евклидовом пространстве, которую можно отождествить с самим $\mathbb{R}^3$. В этом случае $M = 0$.

Тип II. Коммутационные соотношения имеют вид

$$[L_1, L_2] = 0, \quad [L_2, L_3] = L_1, \quad [L_3, L_1] = 0. \quad (35.188)$$

Это – нильпотентная унимодулярная алгебра Ли, которая изоморфна алгебре Ли группы Гейзенберга $\mathbb{H}_3(\mathbb{R})$ (группа верхне треугольных 3×3 матриц с единицами на диагонали, см. пример 8.10.3). В этом случае матрица M нильпотентна, но отлична от нуля (оба собственных значения равны нулю).

Тип III. Коммутационные соотношения имеют вид

$$[L_1, L_2] = -L_3 - L_2, \quad [L_2, L_3] = 0, \quad [L_3, L_1] = L_2 + L_3. \quad (35.189)$$

Это – разрешимая неунимодулярная алгебра Ли.

Перейдем к новому базису

$$\tilde{L}_1 := \frac{1}{2}L_1, \quad \tilde{L}_2 := L_2 + L_3, \quad \tilde{L}_3 := L_2 - L_3.$$

Тогда коммутационные соотношения примут вид

$$[\tilde{L}_1, \tilde{L}_2] = -\tilde{L}_2, \quad [\tilde{L}_1, \tilde{L}_3] = 0, \quad [\tilde{L}_2, \tilde{L}_3] = 0.$$

Сравнение этих коммутационных соотношений с коммутационными соотношениями (8.55) показывает, что алгебра типа III изоморфна прямой сумме $\mathfrak{a}(\mathbb{R}) \oplus \mathbb{R}$, где $\mathfrak{a}(\mathbb{R})$

– алгебра Ли группы аффинных преобразований прямой, натянутой на $\tilde{L}_1$, $\tilde{L}_2$ и $\mathbb{R}$ – алгебра Ли сдвигов прямой. В этом случае одно собственное значение матрицы M равно нулю, а другое – отлично от нуля.

Тип IV. Коммутационные соотношения имеют вид

$$[L_1, L_2] = L_3 - L_2, \quad [L_2, L_3] = 0, \quad [L_3, L_1] = L_3. \quad (35.190)$$

Это – разрешимая неунимодулярная алгебра Ли. Матрица M при этом имеет равные отличные от нуля собственные значения, но не является полупростой.

Тип V. Коммутационные соотношения имеют вид

$$[L_1, L_2] = -L_2, \quad [L_2, L_3] = 0, \quad [L_3, L_1] = L_3. \quad (35.191)$$

Это – разрешимая неунимодулярная алгебра Ли. Матрица M при этом имеет равные отличные от нуля собственные значения и является полупростой.

Тип VI_a. Коммутационные соотношения имеют вид

$$[L_1, L_2] = -L_3 - aL_2, \quad [L_2, L_3] = 0, \quad [L_3, L_1] = L_2 + aL_3. \quad (35.192)$$

Это – разрешимая неунимодулярная (при $a \neq 0$) алгебра Ли. При $a = 0$ алгебра Ли унимодулярна. Алгебра Ли типа VI₀ изоморфна алгебре Ли двумерной группы Пуанкаре $\mathfrak{io}(1, 1)$ (движения двумерной плоскости Минковского), где базисные векторы L_2 и L_3 генерируют трансляции, а L_1 – вращения.

В частном случае $a = 1$ алгебра Ли VI₁ совпадает с алгеброй типа III.

Коммутационные соотношения (35.192) можно упростить. Для этого заметим, что в силу второго коммутационного соотношения произвольные линейные комбинации векторов L_2 и L_3 коммутируют между собой. Совершим линейное преобразование

$$\begin{pmatrix} \tilde{L}_2 \\ \tilde{L}_3 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} L_2 \\ L_3 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} L_2 \\ L_3 \end{pmatrix} = \frac{1}{\det} \begin{pmatrix} D & -B \\ -C & A \end{pmatrix} \begin{pmatrix} \tilde{L}_2 \\ \tilde{L}_3 \end{pmatrix}, \quad (35.193)$$

где $\det := AD - BC \neq 0$. Простые вычисления приводят к следующим коммутаторам:

$$\begin{aligned} [L_1, \tilde{L}_2] &= \frac{1}{\det} (-a \det + AC - BD) \tilde{L}_2 + \frac{1}{\det} (B^2 - A^2) \tilde{L}_3, \\ [L_1, \tilde{L}_3] &= \frac{1}{\det} (C^2 - D^2) \tilde{L}_2 + \frac{1}{\det} (-a \det - AC + BD) \tilde{L}_3. \end{aligned}$$

Положим $B = A$ и $C = -D$ (случай $B = A$ и $C = D$ приводит к вырожденному преобразованию). Тогда коммутационные соотношения принимают вид

$$[L_1, \tilde{L}_2] = -(1 + a) \tilde{L}_2, \quad [L_1, \tilde{L}_3] = (1 - a) \tilde{L}_3.$$

Введем новый базисный вектор

$$\tilde{L}_1 := (1 + a) L_1.$$

Тогда коммутационные соотношения алгебры VI_a примут вид

$$[\tilde{L}_1, \tilde{L}_2] = -\tilde{L}_2, \quad [\tilde{L}_2, \tilde{L}_3] = 0, \quad [\tilde{L}_1, \tilde{L}_3] = h \tilde{L}_3, \quad (35.194)$$

где введен новый параметр

$$h := \frac{1 - a}{1 + a}. \quad (35.195)$$

Функция $h(a)$ монотонно убывает, и при положительных a принимает значения $h \in (-1, 1)$. Поэтому алгебру VI_a , $a > 0$, обозначают также VI_h , $|h| < 1$.

Тип VII_a. Коммутационные соотношения имеют вид

$$[L_1, L_2] = L_3 - aL_2, \quad [L_2, L_3] = 0, \quad [L_3, L_1] = L_2 + aL_3. \quad (35.196)$$

Это – разрешимая неунимодулярная (при $a \neq 0$) алгебра Ли. При $a = 0$ алгебра Ли унимодулярна. Алгебра Ли типа VII_0 изоморфна алгебре Ли неоднородных двумерных вращений $\mathfrak{io}(2)$ (движения двумерной евклидовой плоскости $\mathbb{R}^2$), где базисные векторы L_2 и L_3 генерируют трансляции, а L_1 – вращения.

Коммутационные соотношения (35.196) можно упростить. С этой целью совершим линейное преобразование (35.193). Тогда коммутационные соотношения примут вид

$$\begin{aligned} [L_1, \tilde{L}_2] &= -\frac{1}{\det}(a \det + AC + BD)\tilde{L}_2 + \frac{1}{\det}(A^2 + B^2)\tilde{L}_3, \\ [L_1, \tilde{L}_3] &= -\frac{1}{\det}(C^2 + D^2)\tilde{L}_2 + \frac{1}{\det}(-a \det + AC + BD)\tilde{L}_3. \end{aligned}$$

Положим

$$AC + BD = a \det, \quad \tilde{L}_1 := \frac{\det}{C^2 + D^2} L_1.$$

Тогда

$$[\tilde{L}_1, \tilde{L}_2] = -2a \frac{\det}{C^2 + D^2} \tilde{L}_2 + \frac{A^2 + B^2}{C^2 + D^2} \tilde{L}_3, \quad [\tilde{L}_1, \tilde{L}_3] = -\tilde{L}_2.$$

Пусть выполнено равенство

$$A^2 + B^2 = C^2 + D^2.$$

Это уравнение имеет много решений:

$$\begin{aligned} A &= r \cos \psi, & C &= r \cos \varphi, \\ B &= r \sin \psi, & D &= r \sin \varphi, \end{aligned}$$

где $r > 0$ и $\psi, \varphi \in [0, \pi)$ – произвольные углы. При этом $\det = r^2 \sin(\varphi - \psi)$, и для невырожденности преобразования (35.193) необходимо потребовать выполнение неравенства $\varphi \neq \psi$. Теперь коммутационные соотношения примут вид

$$[\tilde{L}_1, \tilde{L}_2] = -h \tilde{L}_2 + \tilde{L}_3, \quad [\tilde{L}_2, \tilde{L}_3] = 0, \quad [\tilde{L}_1, \tilde{L}_3] = -\tilde{L}_2,$$

где

$$h := 2a \sin(\varphi - \psi).$$

Поскольку $a > 0$ и φ, ψ ($\varphi \neq \psi$) произвольны, то при фиксированном a на модуль h имеется ограничение: $0 < |h| < 2a$.

Тип VIII. Коммутационные соотношения имеют вид

$$[L_1, L_2] = -L_3, \quad [L_2, L_3] = L_1, \quad [L_3, L_1] = L_2. \quad (35.197)$$

Это – полупростая унимодулярная алгебра Ли. Сравнение с коммутационными соотношениями (1.185) показывает, что алгебра типа VIII изоморфна алгебре Лоренца $\mathfrak{so}(2, 1)$. Универсальная накрывающая для группы Лоренца довольно сложно устроена и построена в разделе 8.15. Как многообразие, она некомпактна и диффеоморфна евклидову пространству $\mathbb{R}^3$.

Тип IX. Коммутационные соотношения имеют вид

$$[L_1, L_2] = L_3, \quad [L_2, L_3] = L_1, \quad [L_3, L_1] = L_2. \quad (35.198)$$

Это – полупростая унимодулярная алгебра Ли трехмерных вращений $\mathfrak{so}(3)$ (ср. (1.130)). Как было показано в разделе 1.8, ее универсальной накрывающей является группа Ли $SU(2)$, которая, как многообразие, диффеоморфна трехмерной сфере S^3 . $\square$

Можно доказать, что перечисленные в таблице алгебры Ли неразложимы в прямые суммы, за исключением алгебр типа I и III.

35.5.2 Уравнения движения для однородной вселенной

В настоящем разделе мы покажем, что для однородных моделей вселенных, у которых все пространственные сечения $x^0 := t = \text{const}$ являются главными однородными пространствами, уравнения Эйнштейна сводятся к системе *обыкновенных* дифференциальных уравнений. Это является существенным упрощением задачи нахождения точных решений, хотя построить их в явном виде удается далеко не всегда.

Пусть $\mathbb{H}$ – трехмерное главное однородное пространство с координатами x^μ , $\mu = 1, 2, 3$. Предположим, что на $\mathbb{H}$ определена правоинвариантная метрика, т.е. произвольная невырожденная матрица, заданная в некоторой точке $\mathbb{H}$, разнесена по всему однородному пространству $\mathbb{H}$ с помощью группового действия справа. Для главного однородного пространства такая метрика единственна. Пусть $K = K^\mu \partial_\mu$ – один из векторов Киллинга, действующих на $\mathbb{H}$. Выберем естественные координаты на $\mathbb{R} \times \mathbb{H}$ в виде $\{x^\alpha\} = \{t, x^\mu\}$, $\alpha = 0, 1, 2, 3$. Теперь продолжим действие транзитивной группы преобразований с $\mathbb{H}$ на все $\mathbb{M}$:

$$t \mapsto t' := t, \quad x \mapsto x'^\mu := x^\mu + \epsilon K^\mu(t, x) + o(\epsilon), \quad \epsilon \ll 1. \quad (35.199)$$

То есть группа преобразований пространственных сечений не действует на время.

Определение. Вселенная $\mathbb{M}$ называется *однородной*, если она имеет вид топологического произведения вещественной прямой $\mathbb{R} \ni t$ на трехмерное главное однородное пространство $\mathbb{H} \ni x$, т.е. $\mathbb{M} = \mathbb{R} \times \mathbb{H}$. При этом мы предполагаем, что метрика на $\mathbb{M}$ инвариантна относительно группы транзитивных преобразований, действующих на $\mathbb{H}$ и продолженных на все $\mathbb{M}$ по правилу (35.199). $\square$

Из определения однородной вселенной и теоремы 15.5.1 следует, что в некоторой окрестности каждой точки $\mathbb{M}$ существует такая система координат, что метрика пространства-времени имеет блочно диагональный вид

$$ds^2 = dt^2 + \overset{\circ}{g}_{\mu\nu} dx^\mu dx^\nu, \quad (35.200)$$

где $\overset{\circ}{g}_{\mu\nu}(t, x)$ – инвариантная метрика на однородном пространстве $\mathbb{H}$, зависящая от времени t как от параметра. При этом векторные поля Киллинга не зависят от времени, $K^\mu = K^\mu(x)$.

Отметим отличие метрики (35.200) от метрики однородной и изотропной вселенной. Метрика Фридмана (35.70) в силу изотропии зависит от времени только через масштабный множитель. Теперь мы рассматриваем более широкий класс метрик, т.к. для построения однородной метрики на каждом сечении постоянного времени $t = \text{const}$ в единице группы можно выбрать свою отрицательно определенную матрицу и разнести ее по всему пространству-времени с помощью группового действия.

При этом необходимо только позаботиться только о достаточно гладкой зависимости метрики от времени.

Однородное пространство – это транзитивная группа преобразований $(\mathbb{H}, \mathbb{G})$. Для определенности, мы считаем, что группа преобразований действует справа. В моделях вселенной Фридмана пространственные сечения являются пространствами постоянной кривизны. Группа симметрии действует на них транзитивно, и поэтому они являются однородными пространствами.

По определению пространство $\mathbb{H}$ является главным однородным пространством, и его можно отождествить с трехмерным многообразием группы Ли, $\mathbb{H} = \mathbb{G}$. При этом инвариантные структуры на $\mathbb{G}$ строятся следующим образом. Мы фиксируем произвольную точку $x_0 \in \mathbb{G}$ (например, единицу группы), задаем в этой точке компоненты некоторого тензора и разносим его по всему групповому многообразию с помощью группового действия справа. В результате, по построению, получим инвариантный тензор, заданный на группе Ли $\mathbb{G}$. Здесь важно, что однородное пространство является главным, т.к. в этом случае решение уравнения $x = x_0 a$, где $x, x_0 \in \mathbb{G}$ – две произвольные точки, для $a \in \mathbb{G}$ единственно. Если же однородное пространство не является главным, то в общем случае разнесение тензорных компонент, заданных в точке x_0 , не является единственным и поэтому не определено.

Замечание. Если $\mathbb{H}$ является просто однородным пространством, то в общем случае может существовать несколько элементов группы $\mathbb{G}$, связывающих две точки $\mathbb{H}$: $x = x_0 a = x_0 b$, где $x, x_0 \in \mathbb{H}$ и $a, b \in \mathbb{G}$. Если группа неабелева, то результат переноса метрики из точки x_0 в точку x может оказаться неоднозначным. Например, двумерная сфера $\mathbb{S}^2 \approx \mathrm{SO}(3)/\mathrm{SO}(2)$ является однородным, но не главным однородным пространством. Группа вращений $\mathrm{SO}(3)$ действует на $\mathbb{S}^2$ транзитивно, но не свободно. Поэтому мы не можем получить $\mathrm{SO}(3)$ инвариантную метрику на сфере $\mathbb{S}^2$, выбрав в некоторой точке $x_0 \in \mathbb{S}^2$ произвольную невырожденную матрицу и разнеся ее по сфере с помощью группового действия, т.к. результат перенесения в общем случае зависит от выбора пути на сфере. В то же время на сфере $\mathbb{S}^2$ существует единственная $\mathrm{SO}(3)$ инвариантная метрика – это метрика, индуцированная обычным вложением $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. В этом случае двумерная сфера становится однородным и изотропным пространством (пространством постоянной кривизны).

Трехмерную сферу $\mathbb{S}^3 \approx \mathrm{SO}(4)/\mathrm{SO}(3)$ можно рассматривать как однородное пространство, на котором транзитивно действует группа вращений $\mathrm{SO}(4)$. При этом на ней существует единственная метрика, инвариантная относительно действия группы $\mathrm{SO}(4)$, которая индуцирована обычным вложением $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$. В этом случае сфера $\mathbb{S}^3$ становится пространством постоянной кривизны, и мы возвращаемся к модели однородной и изотропной вселенной, рассмотренной в предыдущих разделах. С другой стороны, сферу $\mathbb{S}^3$ можно рассматривать, как главное однородное пространство, поскольку $\mathbb{S}^3 \approx \mathrm{SU}(2)$. В этом случае на ней можно задать множество метрик, инвариантных относительно действия группы $\mathrm{SU}(2)$ справа (или слева). Для этого достаточно выбрать произвольную невырожденную матрицу (любой сигнатуры), например, в единице группы и разнести ее по групповому многообразию с помощью действия группы справа (слева). Если построенная метрика отрицательно определена, то мы получим однородную модель вселенной, которая и является предметом настоящего раздела. $\square$

В предыдущем разделе мы построили все трехмерные алгебры Ли. Если соответствующая группа Ли является полупростой, то на групповом многообразии существует естественная двусторонне инвариантная метрика – форма Киллинга–Картана.

Из классификации Бианки следует, что таких групп всего две: типы VIII и IX. При этом форма Киллинга–Картана для типа VIII не является знакоопределенной и ее нельзя использовать в качестве инвариантной метрики на пространственных сечениях в космологических моделях. Поэтому в качестве инвариантной метрики мы используем не двусторонне инвариантную метрику, а метрику, инвариантную относительно действия группы справа.

Пусть $R_i(x)$, $i = 1, 2, 3$ – произвольный правоинвариантный базис на группе Ли $\mathbb{G}$, который получается фиксированием некоторого базиса в касательном пространстве к единице группы $R_i(x_0)$ с последующим разнесением по всей группе с помощью группового действия справа (см. раздел 8.3). Алгебра Ли $\mathfrak{g}$ группы Ли $\mathbb{G}$ – это алгебра Ли левоинвариантных векторных полей на $\mathbb{G}$. Базис алгебры Ли $L_i(x)$ строится так же, как и базис правоинвариантных векторных полей, но с помощью группового действия слева. При этом в единице группы, по предположению, векторные поля совпадают, $R_i(x_0) = L_i(x_0)$. В главе 8 было доказано, что базисы удовлетворяют следующим коммутационным соотношениям:

$$[L_i, L_j] = f_{ij}^k L_k, \quad [R_i, R_j] = -f_{ij}^k R_k, \quad (35.201)$$

где f_{ij}^k – структурные константы группы Ли $\mathbb{G}$. Базис алгебры Ли генерирует действие группы Ли справа. В частности, это действие сохраняет правоинвариантные векторные поля. Формально это записывается в виде равенства нулю производной Ли:

$$\mathbf{L}_{L_i} R_j = [L_i, R_j] = 0. \quad (35.202)$$

Построение право- и левоинвариантных векторных полей на группе Ли подробно описано в главе 8.

Если на группе Ли в какой-то окрестности задана система координат x^μ , $\mu = 1, 2, 3$, то базисы право- и левоинвариантных векторных полей можно разложить по координатному базису:

$$R_i = e^\mu_i \partial_\mu, \quad L_i = l^\mu_i \partial_\mu, \quad (35.203)$$

$e^\mu_i(x)$ и $l^\mu_i(x)$ – компоненты соответствующих реперов. В трехмерном случае реперы называют также *триадами*. Инвариантный дуальный базис кокасательных пространств R^i и L^i (базис 1-форм) определяется следующими равенствами:

$$R^i(R_j) = \delta_j^i, \quad L^i(L_j) = \delta_j^i. \quad (35.204)$$

В координатном базисе он имеет следующий вид

$$R^i = dx^\mu e_\mu^i, \quad L^i = dx^\mu l_\mu^i, \quad (35.205)$$

где e_μ^i и l_μ^i – компоненты обратных реперов. Эти базисы удовлетворяют уравнениям Маурера–Картана (8.24):

$$dR^i = \frac{1}{2} R^j \wedge R^k f_{jk}^i, \quad dL^i = -\frac{1}{2} L^j \wedge L^k f_{jk}^i, \quad (35.206)$$

где d и $\wedge$ обозначают внешнюю производную и внешнее произведение форм (см. главу 3).

Пример 35.5.1. Если группа коммутативна, т.е. $f_{ij}^k = 0$, то можно выбрать $R_i = L_i = \partial_i$, где x^i некоторая система координат на $\mathbb{G}$. $\square$

Для нахождения явного вида реперов $e^\mu_i(x)$ и $l^\mu_i(x)$ необходимо знание функции композиции для группы $\mathbb{G}$. К сожалению, она известна только в простейших случаях. Однако для проведения вычислений этого, как правило, и не нужно. Все вычисления можно провести непосредственно в базисе R_i . Этот базис в общем случае неголоном, поскольку коммутатор базисных векторных полей отличен от нуля (35.201). При этом коэффициенты неголономности равны структурным константам, взятым с обратным знаком. Все формулы, необходимые для вычислений, приведены в разделе 6.9.

Ясно, что правоинвариантный базис касательных пространств на всем четырехмерном многообразии $\mathbb{M}$ состоит из векторов $\{\partial_0 := \partial_t, R_i\}$, а форм – из 1-форм $\{dt, R^i\}$. Пусть на $\mathbb{M}$ задано произвольное тензорное поле, например, $T = dx^\alpha T_\alpha^\beta \partial_\beta$. Тогда его можно разложить по правоинвариантному базису:

$$T = dt T_0^0 \partial_0 + dt T_0^i R_i + R^i T_i^0 \partial_0 + R^i T_i^j R_j.$$

Под действием группы $\mathbb{G}$ базисные векторные поля не меняются по построению. Поэтому для того, чтобы тензорное поле было правоинвариантным, необходимо и достаточно, чтобы его компоненты в правоинвариантном базисе зависели только от времени $t = x^0$. Формально условие правой инвариантности тензорного поля записывается в виде равенства нулю производной Ли

$$\mathbf{L}_{L_i} T = 0.$$

При преобразовании координат на $\mathbb{G}$ меняются компоненты триад e^μ_i и l^μ_i по обычному тензорному закону (индекс μ), но не сами векторные поля R_i и L_i . При этом компоненты тензорного поля относительно правоинвариантного базиса ведут себя как скаляры и, следовательно, не зависят от x .

Например, наиболее общая правоинвариантная метрика на $\mathbb{M}$ имеет вид

$$ds^2 = dt^2 + h_{ij} R^i R^j, \quad (35.207)$$

где $h_{ij}(t)$ – произвольная невырожденная симметричная отрицательно определенная матрица, зависящая только от времени. Это значит, что в момент времени t мы выбираем матрицу $h_{ij}(t)$ в единице группы и разносим ее по групповому многообразию с помощью действия группы справа. При этом мы предполагаем, что матрица $h_{ij}(t)$ достаточно гладко зависит от времени.

Приступим к вычислению геометрических характеристик многообразия $\mathbb{M}$ с метрикой (35.207), используя формулы раздела 6.9. Несложные вычисления показывают, что только четыре компоненты линейной связности отличны от нуля

$$\begin{aligned} \omega_{0ij} &= \omega_{i0j} = \frac{1}{2} \dot{h}_{ij}, \\ \omega_{ij0} &= -\frac{1}{2} \dot{h}_{ij}, \\ \omega_{ijk} &= -\frac{1}{2} (f_{ijk} - f_{jki} + f_{kij}), \end{aligned} \quad (35.208)$$

где $f_{ijk} := f_{ij}^l h_{kl}$ и точка обозначает дифференцирование по времени. Теперь можно вычислить компоненты тензора кривизны по формуле (6.112):

$$\begin{aligned} R_{0i0j} &= \frac{1}{2} \ddot{h}_{ij} - \frac{1}{4} \dot{h}_{ik} \dot{h}_{jl} h^{kl}, \\ R_{0ijk} &= -\frac{1}{4} (f_{ik}^l - f_{k^l i} + f_{ik}^l) \dot{h}_{jl} + \frac{1}{4} (f_{ij}^l - f_{j^l i} + f_{ij}^l) \dot{h}_{kl}, \\ R_{ijkl} &= \frac{1}{4} (\dot{h}_{ik} \dot{h}_{jl} - \dot{h}_{il} \dot{h}_{jk}) + \hat{R}_{ijkl}, \end{aligned} \quad (35.209)$$

где тензор кривизны группы Ли имеет вид

$$\begin{aligned} \hat{R}_{ijkl} = & \frac{1}{2} f_{ij}^m f_{klm} + \frac{1}{4} [f_{ik}^m f_{jlm} - f_{jk}^m f_{ilm} + f_{ij}^m (f_{kml} - f_{lmk}) + f_{kl}^m (f_{imj} - f_{jmi}) + \\ & + (f_{ik}^m + f_{ki}^m) (f_{mjl} + f_{mlj}) - (f_{jk}^m + f_{kj}^m) (f_{mil} + f_{mli})]. \end{aligned} \quad (35.210)$$

Следующий шаг состоит в вычислении компонент тензора Риччи:

$$\begin{aligned} R_{00} &= \frac{1}{2} \ddot{h}_{ij} h^{ij} - \frac{1}{4} \dot{h}_{ik} \dot{h}_{jl} h^{ij} h^{kl}, \\ R_{0i} &= \frac{1}{2} f_k^{lk} \dot{h}_{il} - \frac{1}{2} f_i^{kl} \dot{h}_{kl}, \\ R_{ij} &= \frac{1}{2} \ddot{h}_{ij} - \frac{1}{2} \dot{h}_{ik} \dot{h}_{jl} h^{kl} + \frac{1}{4} \dot{h}_{ij} \dot{h}_{kl} h^{kl} + \hat{R}_{ij}, \end{aligned} \quad (35.211)$$

где

$$\hat{R}_{ij} = \frac{1}{2} f_i^{kl} (f_{jkl} + f_{jlk}) + \frac{1}{2} (f^k_{ij} + f^k_{ji}) f_{kl}^l - \frac{1}{4} f^{kl}_i f_{klj}. \quad (35.212)$$

Скалярная кривизна принимает вид

$$R = \ddot{h}_{ij} h^{ij} - \frac{3}{4} \dot{h}_{ik} \dot{h}_{jl} h^{ij} h^{kl} + \frac{1}{4} (\dot{h}_{ij} h^{ij})^2 + \hat{R}, \quad (35.213)$$

где

$$\hat{R} = \frac{1}{4} f^{ijk} f_{ijk} - \frac{1}{2} f^{ijk} f_{kij} + f^{ik}_k f_{ik}^k. \quad (35.214)$$

Для интегрирования уравнений движения удобно ввести тензор внешней кривизны гиперповерхности $t = \text{const}$, который определен формулой (24.44). В рассматриваемом случае блочно диагональной метрики (35.207) – это обычная производная по времени от трехмерной метрики:

$$K_{ij} := -\frac{1}{2} \dot{h}_{ij}, \quad K := K_{ij} h^{ij} = -\frac{1}{\sqrt{|h|}} \frac{d}{dt} \sqrt{|h|}. \quad (35.215)$$

Учитывая формулу

$$\dot{K} = \frac{d}{dt} (K_{ij} h^{ij}) = -\frac{1}{2} \frac{d}{dt} (\dot{h}_{ij} h^{ij}) = -\frac{1}{2} \ddot{h}_{ij} h^{ij} - \frac{1}{2} \dot{h}_{ij} \dot{h}^{ij} = -\frac{1}{2} \ddot{h}_{ij} h^{ij} + 2K_{ij} K^{ij},$$

выражение для нулевой компоненты тензора Риччи (35.211) принимает вид

$$R_{00} = -\dot{K} + K_{ij} K^{ij}. \quad (35.216)$$

Смешанные компоненты тензора Риччи также выражаются через внешнюю кривизну

$$R_{0i} = -f_k^{lk} K_{il} + f_i^{kl} K_{kl}. \quad (35.217)$$

Чтобы упростить выражение для пространственных компонент тензора Риччи, заметим, что справедливы формулы:

$$\begin{aligned} \dot{K}_{ik} h^{kj} &= \frac{d}{dt} (K_{ik} h^{kj}) - K_{ik} \dot{h}^{kj} = \dot{K}_i^j - 2K_{ik} K^{jk}, \\ \frac{1}{\sqrt{|h|}} \frac{d}{dt} (\sqrt{|h|} K_i^j) &= \dot{K}_i^j + \frac{K_i^j}{\sqrt{|h|}} \frac{d}{dt} \sqrt{|h|} = \dot{K}_i^j - K_i^j K, \end{aligned} \quad (35.218)$$

где $h := \det h_{ij}$. Тогда пространственные компоненты тензора Риччи с одним ковариантным и одним контравариантным индексом примут вид

$$R_i^j = -\frac{1}{\sqrt{|h|}} \frac{d}{dt} (\sqrt{|h|} K_i^j) + \hat{R}_i^j. \quad (35.219)$$

Теперь потребуем инвариантности тензора энергии-импульса материи относительно группы $\mathbb{G}$. Вектор скорости будет инвариантен тогда и только тогда, когда его компоненты относительно правоинвариантного базиса будут зависеть только от времени:

$$u = u^\alpha \partial_\alpha = u^0(t) \partial_0 + u^i(t) R_i.$$

Из условия

$$u^2 = (u^0)^2 + u^i u^j h_{ij} = 1$$

находим временную компоненту скорости

$$u^0 = \sqrt{1 - u^i u^j h_{ij}},$$

где мы выбрали положительный корень, т.к. материя движется вперед по времени. Таким образом, инвариантный тензор энергии-импульса в правоинвариантном базисе имеет вид

$$T^{ab} = (\mathcal{E} + \mathcal{P}) u^a u^b - \mathcal{P} g^{ab}, \quad (35.220)$$

где $\{a\} = \{0, i\}$, $a = 0, 1, 2, 3$,

$$\{u^a\} = \{\sqrt{1 - u^k u_k}, u^i\}$$

и метрика имеет блочно диагональный вид (35.207). В правоинвариантном базисе тензор энергии-импульса зависит только от времени t .

Запишем уравнения Эйнштейна в правоинвариантном базисе с одним контра- и одним ковариантным индексом

$$\Phi_a^b := R_a^b - \delta_a^b \Lambda + \frac{1}{2} \left[(\mathcal{E} + \mathcal{P}) u_a u^b - \frac{1}{2} (\mathcal{E} - \mathcal{P}) \delta_a^b \right] = 0. \quad (35.221)$$

С учетом явного вида компонент тензора Риччи (35.216), (35.217) и (35.219) получаем систему уравнений:

$$\begin{aligned} -\dot{K} + K_{ij} K^{ij} &= \Lambda - \frac{1}{2} \left[(\mathcal{E} + \mathcal{P}) (1 - u^i u_i) - \frac{1}{2} (\mathcal{E} - \mathcal{P}) \right], \\ -f_k^{lk} K_{il} + f_i^{kl} K_{kl} &= -\frac{1}{2} (\mathcal{E} + \mathcal{P}) u^0 u_i, \\ -\frac{1}{\sqrt{|h|}} \frac{d}{dt} (\sqrt{|h|} K_i^j) + \hat{R}_i^j &= \delta_i^j \Lambda - \frac{1}{2} \left[(\mathcal{E} + \mathcal{P}) u_i u^j - \frac{1}{2} \delta_i^j (\mathcal{E} - \mathcal{P}) \right]. \end{aligned} \quad (35.222)$$

Таким образом, мы свели уравнения Эйнштейна для однородной вселенной к системе обыкновенных дифференциальных уравнений.

Систему уравнений (35.222) необходимо дополнить уравнениями релятивистской гидродинамики $\nabla_\beta T_a^\beta = 0$, которые являются условиями интегрируемости уравнений Эйнштейна. Поскольку ковариантная производная от репера равна нулю (6.106), то в неголономном базисе уравнения релятивистской гидродинамики примут вид

$$\nabla_b T_a^b = e^\beta_b \partial_\beta T_a^b - \omega_{ba}^c T_c^b + \omega_{bc}^b T_a^c = 0. \quad (35.223)$$

Компоненты линейной связности для однородной вселенной в правоинвариантном базисе уже были посчитаны (35.208). След компонент линейной связности $\omega_a := \omega_{ba}{}^b$ имеет вид

$$\omega_0 = \frac{1}{2} \dot{h}_{ij} h^{ij}, \quad \omega_i = -f_{ji}{}^j. \quad (35.224)$$

Положим в уравнениях (35.223) сначала $a = 0$, а затем $a = i$. Тогда после простых вычислений получим уравнения релятивистской гидродинамики:

$$\begin{aligned} \partial_0 [(\mathcal{E} + \mathcal{P})u_0 u^0 - \mathcal{P}] + \frac{1}{2}(\mathcal{E} + \mathcal{P})\dot{h}_{ij}(h^{ij}u_0 u^0 - u^i u^j) - (\mathcal{E} + \mathcal{P})f_{kj}{}^k u_0 u^j &= 0, \\ \partial_0 [(\mathcal{E} + \mathcal{P})u_i u^0] + \frac{1}{2}(\mathcal{E} + \mathcal{P})\dot{h}_{jk}h^{jk}u_i u^0 - (\mathcal{E} + \mathcal{P})f_{ijk}u^j u^k - (\mathcal{E} + \mathcal{P})f_{kj}{}^k u_i u^j &= 0. \end{aligned} \quad (35.225)$$

Для трехмерного случая структурные константы $f_{ij}{}^k$ можно записать в каноническом виде (35.182), который использовался при классификации Бианки трехмерных алгебр Ли. В каноническом виде след структурных констант имеет вид

$$f_{kj}{}^k = 2a_j.$$

Кроме этого, справедлива формула

$$f_{ijk}u^j u^k = \varepsilon_{ijl}b^{lk}u_k u^j + a_j u^j u_i - a_i u^j u_j.$$

Теперь уравнения релятивистской гидродинамики примут вид

$$\begin{aligned} \partial_0 [(\mathcal{E} + \mathcal{P})u_0 u^0 - \mathcal{P}] + \frac{1}{2}(\mathcal{E} + \mathcal{P})\dot{h}_{ij}(h^{ij}u_0 u^0 - u^i u^j) - 2(\mathcal{E} + \mathcal{P})u_0 a_j u^j &= 0, \\ \partial_0 [(\mathcal{E} + \mathcal{P})u_i u^0] + \frac{1}{2}(\mathcal{E} + \mathcal{P})\dot{h}_{jk}h^{jk}u_i u^0 - (\mathcal{E} + \mathcal{P})(\varepsilon_{ijl}b^{lk}u_k u^j + 3a_j u^j u_i - a_i u^j u_j) &= 0. \end{aligned} \quad (35.226)$$

Конечно, для замыкания системы уравнений в случае баротропной жидкости необходимо дополнить ее уравнением состояния $\mathcal{P} = \mathcal{P}(\mathcal{E})$.

Таким образом, мы получили 15 уравнений (35.222), (35.226) (плюс уравнение состояния) на 11 неизвестных функций h_{ij} , $\mathcal{E}$, $\mathcal{P}$ и u^i . Поскольку уравнения Эйнштейна зависимы, $\partial_b \Phi_a{}^b = 0$, то независимых уравнений движения также одиннадцать.

Система уравнений (35.222) интегрируется в явном виде далеко не всегда. Простейший случай, когда общее решение можно записать явно, рассмотрен в следующем разделе.

35.5.3 Вселенная Казнера

Простейшая и интересная модель однородной, но не изотропной вселенной называется моделью Казнера [287]. Она относится к типу I по классификации Бианки и представляет собой вакуумное решение уравнений Эйнштейна без космологической постоянной.

Однородное пространство типа I – это абелева группа Ли трансляций трехмерного евклидова пространства $\mathbb{R}^3$, которая как многообразие совпадает с $\mathbb{R}^3$. Пусть x^μ , $\mu = 1, 2, 3$ – декартовы координаты в евклидовом пространстве. Поскольку группа абелева, то правоинвариантные векторные поля совпадают с левоинвариантными

векторными полями и координатным базисом: $R_\mu = L_\mu = \partial_\mu$. Поэтому наиболее общая инвариантная метрика (35.200) принимает вид

$$ds^2 = dt^2 + g_{\mu\nu}(t)dx^\mu dx^\nu, \quad (35.227)$$

где $g_{\mu\nu}(t)$ – произвольная симметричная невырожденная отрицательно определенная матрица, зависящая только от времени. Прямая подстановка этой метрики в вакуумные уравнения Эйнштейна приводит к трудоемким вычислениям. Поэтому мы поступим иначе, тем более, что формулы, полученные ниже, еще пригодятся.

Для дальнейших целей рассмотрим более общий случай, когда метрика на пространственных сечениях зависит не только от времени, но и от пространственных координат, $g_{\mu\nu} = g_{\mu\nu}(t, x)$. В разделе 24.4 были получены выражения для тензора кривизны в АДМ параметризации метрики через тензор внешней кривизны $K_{\mu\nu} = K_{\nu\mu}$ вложенной гиперповерхности $\mathbb{H} \hookrightarrow \mathbb{M}$. Поскольку в рассматриваемом случае метрика блочно диагональна, то функции хода и сдвига имеют вид

$$N = 1, \quad N_\mu = 0,$$

что существенно упрощает формулы. Тогда тензор внешней кривизны (24.44) пропорционален производной по времени t от метрики, которую обозначим точкой:

$$K_{\mu\nu} = -\frac{1}{2}\dot{g}_{\mu\nu}. \quad (35.228)$$

Из формул (24.49) следуют выражения для линейно независимых компонент тензора кривизны:

$$\begin{aligned} R_{0\mu 0\nu} &= -\dot{K}_{\mu\nu} - K_\mu{}^\rho K_{\nu\rho}, \\ R_{\mu\nu\rho 0} &= \hat{\nabla}_\mu K_{\nu\rho} - \hat{\nabla}_\nu K_{\mu\rho}, \\ R_{\mu\nu\rho\sigma} &= \hat{R}_{\mu\nu\rho\sigma} + K_{\mu\rho}K_{\nu\sigma} - K_{\mu\sigma}K_{\nu\rho}, \end{aligned} \quad (35.229)$$

где $\hat{\nabla}_\mu$ и $\hat{R}_{\mu\nu\rho\sigma}$ – трехмерные ковариантная производная и тензор кривизны, построенные по метрике $g_{\mu\nu}$. Соответствующий тензор Риччи (24.51) имеет следующие линейно независимые компоненты:

$$\begin{aligned} R_{00} &= -\hat{g}^{\mu\nu}\dot{K}_{\mu\nu} - K^{\mu\nu}K_{\mu\nu}, \\ R_{0\mu} &= \hat{\nabla}_\nu K^\nu{}_\mu - \partial_\mu K, \\ R_{\mu\nu} &= \hat{R}_{\mu\nu} - \dot{K}_{\mu\nu} + K_{\mu\nu}K - 2K_\mu{}^\rho K_{\nu\rho}, \end{aligned} \quad (35.230)$$

где $\hat{g}^{\mu\nu}$ – матрица, обратная к $g_{\mu\nu}$, с помощью которой мы поднимаем пространственные индексы, и

$$K := K^\mu{}_\mu = -\frac{1}{2}\hat{g}^{\mu\nu}\dot{g}_{\mu\nu} = -\frac{\dot{\hat{e}}}{\hat{e}} = -\frac{d}{dt}\ln\hat{e}. \quad (35.231)$$

– след тензора внешней кривизны. Выше $\hat{e} := \sqrt{-\det g_{\mu\nu}}$ – форма пространственного объема, и была использована формула (6.52).

Для скалярной кривизны (24.53) получаем выражение

$$R = \hat{R} - 2\hat{g}^{\mu\nu}\dot{K}_{\mu\nu} - 3K^{\mu\nu}K_{\mu\nu} + K^2. \quad (35.232)$$

Поскольку

$$\hat{g}^{\mu\nu}\dot{K}_{\mu\nu} = \frac{d}{dt}(\hat{g}^{\mu\nu}K_{\mu\nu}) - \frac{d\hat{g}^{\mu\nu}}{dt}K_{\mu\nu} = \dot{K} - 2K^{\mu\nu}K_{\mu\nu},$$

то выражения для R_{00} и R можно упростить:

$$\begin{aligned} R_{00} &= -\dot{K} + K^{\mu\nu} K_{\mu\nu}, \\ R &= \hat{R} - 2\dot{K} + K^{\mu\nu} K_{\mu\nu} + K^2. \end{aligned} \quad (35.233)$$

Заметим также, что справедлива формула

$$\frac{1}{\hat{e}} \frac{d}{dt} (\hat{e} K_\mu^\nu) = \dot{K}_\mu^\nu + \frac{\dot{\hat{e}}}{\hat{e}} K_\mu^\nu = \dot{K}_\mu^\nu - K K_\mu^\nu.$$

Поэтому выражение для пространственной компоненты тензора Риччи с одним верхним и одним нижним индексом также упрощается:

$$R_\mu^\nu = \hat{R}_\mu^\nu - \frac{1}{\hat{e}} \frac{d}{dt} (\hat{e} K_\mu^\nu) \quad (35.234)$$

Для анализа особенностей полезно знать выражение для квадрата тензора кривизны, который является геометрическим инвариантом. Несложные вычисления дают следующий ответ

$$\begin{aligned} R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} &= 4\dot{K}_{\mu\nu} \dot{K}_{\rho\sigma} \hat{g}^{\mu\rho} \hat{g}^{\nu\sigma} + 8\dot{K}_{\mu\nu} K^{\mu\rho} K_\rho^\nu + 2K_\mu^\nu K_\nu^\rho K_\rho^\sigma K_\sigma^\mu + 2(K_\mu^\nu K_\nu^\mu)^2 + \\ &+ 8\hat{\nabla}_\mu K_{\nu\rho} \hat{\nabla}^\mu K^{\nu\rho} - 8\hat{\nabla}_\mu K_{\nu\rho} \hat{\nabla}^\nu K^{\mu\rho} + \hat{R}^{\mu\nu\rho\sigma} \hat{R}_{\mu\nu\rho\sigma} + 4\hat{R}_{\mu\nu\rho\sigma} K^{\mu\rho} K^{\nu\sigma}. \end{aligned} \quad (35.235)$$

Теперь вернемся к решению вакуумных уравнений Эйнштейна для вселенной Казнера. Поскольку пространственная часть метрики (35.227) не зависит от x , то $R_{\mu 0} = 0$. Поэтому необходимо решить только два уравнения $R_{00} = 0$ и $R_\mu^\nu = 0$. Из выражений (35.233) и (35.234) следует, что полная система уравнений движения для вселенной Казнера имеет вид

$$-\dot{K} + K^{\mu\nu} K_{\mu\nu} = 0, \quad (35.236)$$

$$\frac{d}{dt} (\hat{e} K_\mu^\nu) = 0. \quad (35.237)$$

Конечно, эта система уравнений является частным случаем системы уравнений Эйнштейна (35.222) при отсутствии полей материи, $\Lambda = 0$ и $f_{ij}^k = 0$.

Второе уравнение имеет общее решение

$$\hat{e} K_\mu^\nu = -\lambda_\mu^\nu = \text{const},$$

где λ_μ^ν – некоторая постоянная матрица. Свертка полученного решения по индексам μ и ν приводит к равенствам

$$K = -\frac{\lambda}{\hat{e}} = -\frac{\dot{\hat{e}}}{\hat{e}}, \quad \lambda := \lambda_\mu^\mu,$$

где учтено выражение для скалярной внешней кривизны (35.231). Отсюда вытекает, что элемент объема является линейной функцией времени:

$$\hat{e} = \lambda t, \quad (35.238)$$

где мы положили несущественную постоянную интегрирования равной нулю. Растяжкой координат всегда можно обратить след матрицы λ в единицу. Поэтому, не ограничивая общности, положим $\hat{e} = t$. Тогда $\lambda = 1$ и $K = -1/t$. В этом случае

$$K_\mu^\nu = -\frac{1}{t} \lambda_\mu^\nu, \quad (35.239)$$

где $\lambda_\mu{}^\nu$ – постоянная матрица с единичным следом. Подстановка этой формулы в уравнение $R_{00} = 0$ приводит к еще одному условию на неизвестную матрицу

$$\lambda_\mu{}^\nu \lambda_\nu{}^\mu = 1. \quad (35.240)$$

Возвращаясь к определению внешней кривизны (35.228), получаем обыкновенное линейное однородное дифференциальное уравнение на метрику

$$\dot{g}_{\mu\nu} = \frac{2}{t} \lambda_\mu{}^\rho g_{\rho\nu}. \quad (35.241)$$

Это уравнение определяет метрику с точностью до умножения на произвольную отличную от нуля постоянную. Несущественная постоянная интегрирования соответствует сдвигу времени, и будет положена равной нулю.

Обозначим через p_1 , p_2 и p_3 собственные значения матрицы $\lambda_\mu{}^\nu$. Будем считать, что они вещественны и различны. Пусть $n_\mu^{(1)}$, $n_\mu^{(2)}$ и $n_\mu^{(3)}$ соответствующие собственные ковекторы:

$$\lambda_\mu{}^\rho n_\rho^{(i)} = p_i n_\mu^{(i)}, \quad i = 1, 2, 3.$$

Тогда решение уравнения (35.241) можно представить в виде

$$g_{\mu\nu} = -t^{2p_1} n_\mu^{(1)} n_\nu^{(1)} - t^{2p_2} n_\mu^{(2)} n_\nu^{(2)} - t^{2p_3} n_\mu^{(3)} n_\nu^{(3)}. \quad (35.242)$$

Здесь мы выбрали отрицательно определенную метрику в соответствии с нашим соглашением. Теперь выберем векторы $n^{(i)}$ в качестве направляющих векторов новых координатных осей, которые обозначим x, y, z , и нормируем их на единицу. В результате получим *метрику Казнера*

$$ds^2 = dt^2 - t^{2p_1} dx^2 - t^{2p_2} dy^2 - t^{2p_3} dz^2, \quad (35.243)$$

где p_1, p_2 и p_3 – произвольные различные числа, удовлетворяющие двум условиям:

$$p_1 + p_2 + p_3 = 1, \quad p_1^2 + p_2^2 + p_3^2 = 1. \quad (35.244)$$

Первое равенство вытекает из требования $\lambda = 1$, а второе – из условия (35.240).

Чтобы представить себе область допустимых значений параметров p_1, p_2 и p_3 , заметим, что первое условие (35.244) определяет плоскость, а второе – сферу в евклидовом пространстве $(p_1, p_2, p_3) \in \mathbb{R}^3$. Поэтому область допустимых значений параметров представляет собой пересечение плоскости со сферой. По предположению, собственные числа должны быть различны. Расположим их в порядке возрастания $p_1 < p_2 < p_3$. Чтобы найти интервалы возможных значений собственных чисел, рассмотрим случай совпадения двух собственных чисел. Это возможно только для двух троек чисел $(-1/3, 2/3, 2/3)$ и $(0, 0, 1)$. Отсюда следует, что одно из собственных чисел должно быть отрицательно, а два других – положительны, причем их значения лежат в интервалах

$$-\frac{1}{3} \leq p_1 \leq 0, \quad 0 \leq p_2 \leq \frac{2}{3}, \quad \frac{2}{3} \leq p_3 \leq 1. \quad (35.245)$$

Таким образом, метрика Казнера (35.243) описывает однородное, но не изотропное пространство-время. Форма пространственного объема $\hat{e} = t$ линейно растет с течением времени. Это значит, что в целом вселенная расширяется. При этом линейные расстояния вдоль оси x уменьшаются, а вдоль осей y и z – увеличиваются.

Рассмотрим некоторые свойства тензора кривизны для метрики Казнера. Для вычислений нам понадобится явный вид тензора внешней кривизны

$$K_{\mu}^{\nu} = -\frac{1}{t} \begin{pmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{pmatrix}, \quad K_{\mu\nu} = \begin{pmatrix} p_1 t^{2p_1-1} & 0 & 0 \\ 0 & p_2 t^{2p_2-1} & 0 \\ 0 & 0 & p_3 t^{2p_3-1} \end{pmatrix}. \quad (35.246)$$

Простые вычисления показывают, что скалярная кривизна (35.232) для метрики Казнера равна нулю, $R = 0$. Это естественно, т.к. для метрики Казнера тензор Риччи равен нулю, $R_{\alpha\beta} = 0$ (вакуумные уравнения движения). Однако это не значит, что вселенная Казнера является плоской. Действительно, квадрат тензора кривизны (35.235), который является геометрическим инвариантом, принимает вид

$$R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} = 4\dot{K}_{\mu\nu}\dot{K}_{\rho\sigma}\hat{g}^{\mu\rho}\hat{g}^{\nu\sigma} + 8\dot{K}_{\mu\nu}K^{\mu\rho}K^{\nu}_{\rho} + 2K_{\mu}^{\nu}K_{\nu}^{\rho}K_{\rho}^{\sigma}K_{\sigma}^{\mu} + 2(K_{\mu}^{\nu}K_{\nu}^{\mu})^2.$$

Отметим, что в рассматриваемом случае тензор кривизны совпадает с тензором Вейля. Подстановка в это выражение тензора внешней кривизны (35.246) приводит к следующему результату

$$R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} = \frac{4}{t^4} (2 - 2p_1^3 - 2p_2^3 - 2p_3^3 - p_1^2 p_2^2 - p_2^2 p_3^2 - p_1^2 p_3^2). \quad (35.247)$$

В общем случае правая часть обращается в бесконечность при $t = 0$. То есть решение Казнера содержит неустранимую особенность, где тензор кривизны расходится. Заметим, что при $t = 0$ метрика Казнера (35.243) также имеет особенность, т.к.

$$g_{11} \rightarrow \infty, \quad g_{22} \rightarrow 0, \quad g_{33} \rightarrow 0, \quad \det g \rightarrow 0.$$

Эта особенность соответствует большому взрыву.

Посмотрим, существуют ли среди метрик Казнера плоские. В плоском случае полный тензор кривизны должен обратиться в нуль, $R_{\alpha\beta\gamma\delta} = 0$. Для этого необходимо, чтобы $R_{\mu\nu\rho\sigma} = 0$. Из вида пространственных компонент тензора кривизны (35.229) вытекает необходимое условие того, чтобы метрика Казнера была плоской:

$$K_{11}K_{22} = 0, \quad K_{11}K_{33} = 0, \quad K_{22}K_{33} = 0.$$

Если собственные числа упорядочены, то эта система уравнений имеет единственное решение $p_1 = p_2 = 0$ и $p_3 = 1$. В этом пределе метрика Казнера принимает вид

$$ds^2 = dt^2 - dx^2 - dy^2 - t^2 dz^2.$$

После замены координат

$$\tau := t \operatorname{ch} z, \quad \zeta := t \operatorname{sh} z,$$

метрика Казнера переходит в метрику Лоренца.

Случаи совпадающих собственных значений матрицы λ_{μ}^{ν} рассмотрены в [169], §117.

Глава 36

Геометрическая теория дефектов

Многие твердые тела обладают кристаллической структурой. Однако в природе идеальных кристаллов нет, и большинство их физических свойств, таких как пластичность, плавление, рост и др., определяется дефектами кристаллической структуры. Поэтому изучение дефектов является актуальной научной проблемой, важной, в первую очередь, для приложений. Интенсивные экспериментальные и теоретические исследования дефектов в кристаллах начались в 30-е годы прошлого века и продолжаются по сей день. Несмотря на десятки монографий и тысячи статей, фундаментальная теория дефектов в настоящее время отсутствует.

Один из перспективных подходов к созданию теории дефектов основан на геометрии Римана–Картана, которая задается нетривиальной метрикой и кручением. В этом подходе кристалл рассматривается, как непрерывная упругая среда со спиновой структурой. Если векторное поле смещений является гладкой функцией, то в кристалле присутствуют только упругие напряжения, которые соответствуют диффеоморфизмам плоского евклидова пространства. Если поле смещений имеет разрывы, то мы говорим, что в среде есть дефекты упругой структуры. Дефекты упругой структуры называются дислокациями и приводят к возникновению нетривиальной геометрии. А именно, они соответствуют отличному от нуля тензору кручения, который равен поверхностной плотности вектора Бюргерса.

Идея связать кручение с дислокациями возникла в пятидесятые годы [288, 289, 290, 291]. Интересно отметить, что Э. Картан ввел понятие кручения в геометрию [292], проводя аналогию с механикой упругой среды.

Параллельно с изучением дислокаций шло интенсивное изучение другого типа дефектов. Дело в том, что многие тела обладают не только упругими свойствами, но и спиновой структурой. Например, ферромагнетики, жидкие кристаллы, спиновые стекла и др. В этом случае существуют дефекты в спиновой структуре, которые называются дисклинациями [293]. Они возникают тогда, когда поле директора имеет разрывы. Наличие дисклинаций также связано с нетривиальной геометрией. А именно, тензор кривизны равен поверхностной плотности вектора Франка.

Геометрическая теория статического распределения дефектов, которая с единой точки зрения описывает оба типа дефектов – дислокации и дисклинации, была предложена в [294]. В отличие от других подходов единственными независимыми переменными в данной модели являются репер и $SO(3)$ связность. Тензоры кручения и кривизны имеют прямой физический смысл поверхностной плотности, соответственно, линейных дислокаций и дисклинаций. Для репера и $SO(3)$ связности постулируются ковариантные уравнения равновесия, такие же, как и в модели гравитации с кручением. Поскольку любое решение уравнений равновесия определено с точностью

до общих преобразований координат и локальных $\mathbb{SO}(3)$ вращений, то для однозначного определения решения необходимо задать систему координат (зафиксировать калибровку). Для этой цели была предложена упругая калибровка для репера [295] и лоренцева калибровка для $\mathbb{SO}(3)$ связности [296].

Подчеркнем, что в данном подходе понятие векторного поля смещений и поля угла поворота полностью отсутствует. Эти переменные могут быть определены только в тех областях среды, в которых дефекты отсутствуют. В этом случае уравнения равновесия для репера и $\mathbb{SO}(3)$ связности тождественно удовлетворяются, упругая калибровка сводится к уравнениям нелинейной теории упругости для вектора смещений, а лоренцева калибровка – к уравнениям главного кирального $\mathbb{SO}(3)$ поля. Другими словами, для фиксирования калибровки выбираются две фундаментальные модели: теория упругости и модель главного кирального $\mathbb{SO}(3)$ поля.

Важно отметить, что геометрическая теория дефектов описывает как отдельные дефекты, так и их непрерывное распределение. Отдельные дислокации можно описать в рамках классической теории упругости путем наложения сложных граничных условий. Однако при непрерывном распределении дислокаций векторное поле смещений не существует, т.к. оно имеет разрывы в каждой точке. Следовательно, описание непрерывного распределения дислокаций в рамках классической теории упругости невозможно. Это же относится и к описанию непрерывного распределения дисклинаций. В геометрической теории дефектов динамическими переменными являются репер и $\mathbb{SO}(3)$ связность. Эти переменные существуют и являются достаточно гладкими функциями и для непрерывного распределения дефектов.

36.1 Упругие деформации

Рассмотрим трехмерную бесконечную упругую среду. Мы предполагаем, что в отсутствие дефектов недеформированная среда инвариантна относительно трансляций и вращений в некоторой системе координат. Тогда в этой системе координат y^i , $i = 1, 2, 3$, среда описывается евклидовой метрикой $\delta_{ij} = \text{diag}(+, +, +)$, и система координат является декартовой. Таким образом, в недеформированном состоянии мы имеем евклидово пространство $\mathbb{R}^3$ с заданной декартовой системой координат. Мы также предполагаем, что кручение в среде равно нулю.

Для того, чтобы определить векторное поле смещений, евклидово пространство $\mathbb{R}^3$, помимо структуры многообразия, должно быть снабжено структурой векторного пространства.

Определение. Пусть некоторая точка среды имеет в основном состоянии координаты y^i . После деформации данная точка в исходной системе будет иметь координаты

$$y^i \mapsto x^i(y) := y^i + u^i(x), \quad (36.1)$$

где $u^i(x)$ – векторное поле смещений, рис. 36.1. □

В теории упругости приняты обратные обозначения. Обычно пишут $x^i \mapsto y^i = x^i + u^i(x)$. Это – эквивалентные записи, поскольку обе системы координат x^i и y^i покрывают все $\mathbb{R}^3$. Однако в теории дефектов, которая рассматривается в следующих разделах, ситуация другая. В общем случае, только в конечном состоянии (после создания дислокаций) упругая среда заполняет все евклидово пространство $\mathbb{R}^3$. При наличии дислокаций координаты y^i , как правило, не покрывают все $\mathbb{R}^3$, т.к. часть

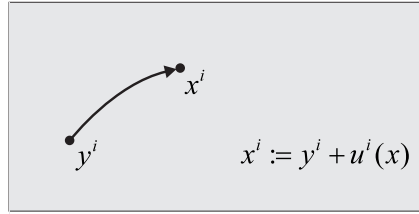


Рис. 36.1: Упругие деформации.

среды может быть удалена или, наоборот, добавлена. Поэтому предпочтительнее является система координат, связанная с точками среды после упругой деформации и образования дефектов.

Здесь и в дальнейшем все поля предполагаются достаточно гладко зависящими от координат x , которые являются координатами точек среды после деформации и, по предположению, покрывают все евклидово пространство $\mathbb{R}^3$.

В линейной теории упругости относительные деформации предполагаются малыми, $\partial_j u^i \ll 1$, и функции $u^i(x)$ – компоненты векторного поля смещений – являются основными переменными теории упругости.

В отсутствие дефектов мы полагаем, что поле смещений является достаточно гладким векторным полем в евклидовом пространстве $\mathbb{R}^3$. Наличие у поля смещений разрывов и (или) особенностей интерпретируется, как наличие дефектов в упругой среде.

В дальнейшем мы будем рассматривать только статические деформации, когда поле смещений u^i не зависит от времени. Тогда основные уравнения равновесия упругой среды для малых деформаций имеют вид (см., например, [297], глава I, §2.4)

$$\partial_j \sigma^{ji} + f^i = 0, \quad (36.2)$$

$$\sigma^{ij} = \lambda \delta^{ij} \epsilon_k^k + 2\mu \epsilon^{ij}, \quad (36.3)$$

где σ^{ji} – тензор напряжений (i -я компонента упругой силы, действующей на единичную площадку с нормалью n^j), который предполагается симметричным. Тензор малых деформаций ϵ_{ij} задается симметризованной частной производной от вектора смещений

$$\epsilon_{ij} := \frac{1}{2}(\partial_i u_j + \partial_j u_i). \quad (36.4)$$

Опускание и подъем латинских индексов проводится с помощью евклидовой метрики δ_{ij} и ее обратной δ^{ij} . Буквы λ и μ обозначают постоянные, характеризующие упругие свойства среды, и называются *коэффициентами Ламе*. Функции $f^i(x)$ описывают суммарную плотность неупругих сил внутри среды. В дальнейшем мы предполагаем, что такие силы отсутствуют: $f^i(x) = 0$. Уравнение (36.2) есть второй закон Ньютона, а равенство (36.3) представляет собой *закон Гука*, связывающий напряжения и деформации.

В декартовой системе координат при малых деформациях различие между верхними и нижними индексами пропадает, поскольку подъем и опускание индексов проводится с помощью евклидовой метрики. По этой причине в теории упругости об этом различии часто забывают, что вполне оправдано. Однако при наличии дефектов понятие декартовой системы координат и евклидовой метрики отсутствует, а подъем и опускание индексов производится с помощью римановой метрики. Поэтому мы делаем различие между верхними и нижними индексами, как это принято в

дифференциальной геометрии, имея в виду последующий переход к упругим средам с дефектами.

После подстановки закона Гука (36.3) в закон Ньютона (36.2) при $f^i = 0$ возникает уравнение второго порядка на вектор смещения:

$$(1 - 2\sigma)\Delta u_i + \partial_i \partial_j u^j = 0, \quad (36.5)$$

где

$$\sigma := \frac{\lambda}{2(\lambda + \mu)}$$

– безразмерный коэффициент Пуассона (из физических соображений [169] на него есть ограничения $-1 \leq \sigma \leq 1/2$) и $\Delta := \delta^{ij} \partial_i \partial_j$ – трехмерный оператор Лапласа.

В линейной теории упругости основной задачей является решение системы уравнений второго порядка для вектора смещений (36.5) при некоторых граничных условиях. Множество известных решений находится в прекрасном согласии с экспериментом. Поэтому можно сказать, что уравнения (36.3) и (36.2) имеют хорошую экспериментальную основу.

Посмотрим на упругие деформации с точки зрения дифференциальной геометрии. С математической точки зрения отображение (36.1) представляет собой диффеоморфизм евклидова пространства $\mathbb{R}^3$. При этом евклидова метрика δ_{ij} индуцирована возвратом отображения $y^i \mapsto x^i$. Это значит, что в деформированном состоянии метрика в линейном приближении, $\partial_i u^j \ll 1$, имеет вид

$$g_{ij}(x) = \frac{\partial y^k}{\partial x^i} \frac{\partial y^l}{\partial x^j} \delta_{kl} \approx \delta_{ij} - \partial_i u_j - \partial_j u_i = \delta_{ij} - 2\epsilon_{ij}, \quad (36.6)$$

т.е. определяется тензором малых деформаций (36.4). Заметим, что в линейном приближении выполнены равенства $\epsilon_{ij}(x) = \epsilon_{ij}(y)$ и $\partial u_j / \partial x^i = \partial u_j / \partial y^i$.

Элемент объема среды определяется определителем метрики и имеет вид $\sqrt{\det g_{ij}}$. Поэтому в общем случае упругие деформации сопровождаются сжатием или растяжением среды. В линейном приближении изменение объема определяется тензором деформаций:

$$1 = \sqrt{\det \delta_{ij}} \mapsto \sqrt{\det g_{ij}} \approx \sqrt{1 - 2\epsilon_{ij}} \approx 1 - \epsilon, \quad (36.7)$$

где

$$\epsilon := \epsilon_i^i \quad (36.8)$$

– *дилатация*. Мы видим, что если $\epsilon > 0$, то среда при упругой деформации сжимается. Если $\epsilon < 0$, то происходит растяжение. Объемы сохраняются, если дилатация равна нулю, $\epsilon = 0$.

В римановой геометрии метрика однозначно определяет связность Леви-Чевиты $\tilde{\Gamma}_{ij}^k(x)$ (символы Кристоффеля) (4.20). По этим символам можно вычислить тензор кривизны (6.74). Этот тензор тождественно равен нулю, $\tilde{R}_{ijk}^l(x) = 0$, т.к. кривизна евклидова пространства равна нулю, а отображение $y^i \mapsto x^i$ – диффеоморфизм. По этой же причине равен нулю и тензор кручения. Таким образом, упругая деформация среды соответствует тривиальной геометрии Римана–Картана, т.к. тензоры кривизны и кручения равны нулю.

Физическая интерпретация метрики (36.6) следующая. Внешний наблюдатель фиксирует декартову систему координат, соответствующую основному недеформированному состоянию среды. Затем происходит деформация среды, и наблюдатель обнаруживает, что в этой системе координат метрика становится нетривиальной.

Если предположить, что упругие возмущения в среде (фононы) распространяются вдоль экстремалей (линий наименьшей длины), то в деформированном состоянии их траектории будут определяться уравнениями (16.21). Поскольку символы Кристоффеля нетривиальны, $\tilde{\Gamma}_{jk}^i \neq 0$, то траекториями фононов уже не будут прямые линии. В этом смысле метрика (36.6) наблюдаема. Здесь проявляется существенная роль декартовой системы координат y^i , определяемой недеформированным состоянием, и с которой связан процесс измерения.

Предположим, что некоторому напряженному состоянию упругой среды без дефектов соответствует метрика $g_{ij}(x)$, заданная в декартовой системе координат. В этом случае вектор смещения определяется системой дифференциальных уравнений в частных производных (36.6). Условием интегрируемости данной системы уравнений при заданной левой части является равенство нулю тензора кривизны. Это является следствием теоремы 6.7.3. В линейном приближении данные условия известны в теории упругости, как *условия интегрируемости Сен-Венана*.

Наряду с линейной рассматривается нелинейная теория упругости. Нелинейность в теорию упругости, как правило, вводится двумя способами. Во-первых, можно изменить определение тензора деформаций. Положим

$$\epsilon_{ij} := \frac{1}{2}(\delta_{ij} - g_{ij}). \quad (36.9)$$

Тогда тензор деформаций задается бесконечным рядом по вектору смещений. Во-вторых, линейный закон Гука (36.3) можно заменить на нелинейное соотношение. Здесь вариантов очень много.

Сделаем замечание, важное для дальнейшего рассмотрения. При надлежащих граничных условиях решение уравнений теории упругости (36.2), (36.3) единственно. С геометрической точки зрения это означает, что теория упругости фиксирует диффеоморфизмы, поскольку векторное поле смещений параметризует диффеоморфизмы. Этот факт будет использован в геометрической теории дефектов. Уравнения нелинейной теории упругости в смысле (36.9), переписанные для метрики или репера, будут использованы для фиксирования системы координат.

36.2 Дислокации

Начнем с описания линейных дислокаций в упругой среде (см., например, [297, 298]). Простейшие и наиболее распространенные примеры прямолинейных дислокаций изображены на рис. 36.2. Они образованы следующим образом. Разрежем среду по полуплоскости $x^2 = 0$, $x^1 > 0$. Сдвинем верхнюю часть среды, расположенную над разрезом $x^2 > 0$, $x^1 > 0$, на вектор $\mathbf{b}$ по направлению к оси дислокации x^3 и склеим берега разреза. Вектор $\mathbf{b}$ называется *вектором Бюргерса*. В общем случае вектор Бюргерса может быть не постоянен на разрезе. Для краевой дислокации он меняется от нуля до постоянного значения $\mathbf{b}$ по мере удаления от оси дислокации. После склейки среда придет в некоторое равновесное состояние, которое называется *краевой дислокацией*, изображенной на рис. 36.2a.

Та же краевая дислокация возникнет, если сдвинуть нижний берег разреза в противоположном направлении от линии дислокации и затем произвести склейку.

Одну и ту же дислокацию можно получить разными способами. Например, если в рассмотренном примере краевой дислокации вектор Бюргерса перпендикулярен плоскости разреза и направлен от него, то перед склейкой образовавшуюся полость

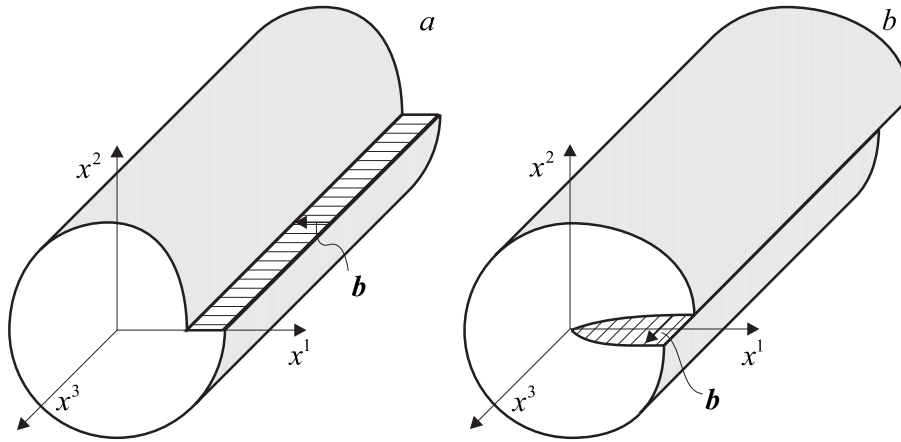


Рис. 36.2: Прямолинейные дислокации. (a) Краевая дислокация. Вектор Бюргерса $\mathbf{b}$ перпендикулярен оси дислокации. (b) Винтовая дислокация. Вектор Бюргерса $\mathbf{b}$ параллелен оси дислокации.

необходимо заполнить веществом. В результате, как нетрудно представить, также возникает краевая дислокация, только повернутая вокруг оси x^3 на угол $\pi/2$. Этот пример показывает, что характеристикой дислокации является не поверхность разреза, а линия дислокации (край разреза) и вектор Бюргерса.

Если произвести такой же разрез среды по полуплоскости, сдвинуть нижнюю часть среды вдоль оси x^3 и произвести склейку, то получим *винтовую* дислокацию, рис. 36.2b. В этом случае вектор Бюргерса параллелен линии дислокации. Эту же винтовую дислокацию можно получить, сдвинув перед склейкой верхний берег разреза в противоположном направлении.

С топологической точки зрения среда, содержащая дислокации, которых может быть несколько или даже бесконечное число, представляет собой евклидово пространство $\mathbb{R}^3$. В отличие от упругих деформаций при наличии дислокаций вектор смещений перестает быть гладкой функцией, т.к. имеются поверхности разреза. В то же время мы предполагаем, что частные производные от вектора смещений $\partial_j u^i$ (*тензор дисторсии*) являются достаточно гладкими функциями на поверхности разреза. С физической точки зрения такое предположение оправдано тем, что эти производные определяют тензор деформаций (36.4). В свою очередь частные производные от тензора деформаций должны существовать и быть непрерывными функциями в равновесном состоянии везде, за исключением, может быть, линий дислокаций, т.к. в противном случае уравнение равновесия (36.2) не имеет смысла. Поскольку тензор деформаций определяет индуцированную метрику (36.6), то мы предполагаем, что метрика и репер являются достаточно гладкими функциями во всем евклидовом пространстве $\mathbb{R}^3$, за исключением, возможно, линий дислокаций.

Основная идея геометрического подхода сводится к следующему. Для описания отдельных дислокаций в рамках теории упругости необходимо решить уравнения равновесия для вектора смещений с определенными граничными условиями на разрезах. Для небольшого числа дислокаций это возможно. Однако с увеличением количества дислокаций граничные условия настолько усложняются, что решение задачи представляется нереальным. Кроме того, одну и ту же дислокацию можно создать с помощью различных разрезов, что приводит к неоднозначности для векторного поля смещений. Другим недостатком этого подхода является его непригодность для описания непрерывного распределения дислокаций, т.к. в этом случае векторное поле

смещений просто не существует, поскольку должно иметь разрывы в каждой точке. В геометрическом подходе основная переменная – репер, который, по предположению, является достаточно гладкой функцией везде за исключением, возможно, линий дислокаций. Для репера постулируются новые уравнения (см. раздел 36.4). В геометрическом подходе переход от конечного числа дислокаций к их непрерывному распределению происходит просто и естественно. При этом происходит сглаживание сингулярностей на линиях дислокаций аналогично тому, как сглаживается распределение масс для точечных частиц при переходе к непрерывной среде.

Перейдем к построению формализма геометрического подхода. Поскольку при наличии дефектов в равновесном состоянии в общем случае у нас нет никакой симметрии, то отсутствует понятие выделенной декартовой системы координат. Поэтому рассмотрим произвольную систему координат x^μ , $\mu = 1, 2, 3$, в $\mathbb{R}^3$. Для нумерации координат мы используем теперь греческие буквы, поскольку допускаем произвольные преобразования координат. Тогда вектор Бюргерса можно выразить в виде интеграла от вектора смещения:

$$\oint_C dx^\mu \partial_\mu u^i(x) = - \oint_C dx^\mu \partial_\mu y^i(x) = -b^i, \quad (36.10)$$

где C – замкнутый контур, охватывающий линию дислокации, рис. 36.3. Этот ин-

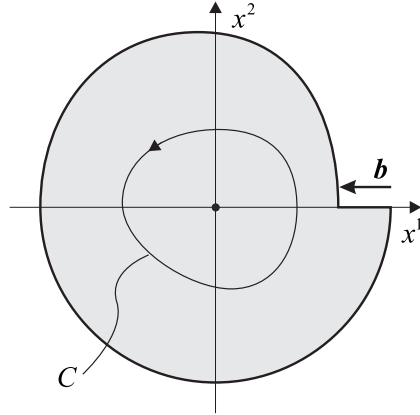


Рис. 36.3: Сечение среды с краевой дислокацией. C – контур интегрирования для вектора Бюргерса $\mathbf{b}$.

теграл инвариантен относительно произвольных преобразований координат $x^\mu \mapsto x^{\mu'}(x)$ и ковариантен относительно глобальных $\mathbb{SO}(3)$ вращений координат y^i . Здесь компоненты векторного поля смещений $u^i(x)$ рассматриваются относительно ортонормального базиса касательного пространства $u = u^i e_i$. Если бы мы рассматривали компоненты векторного поля смещений относительно координатного базиса $u = u^\mu \partial_\mu$, то инвариантность интеграла (36.10) относительно общих преобразований координат была бы нарушена.

В геометрическом подходе вместо частных производных от векторного поля смещений $\partial_\mu u^i$ мы вводим новую независимую переменную – репер – следующим образом

$$e_\mu^i(x) := \begin{cases} \partial_\mu y^i, & \text{вне разреза,} \\ \lim \partial_\mu y^i, & \text{на разрезе.} \end{cases} \quad (36.11)$$

По построению, репер является непрерывной функцией на разрезе, поскольку в равновесном состоянии среды упругие силы на обоих берегах разреза должны быть

равны. Более того, мы предполагаем, что он является достаточно гладким. Отметим, что, если бы мы просто определили репер, как частную производную $\partial_\mu y^i$, то у репера была бы δ -образная особенность на разрезе, т.к. функции $y^i(x)$ испытывают скачок. Эти δ -функции в определении (36.11) отброшены.

Вектор Бюргерса по теореме Стокса можно представить в виде интеграла по поверхности S , натянутой на замкнутый контур C ,

$$\oint_C dx^\mu e_\mu^i = \iint_S dx^\mu \wedge dx^\nu (\partial_\mu e_\nu^i - \partial_\nu e_\mu^i) = b^i, \quad (36.12)$$

где $dx^\mu \wedge dx^\nu$ – элемент площади поверхности. Из определения репера (36.11) следует, что подынтегральное выражение равно нулю всюду, за исключением линии дислокации. Для краевой дислокации с постоянным вектором Бюргерса подынтегральное выражение имеет δ -образную особенность на оси z . Критерием наличия дислокации является нарушение условия интегрируемости системы уравнений $\partial_\mu y^i = e_\mu^i$:

$$\partial_\mu e_\nu^i - \partial_\nu e_\mu^i \neq 0. \quad (36.13)$$

Если дислокаций нет, то функции $y^i(x)$ существуют во всем $\mathbb{R}^3$ согласно лемме Пуанкаре и определяют переход к декартовой системе координат.

В геометрической теории дефектов поле e_μ^i отождествляется с репером. Далее, сравним подынтегральное выражение в (36.12) с выражением для тензора кручения в переменных Картана (5.56)

$$T_{\mu\nu}^i = \partial_\mu e_\nu^i - \partial_\nu e_\mu^i - e_\mu^j \omega_{\nu j}^i + e_\nu^j \omega_{\mu j}^i. \quad (36.14)$$

Они отличаются только слагаемыми, содержащими $\mathbb{SO}(3)$ связность. Это дает основание принять следующий постулат. В геометрической теории дефектов вектор Бюргерса, соответствующий некоторой поверхности S , определяется интегралом от тензора кручения:

$$b^i := \iint_S dx^\mu \wedge dx^\nu T_{\mu\nu}^i. \quad (36.15)$$

Это определение инвариантно относительно общих преобразований координат x^μ и ковариантно относительно глобальных вращений. Таким образом, в геометрической теории дефектов тензор кручения имеет прямой физический смысл: он равен поверхностной плотности вектора Бюргерса.

В геометрической теории дефектов выражение для вектора Бюргерса (36.15) является постулатом, и все предыдущие построения – не более, чем наводящие соображения.

Физический смысл $\mathbb{SO}(3)$ связности будет дан в разделе 36.3, а сейчас мы покажем, каким образом данное определение сводится к выражению для вектора Бюргерса (36.12), полученному в рамках теории упругости. Если тензор кривизны $\mathbb{SO}(3)$ связности равен нулю, то по теореме 5.5.1 связность локально тривиальна, и существует такое $\mathbb{SO}(3)$ вращение, что $\omega_{\mu j}^i = 0$. В этом случае мы возвращаемся к выражению (36.12).

Если $\mathbb{SO}(3)$ связность равна нулю и репер является достаточно гладким, то с каждым контуром можно однозначно связать вектор Бюргерса. При этом его можно выразить как поверхностный интеграл от тензора кручения. Поверхностный интеграл зависит только от граничного контура, но не от поверхности в силу теоремы Стокса.

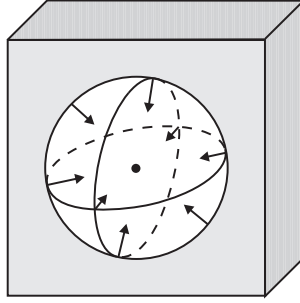


Рис. 36.4: Точечный дефект – вакансия, возникает, когда из среды вырезается шар и граничная сфера стягивается в точку.

Мы показали, что наличие линейных дислокаций приводит к нетривиальному кручению. В геометрической теории дефектов равенство нулю тензора кручения $T_{\mu\nu}^i = 0$ естественно рассматривать, как критерий отсутствия дислокаций. Тогда под термин дислокация попадают не только линейные дислокации, но и, по сути дела, произвольные дефекты в упругой среде. Они получаются с помощью произвольных разрезов среды и склеек. При этом часть среды может быть удалена или, наоборот, добавлена, если образовались полости. Например, точечные дефекты – *вакансии и примеси* также являются дислокациями. В первом случае из евклидова пространства $\mathbb{R}^3$ вырезается шар вещества, граничная сфера стягивается в точку и склеивается (рис. 36.4). В случае примеси точка евклидова пространства раздувается до сферы и образовавшаяся шаровидная полость заполняется таким же веществом. Точечные дефекты характеризуются массой удаленного или добавленного вещества, которая также определяется репером:

$$M := \rho_0 \int_{\mathbb{R}^3} dx \left(\det e_\mu^i - \det \overset{\circ}{e}_\mu^i \right), \quad \overset{\circ}{e}_\mu^i := \partial_\mu y^i, \quad (36.16)$$

где $y^i(x)$ – функции перехода к декартовой системе координат в $\mathbb{R}^3$, ρ_0 – плотность массы среды в равновесном состоянии, которую мы предполагаем постоянной. Масса определяется разностью двух интегралов, каждый из которых по отдельности расходится. Первый интеграл равен объему среды с дефектами, а второй интеграл – объему евклидова пространства. Тензор кручения для вакансий и примесей равен нулю всюду за исключением одной точки, где он имеет δ -образную особенность. Для точечных дефектов понятие вектора Бюргерса отсутствует.

Согласно данному определению масса примеси положительна, т.к. вещество добавляется к среде, а масса вакансии отрицательна, поскольку часть вещества удалена.

В трехмерном пространстве помимо точечных и линейных дислокаций существуют также поверхностные дефекты. В геометрическом подходе все они относятся к дислокациям, т.к. связаны с нетривиальным тензором кручения.

36.3 Дисклинации

В предыдущем разделе мы связали дислокации в упругой среде с тензором кручения. При этом была введена $\mathbb{SO}(3)$ связность. Ниже мы покажем, что кривизна $\mathbb{SO}(3)$ связности определяет поверхностную плотность *вектора Франка*, характеризующего другие хорошо известные дефекты – линейные *дисклинации* в спиновой структуре среды [297].

Пусть в каждой точке среды задано единичное векторное поле $n^i(x)$, $n^i n_i = 1$. Например, для ферромагнетиков n^i имеет смысл магнитного момента, связанного с каждой точкой среды (рис. 36.5a). Для нематических жидких кристаллов единичное векторное поле n^i вместе с отношением эквивалентности $n^i \sim -n^i$ описывает поле директора (рис. 36.5b).

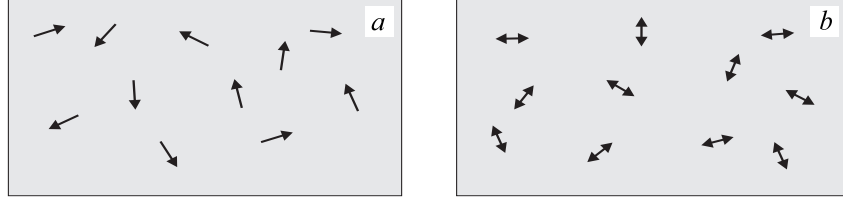


Рис. 36.5: Примеры сред со спиновой структурой: (a) ферромагнетик, (b) нематический жидкий кристалл.

Зафиксируем некоторое направление в среде n_0^i . Тогда поле $n^i(x)$ в точке x может быть однозначно задано полем $\omega^{ij}(x) = -\omega^{ji}(x)$, принимающем значения в алгебре вращений $\mathfrak{so}(3)$ (угол вращения):

$$n^i = n_0^j S_j^i(\omega),$$

где $S_j^i \in \mathbb{SO}(3)$ – матрица вращений, соответствующая элементу алгебры ω^{ij} . Здесь мы используем следующую параметризацию группы трехмерных вращений элементами ее алгебры (1.132),

$$S_i^j = (e^{(\omega\varepsilon)})_i^j = \cos \omega \delta_i^j + \frac{(\omega\varepsilon)_i^j}{\omega} \sin \omega + \frac{\omega_i \omega^j}{\omega^2} (1 - \cos \omega) \in \mathbb{SO}(3), \quad (36.17)$$

где $(\omega\varepsilon)_i^j := \omega^k \varepsilon_{ki}^j$ и $\omega := \sqrt{\omega^i \omega_i}$ – модуль вектора ω^i . Псевдовектор $\omega^k := \omega_{ij} \varepsilon^{ijk}/2$, где ε^{ijk} – полностью антисимметричный тензор третьего ранга, $\varepsilon^{123} = 1$, направлен вдоль оси вращения, а его длина равна углу поворота. Поле $\omega^{ij}(x) \in \mathfrak{so}(3)$ мы будем называть *спиновой структурой* вещества.

Если среда обладает спиновой структурой, то в ней возможны дефекты, которые называются дисклинациями. Для линейных дисклинаций, параллельных оси x^3 , векторное поле n лежит в перпендикулярной плоскости x^1, x^2 . Простейшие примеры линейных дисклинаций показаны на рис. 36.6. Каждая линейная дисклинация характеризуется *вектором Франка* с компонентами

$$\Theta_i := \varepsilon_{ijk} \Omega^{jk}, \quad (36.18)$$

где

$$\Omega^{ij} := \oint_C dx^\mu \partial_\mu \omega^{ij}, \quad (36.19)$$

и интеграл берется по замкнутому контуру C , охватывающему линию дисклинации. Длина вектора Франка равна полному углу поворота поля n^i при обходе вокруг дисклинации.

Векторное поле $n^i(x)$ задает отображение евклидова пространства в сферу $\mathbb{R}^3 \rightarrow \mathbb{S}^2$. Для линейных дисклинаций, параллельных оси x^3 , это отображение сужено до отображения плоскости $\mathbb{R}^2$ в окружность $\mathbb{S}^1$. Ясно, что в этом случае полный угол поворота $|\Theta| := \sqrt{\Theta^i \Theta_i}$ должен быть кратен 2π .

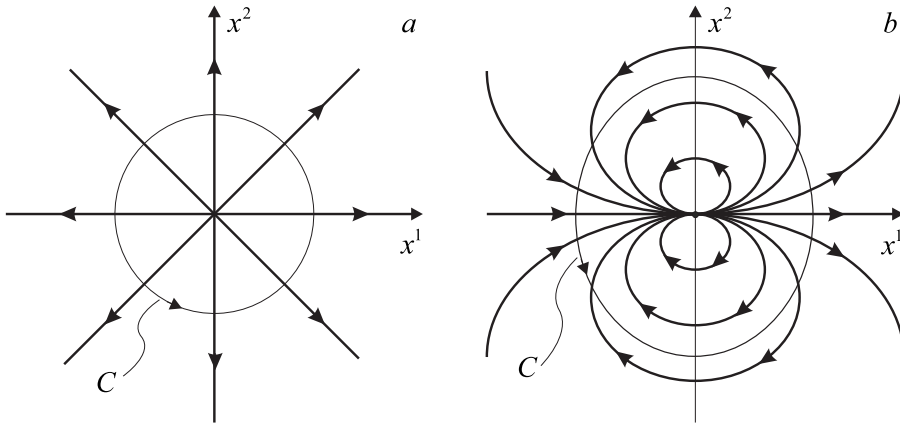


Рис. 36.6: Распределение единичного векторного поля в плоскости x^1, x^2 для линейных дисклиний, параллельных оси x^3 , при $|\Theta| = 2\pi$ (a) и $|\Theta| = 4\pi$ (b).

Для нематических жидких кристаллов у нас есть отношение эквивалентности $n^i \sim -n^i$. Поэтому для линейных дисклиний, параллельных оси x^3 , поле директора задает отображение плоскости в проективную прямую $\mathbb{R}^2 \rightarrow \mathbb{RP}^1$. В этом случае длина вектора Франка должна быть кратна π . Соответствующие примеры простейших дисклиний показаны на рис. 36.7.

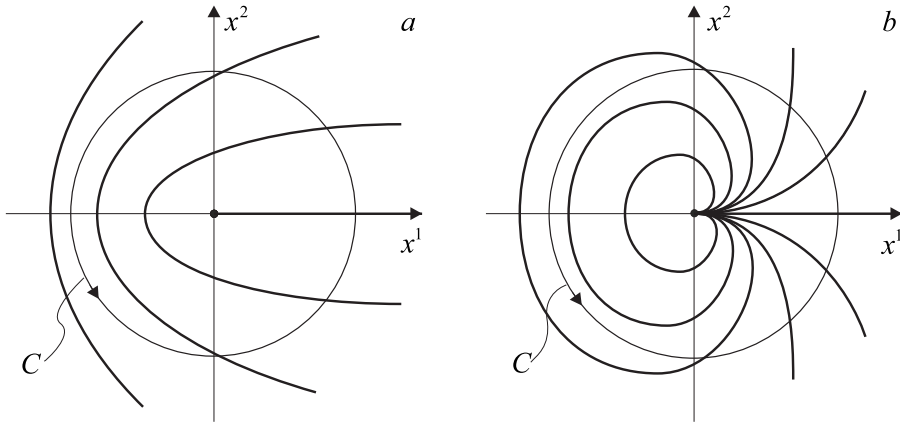


Рис. 36.7: Распределение поля директора в плоскости x^1, x^2 для линейных дисклиний, параллельных оси x^3 , при $|\Theta| = \pi$ (a) и $|\Theta| = 3\pi$ (b).

Как и в случае поля смещений для дислокаций, поле $\omega^{ij}(x)$, принимающее значения в алгебре Ли $\mathfrak{so}(3)$, при наличии дисклиний не является непрерывной функцией в $\mathbb{R}^3$. Сделаем разрез в $\mathbb{R}^3$, ограниченный линией дисклинии. Тогда можно считать поле $\omega^{ij}(x)$ достаточно гладким во всем пространстве за исключением разреза. Предположим, что все частные производные от $\omega^{ij}(x)$ имеют одинаковый предел при приближении к разрезу с обеих сторон. Тогда определим новое поле

$$\omega_\mu^{ij} := \begin{cases} \partial_\mu \omega^{ij}, & \text{вне разреза,} \\ \lim \partial_\mu \omega^{ij}, & \text{на разрезе.} \end{cases} \quad (36.20)$$

По построению функции ω_μ^{ij} являются достаточно гладкими везде, за исключением, возможно, линии дисклинии. Тогда вектор Франка можно представить поверх-

ностным интегралом

$$\Omega^{ij} = \oint_C dx^\mu \omega_\mu^{ij} = \iint_S dx^\mu \wedge dx^\nu (\partial_\mu \omega_\nu^{ij} - \partial_\nu \omega_\mu^{ij}), \quad (36.21)$$

где S – произвольная поверхность с краем C . Если поле ω_μ^{ij} задано, то условием интегрируемости системы уравнений $\partial_\mu \omega_\nu^{ij} = \omega_\mu^{ij}$ является равенство

$$\partial_\mu \omega_\nu^{ij} - \partial_\nu \omega_\mu^{ij} = 0. \quad (36.22)$$

Это равенство дает критерий отсутствия дисклинаций.

В геометрической теории дефектов мы отождествляем поле ω_μ^{ij} с $\mathbb{SO}(3)$ -связностью. В выражении для кривизны (5.57)

$$R_{\mu\nu}^{ij} = \partial_\mu \omega_\nu^{ij} - \partial_\nu \omega_\mu^{ij} - \omega_\mu^{ik} \omega_{\nu k}^j + \omega_\nu^{ik} \omega_{\mu k}^j \quad (36.23)$$

слагаемые с производными совпадают с (36.22) поэтому мы постулируем ковариантный критерий отсутствия дисклинаций в виде равенства нулю тензора кривизны $\mathbb{SO}(3)$ связности:

$$R_{\mu\nu}^{ij} = 0.$$

Одновременно мы даем физическую интерпретацию тензору кривизны как поверхностной плотности вектора Франка:

$$\Omega^{ij} := \iint_S dx^\mu \wedge dx^\nu R_{\mu\nu}^{ij}. \quad (36.24)$$

Это определение приводит к предыдущему выражению для вектора Франка (36.21) в случае, когда вращение вектора n происходит в фиксированной плоскости. В этом случае вращения сужены до подгруппы $\mathbb{SO}(2) \subset \mathbb{SO}(3)$. Поскольку группа вращений плоскости $\mathbb{SO}(2)$ является абелевой, то квадратичные слагаемые в выражении для кривизны (5.57) исчезают и мы получаем прежнее выражение для вектора Франка (36.21). В этом можно убедиться с помощью прямой проверки.

Предложение 36.3.1. *Если вращение директора происходит в фиксированной плоскости, то тензор кривизны линеен по компонентам $\mathbb{SO}(3)$ связности.*

Доказательство. В трехмерном пространстве существует полностью антисимметричный тензор третьего ранга ε_{ijk} , $\varepsilon_{123} := 1$, (см. приложение 37.4). Поэтому $\mathbb{SO}(3)$ связность ω_μ^{ij} можно параметризовать полем $\omega_{\mu k}$, имеющим только два индекса:

$$\omega_{\mu k} := \frac{1}{2} \omega_\mu^{ij} \varepsilon_{ijk} \quad \Leftrightarrow \quad \omega_\mu^{ij} = \varepsilon^{ijk} \omega_{\mu k}. \quad (36.25)$$

Соответствующий тензор кривизны (36.23) имеет вид

$$R_{\mu\nu k} := R_{\mu\nu}^{ij} \varepsilon_{ijk} = \partial_\mu \omega_{\nu k} - \partial_\nu \omega_{\mu k} + 2\omega_{\mu i} \omega_{\nu j} \varepsilon^{ij}_k. \quad (36.26)$$

Если вращение поля директора происходит в одной плоскости, например, x, y , то компоненты спиновой структуры, имеющие индекс вдоль оси z , равны нулю, $\omega^{i3} = -\omega^{3i}$. Поэтому $\omega_{\mu 1} = 0$ и $\omega_{\mu 2} = 0$. Тогда из выражения (36.26) следует, что последнее слагаемое равно нулю. $\square$

В геометрической теории дефектов равенство (36.24) является постулатом, а все предыдущие построения – наводящими соображениями.

Таким образом, мы описали среду с дислокациями (дефекты упругой среды) и дисклинациями (дефекты спиновой структуры) в рамках геометрии Римана–Картана. При этом мы отождествили тензор кручения с поверхностной плотностью дислокаций, а тензор кривизны – с поверхностной плотностью дисклинаций. Соответствие между физическими и геометрическими объектами подытожено в Таблице 36.1.

Упругие деформации	$R_{\mu\nu}{}^{ij} = 0$	$T_{\mu\nu}{}^i = 0$
Дислокации	$R_{\mu\nu}{}^{ij} = 0$	$T_{\mu\nu}{}^i \neq 0$
Дисклинации	$R_{\mu\nu}{}^{ij} \neq 0$	$T_{\mu\nu}{}^i = 0$
Дислокации и дисклинации	$R_{\mu\nu}{}^{ij} \neq 0$	$T_{\mu\nu}{}^i \neq 0$

Таблица 36.1: Соответствие между физическими и геометрическими объектами в геометрической теории дефектов.

36.4 Свободная энергия

До сих пор мы обсуждали только соответствие между физическими и геометрическими объектами. Для завершения построения геометрической теории дефектов необходимо постулировать уравнения равновесия, описывающие статическое распределение дефектов в среде. Единственными и независимыми переменными в геометрическом подходе являются репер $e_\mu{}^i$ и $\mathbb{SO}(3)$ связность $\omega_\mu{}^{ij}$. В отличие от предыдущих геометрических подходов мы полностью отказываемся от поля смещений u^i и спиновой структуры ω^{ij} как полей, входящих в систему уравнений равновесия. В общем случае непрерывного распределения дефектов их просто не существует. Тем не менее на определенном этапе и при определенных условиях их можно восстановить, но об этом пойдет речь в следующих разделах.

Выражение для свободной энергии было получено в статье [294]. Мы предполагаем, что уравнения равновесия должны быть ковариантны относительно общих преобразований координат и локальных $\mathbb{SO}(3)$ вращений, не выше второго порядка и следовать из вариационного принципа. Мы также требуем инвариантности уравнений равновесия относительно пространственных отражений. Тогда выражение для свободной энергии, приводящее к уравнениям равновесия, должно быть равно интегралу по объему от скалярной функции (лагранжиана), квадратичной по тензору кручения и кривизны, которые заданы равенствами (36.14) и (36.23). В трехмерном пространстве существует три независимых инварианта, квадратичных по тензору кручения, и три независимых инварианта, квадратичных по тензору кривизны. Сюда же можно добавить скалярную кривизну и “космологическую” постоянную Λ . Таким образом, мы получаем восьмипараметрический лагранжиан общего вида

$$\begin{aligned} \frac{1}{\sqrt{g}}L = & -\kappa R + \frac{1}{4}T_{ijk}(\beta_1 T^{ijk} + \beta_2 T^{kij} + \beta_3 T^j \delta^{ik}) + \\ & + \frac{1}{4}R_{ijkl}(\gamma_1 R^{ijkl} + \gamma_2 R^{klji} + \gamma_3 R^{ik} \delta^{jl}) - \Lambda, \quad \sqrt{g} = \det e_\mu{}^i, \end{aligned} \quad (36.27)$$

где κ , $\beta_{1,2,3}$ и $\gamma_{1,2,3}$ – некоторые постоянные, введен след тензора кручения $T_j := T_{ij}{}^i$, тензор Риччи $R_{ik} := R_{ijk}{}^j$ и скалярная кривизна $R := R_i{}^i$. Здесь и в дальнейшем переход от греческих индексов к латинским и наоборот осуществляется с помощью репера и его обратного. Например,

$$R_{ijkl} := R_{\mu\nu\kappa\lambda} e^\mu{}_i e^\nu{}_j e^\kappa{}_k e^\lambda{}_l, \quad T_{ijk} := T_{\mu\nu\kappa} e^\mu{}_i e^\nu{}_j e^\kappa{}_k.$$

Спецификой трех измерений является то обстоятельство, что полный тензор кривизны взаимно однозначно определяется своим тензором Риччи (17.24) и имеет три неприводимые компоненты. Поэтому лагранжиан содержит только три независимых инварианта, квадратичных по тензору кривизны. Лагранжиан Гильберта–Эйнштейна

$\tilde{R}$, который также приводит к уравнениям второго порядка для репера, нет никакой необходимости добавлять к свободной энергии (36.27) ввиду тождества (6.96).

Таким образом, наиболее общий лагранжиан зависит от восьми постоянных и приводит к очень сложным уравнениям равновесия. В настоящее время мы не знаем, какие именно значения постоянных описывают те или иные среды. Поэтому, чтобы упростить ситуацию, сделаем физически разумные предположения. А именно, потребуем, чтобы уравнения равновесия допускали следующие три типа решений.

1. Существуют решения, описывающую среду только с дислокациями,
 $R_{\mu\nu}{}^{ij} = 0, T_{\mu\nu}{}^i \neq 0$.
2. Существуют решения, описывающие среду только с дисклинациями,
 $R_{\mu\nu}{}^{ij} \neq 0, T_{\mu\nu}{}^i = 0$.
3. Существуют решения, описывающие среду без дислокаций и дисклинаций,
 $R_{\mu\nu}{}^{ij} = 0, T_{\mu\nu}{}^i = 0$.

Оказывается, что эти простые предположения уменьшают число независимых параметров в выражении для свободной энергии с восьми до двух. Докажем это утверждение. Лагранжиан (36.27) приводит к следующим уравнениям равновесия:

$$\begin{aligned} \frac{1}{\sqrt{g}} \frac{\delta L}{\delta e_\mu{}^i} = & \kappa (Re^\mu{}_i - 2R_i{}^\mu) + \beta_1 \left(\nabla_\nu T^{\nu\mu}{}_i - \frac{1}{4} T_{jkl} T^{jkl} e^\mu{}_i + T^{\mu jk} T_{ijk} \right) + \\ & + \beta_2 \left(-\frac{1}{2} \nabla_\nu (T_i{}^{\mu\nu} - T_i{}^{\nu\mu}) - \frac{1}{4} T_{jkl} T^{ljk} e^\mu{}_i - \frac{1}{2} T^{j\mu k} T_{kij} + \frac{1}{2} T^{jk\mu} T_{kij} \right) + \\ & + \beta_3 \left(-\frac{1}{2} \nabla_\nu (T^\nu e^\mu{}_i - T^\mu e^\nu{}_i) - \frac{1}{4} T_j T^j e^\mu{}_i + \frac{1}{2} T^\mu T_i + \frac{1}{2} T^j T_{ij}{}^\mu \right) + \\ & + \gamma_1 \left(-\frac{1}{4} R_{jklm} R^{jklm} e^\mu{}_i + R^{\mu jkl} R_{ijkl} \right) + \\ & + \gamma_2 \left(-\frac{1}{4} R_{jklm} R^{lmjk} e^\mu{}_i + R^{kl\mu j} R_{ijkl} \right) + \\ & + \gamma_3 \left(-\frac{1}{4} R_{jk} R^{jk} e^\mu{}_i + \frac{1}{2} R^{\mu j} R_{ij} + \frac{1}{2} R^{jk} R_{jik}{}^\mu \right) + \Lambda e^\mu{}_i = 0, \end{aligned} \quad (36.28)$$

$$\begin{aligned} \frac{1}{\sqrt{g}} \frac{\delta L}{\delta \omega_\mu{}^{ij}} = & \kappa \left(\frac{1}{2} T_{ij}{}^\mu + T_i e^\mu{}_j \right) + \beta_1 \frac{1}{2} T^\mu{}_{ji} + \beta_2 \frac{1}{4} (T_i{}^\mu{}_j - T_{ij}{}^\mu) + \\ & + \beta_3 \frac{1}{4} T_j e^\mu{}_i + \gamma_1 \frac{1}{2} \nabla_\nu R^{\nu\mu}{}_{ij} + \gamma_2 \frac{1}{2} \nabla_\nu R_{ij}{}^{\nu\mu} + \\ & + \gamma_3 \frac{1}{4} \nabla_\nu (R^\nu{}_i e^\mu{}_j - R^\mu{}_i e^\nu{}_j) - (i \leftrightarrow j) = 0, \end{aligned} \quad (36.29)$$

где ковариантная производная действует с $\mathbb{SO}(3)$ связностью на латинские индексы и с символами Кристоффеля – на греческие. Например,

$$\begin{aligned} \nabla_\nu T^{\rho\mu}{}_i &:= \partial_\nu T^{\rho\mu}{}_i + \tilde{\Gamma}_{\nu\sigma}{}^\rho T^{\sigma\mu}{}_i + \tilde{\Gamma}_{\nu\sigma}{}^\mu T^{\rho\sigma}{}_i - \omega_{\nu i}{}^j T^{\rho\mu}{}_j, \\ \nabla_\nu R^{\rho\mu}{}_{ij} &:= \partial_\nu R^{\rho\mu}{}_{ij} + \tilde{\Gamma}_{\nu\sigma}{}^\rho R^{\sigma\mu}{}_{ij} + \tilde{\Gamma}_{\nu\sigma}{}^\mu R^{\rho\sigma}{}_{ij} - \omega_{\nu i}{}^k R^{\rho\mu}{}_{kj} - \omega_{\nu j}{}^k R^{\rho\mu}{}_{ik}. \end{aligned}$$

Первое требование к классу решений уравнений равновесия состоит в том, чтобы они допускали решения, описывающие наличие в среде только дислокаций. Это означает наличие решений с нулевым тензором кривизны, что соответствует отсутствию дисклинаций. Подстановка условия $R_{\mu\nu}{}^{ij} = 0$ в уравнение (36.29) для $\mathbb{SO}(3)$

связности дает три независимых уравнения для неприводимых компонент тензора кручения:

$$\begin{aligned}(12\kappa + 2\beta_1 - \beta_2 - 2\beta_3)T_i &= 0, \\ (\kappa - \beta_1 - \beta_2)T^* &= 0, \\ (4\kappa + 2\beta_1 - \beta_2)W_{ijk} &= 0.\end{aligned}\tag{36.30}$$

Здесь T_i (= след), T^* (= полностью антисимметричная часть) и W_{ijk} – неприводимые компоненты тензора кручения:

$$T_{ijk} = W_{ijk} + T^* \epsilon_{ijk} + \frac{1}{2}(\delta_{ik}T_j - \delta_{jk}T_i),$$

где

$$\begin{aligned}T^* &:= \frac{1}{6}T_{ijk}\epsilon^{ijk}, & T_j &:= T_{ij}{}^i, \\ W_{ijk} &:= T_{ijk} - T^* \epsilon_{ijk} - \frac{1}{2}(\delta_{ik}T_j - \delta_{jk}T_i), & W_{ijk}\epsilon^{ijk} &= 0, & W_{ij}{}^i &= 0.\end{aligned}$$

В общем случае при наличии дислокаций все неприводимые компоненты тензора кручения отличны от нуля ($T_i, T^*, W_{ijk} \neq 0$) и уравнения (36.30) имеют единственное решение

$$\beta_1 = -\kappa, \quad \beta_2 = 2\kappa, \quad \beta_3 = 4\kappa.\tag{36.31}$$

При этих константах связи первые четыре слагаемые в лагранжиане (36.27) с точностью до дивергенции равны лагранжиану Гильберта-Эйнштейна $\kappa\tilde{R}(e)$ как следствие тождества (6.96). При этом уравнение (36.28) сводится к уравнениям Эйнштейна с космологической постоянной

$$\tilde{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\tilde{R} - \frac{\Lambda}{2\kappa}g_{\mu\nu} = 0.\tag{36.32}$$

Таким образом, первое условие выполнено.

Согласно второму условию уравнения равновесия должны допускать решения с нулевым кручением $T_{\mu\nu}{}^i = 0$. В этом случае тензор кривизны обладает дополнительной симметрией относительно перестановки первой и второй пары индексов, $R_{ijkl} = R_{klij}$, и уравнение (36.29) принимает вид

$$\begin{aligned}(\gamma_1 + \gamma_2 + \frac{1}{4}\gamma_3)\nabla_\nu (R^{s\nu}{}_i e^\mu{}_j - R^{s\mu}{}_i e^\nu{}_j - R^{s\nu}{}_j e^\mu{}_i + R^{s\mu}{}_j e^\nu{}_i) + \\ + \frac{1}{6}(\gamma_1 + \gamma_2 + 4\gamma_3)(e^\nu{}_i e^\mu{}_j - e^\mu{}_i e^\nu{}_j)\nabla_\nu R = 0.\end{aligned}\tag{36.33}$$

Здесь мы разложили тензор Риччи на неприводимые компоненты, выделив след и антисимметричную часть:

$$R_{ij} = R^s{}_{ij} + R^A{}_{ij} + \frac{1}{3}R\delta_{ij},$$

где

$$R^s{}_{ij} := \frac{1}{2}(R_{ji} + R_{ij}) - \frac{1}{3}\delta_{ij}R, \quad R^A{}_{ij} := \frac{1}{2}(R_{ij} - R_{ji}).$$

Напомним, что при нулевом кручении тензор Риччи симметричен: $R^A{}_{ij} = 0$. Свертка уравнения (36.33) с $e_\mu{}^j$ приводит к уравнению

$$(\gamma_1 + \gamma_2 + \frac{1}{4}\gamma_3)\nabla_\nu R^{s\nu}{}_\mu + \frac{1}{3}(\gamma_1 + \gamma_2 + 4\gamma_3)\nabla_\mu R = 0.$$

В общем случае отличной от нуля кривизны ковариантные производные $\nabla_\nu R^{s\nu}_\mu$ и $\nabla_\mu R$ отличны от нуля и независимы. Поэтому мы получаем два уравнения на константы связи:

$$\gamma_1 + \gamma_2 + \frac{1}{4}\gamma_3 = 0, \quad \gamma_1 + \gamma_2 + 4\gamma_3 = 0,$$

которые имеют единственное решение

$$\gamma_1 = -\gamma_2 = \gamma, \quad \gamma_3 = 0. \quad (36.34)$$

В этом случае уравнение (36.28) для репера при ненулевом кручении также сводится к уравнениям Эйнштейна (36.32).

Последнее требование наличия решений с нулевой кривизной и кручением выполняется только при нулевой космологической постоянной $\Lambda = 0$.

Таким образом, простые и физически осмысленные требования определяют двухпараметрический лагранжиан

$$\frac{1}{\sqrt{g}}L = -\kappa\tilde{R} + 2\gamma R^A_{ij}R^{Aij}, \quad (36.35)$$

который равен сумме лагранжиана Гильберта–Эйнштейна для репера и квадрата антисимметричной компоненты тензора Риччи. Отметим, что $\tilde{R}(e)$ и $R^A_{ij}(e, \omega)$ построены для различных метрических связностей.

Действие (свободная энергия) в геометрической теории дефектов имеет вид

$$S := \int dx \sqrt{g} (-\kappa\tilde{R} + 2\gamma R^A_{ij}R^{Aij}). \quad (36.36)$$

Конечно, при решении конкретных задач к этому действию необходимо добавить слагаемые с соответствующими источниками.

Заметим, что если отказаться от требования инвариантности свободной энергии относительно пространственных отражений, то возникают дополнительные инварианты, например, $R_{ij}T_k\epsilon^{ijk}$. Эта возможность требует отдельного исследования.

Выражение (36.35) в геометрической теории дефектов определяет плотность свободной энергии и приводит к уравнениям равновесия (уравнения Эйлера–Лагранжа). В геометрическом подходе вектор смещения и спиновая структура вообще не входят в выражение для свободной энергии (36.35).

36.5 Фиксирование калибровки

В геометрическом подходе единственными переменными являются репер e_μ^i и $\mathbb{SO}(3)$ связность ω_μ^j . Поле смещений u^i и спиновую структуру ω_i^j можно ввести только в тех областях среды, где дефекты отсутствуют. Действительно, из условия отсутствия дисклинаций ($R_{\mu\nu}^i = 0$) следует, что $\mathbb{SO}(3)$ связность является чистой калибровкой (12.2), т.е. существует спиновая структура ω_i^j . Если, вдобавок, отсутствуют дислокации ($T_{\mu\nu}^i = 0$), то существует поле смещений $u^i(x)$ такое, что репер равен частным производным, $e_\mu^i = \partial_\mu u^i$. В этом и только в этом случае можно ввести поле смещений и спиновую структуру. Ниже мы покажем, что это можно сделать таким образом, чтобы были выполнены уравнения нелинейной теории упругости и уравнения главного кирального $\mathbb{SO}(3)$ поля.

Уравнения Эйлера–Лагранжа для свободной энергии (36.35) ковариантны относительно общих преобразований координат в $\mathbb{R}^3$ и локальных $\mathbb{SO}(3)$ вращений. Это

значит, что любое решение уравнений равновесия определено с точностью до диффеоморфизмов и локальных вращений. Чтобы геометрическая теория дефектов имела предсказательную силу, необходимо зафиксировать систему координат (выбрать калибровку, как принято говорить в теории калибровочных полей). Это позволит из каждого класса эквивалентных решений выбрать по одному представителю. После этого мы скажем, что данное решение уравнений Эйлера–Лагранжа описывает распределение дефектов в среде в лабораторной системе координат.

Начнем с фиксирования диффеоморфизмов. Для этой цели выберем упругую калибровку, которая была предложена в [295]. Данный вопрос имеет принципиальное значение, и мы остановимся на нем подробно.

Из уравнений линейной теории упругости в отсутствие неупругих сил, $f^i = 0$, следует нековариантное уравнение второго порядка на вектор смещения (36.5). Его можно переписать в терминах индуцированной метрики (36.6), для которой мы получим нековариантное уравнение первого порядка. Это уравнение мы и выберем в качестве калибровочного условия, фиксирующего диффеоморфизмы. Отметим, что калибровочное условие не определено однозначно, т.к. индуцированная метрика нелинейна по вектору смещений, и различные уравнения для метрики могут иметь одно и то же линейное приближение. Приведем, например, два возможных выбора:

$$g^{\mu\nu} \overset{\circ}{\nabla}_\mu g_{\nu\rho} + \frac{\sigma}{1-2\sigma} g^{\mu\nu} \overset{\circ}{\nabla}_\rho g_{\mu\nu} = 0, \quad (36.37)$$

$$\overset{\circ}{g}^{\mu\nu} \overset{\circ}{\nabla}_\mu g_{\nu\rho} + \frac{\sigma}{1-2\sigma} \overset{\circ}{\nabla}_\rho g^T = 0, \quad (36.38)$$

где $\overset{\circ}{g}_{\mu\nu}$ – евклидова метрика в той системе координат, в которой проводится решение задачи, и введено обозначение для следа метрики $g^T := \overset{\circ}{g}^{\mu\nu} g_{\mu\nu}$.

Калибровочные условия (36.37), (36.38), несмотря на их формальный вид, нарушают общую ковариантность, поскольку понимаются следующим образом. Метрика $\overset{\circ}{g}_{\mu\nu}$ – евклидова метрика, записанная в произвольной системе координат, например, в цилиндрической или сферической. Ковариантная производная $\overset{\circ}{\nabla}_\mu$ построена по символам Кристоффеля, соответствующим метрике $\overset{\circ}{g}_{\mu\nu}$, и, следовательно, $\overset{\circ}{\nabla}_\mu \overset{\circ}{g}_{\nu\rho} = 0$. Метрика $g_{\mu\nu}$ – это метрика, описывающая дислокации (точное решение уравнений равновесия для свободной энергии (36.35)). Если мы хотим записать калибровочное условие, например, для декартовой системы координат, то выбираем $\overset{\circ}{g}^{\mu\nu} = \delta^{\mu\nu}$. Тогда калибровочное условие (36.38) примет вид

$$\delta^{\mu\nu} \partial_\mu g_{\nu\rho} + \frac{\sigma}{1-2\sigma} \partial_\rho (\delta^{\mu\nu} g_{\mu\nu}) = 0. \quad (36.39)$$

Если сюда подставить линейное приближение для метрики (36.6), то получим в точности уравнение (36.5). Уравнение (36.39) явно нарушает общую ковариантность, что и требуется от калибровочного условия, фиксирующего систему координат.

Калибровочные условия (36.37) и (36.38) отличаются тем, что в первом случае производится свертка с метрикой дислокации $g^{\mu\nu}$, а во втором случае – с евклидовой метрикой $\overset{\circ}{g}^{\mu\nu}$, что не меняет линейного приближения. Оба калибровочных условия в линейном приближении по вектору смещений (36.6) приводят к уравнению (36.5), что уже было продемонстрировано в декартовой системе координат.

С геометрической точки зрения мы имеем следующее. Среда с дислокациями диффеоморфна евклидову пространству $\mathbb{R}^3$, снабженному двумя метриками $\overset{\circ}{g}_{\mu\nu}$ и

$g_{\mu\nu}$. Метрика $\overset{\circ}{g}_{\mu\nu}$ является плоской евклидовой метрикой записанной в произвольной системе координат. Метрика $g_{\mu\nu}$ не является плоской и описывает распределение дислокаций в той же системе координат. На самом деле метрика $\overset{\circ}{g}_{\mu\nu}$ нужна только для фиксирования системы координат, в которой измеряется метрика $g_{\mu\nu}$.

Если решение уравнений равновесия удовлетворяет одному из условий (36.37) или (36.38), записанному, например, в цилиндрической системе координат, то будем говорить, что найдено решение в цилиндрической системе координат. При этом мы предполагаем, что распределение дислокаций упругой среды в лабораторной цилиндрической системе координат описывается именно этим решением. Аналогично можно искать решения в декартовой, сферической или другой системе координат.

Калибровочные условия можно записать и для репера e_μ^i , который определен уравнением $g_{\mu\nu} := e_\mu^i e_\nu^j \delta_{ij}$. Здесь возникает дополнительный произвол, т.к. репер определен с точностью до локальных вращений. Эта инвариантность приводит к тому, что для репера возможны различные линейные приближения по вектору смещений. Рассмотрим две возможности в декартовой системе координат:

$$e_{\mu i} \approx \delta_{\mu i} - \partial_\mu u_i, \quad (36.40)$$

$$e_{\mu i} \approx \delta_{\mu i} - \frac{1}{2}(\partial_\mu u_i + \partial_i u_\mu), \quad (36.41)$$

где индекс опущен с помощью символа Кронекера. Этим возможностям и условию (36.38) на метрику соответствует два калибровочных условия на репер:

$$\overset{\circ}{g}^{\mu\nu} \overset{\circ}{\nabla}_\mu e_{\nu i} + \frac{1}{1-2\sigma} \overset{\circ}{e}^\mu_i \overset{\circ}{\nabla}_\mu e^\tau = 0, \quad (36.42)$$

$$\overset{\circ}{g}^{\mu\nu} \overset{\circ}{\nabla}_\mu e_{\nu i} + \frac{\sigma}{1-2\sigma} \overset{\circ}{e}^\mu_i \overset{\circ}{\nabla}_\mu e^\tau = 0, \quad (36.43)$$

где $e^\tau := \overset{\circ}{e}^\mu_i e_\mu^i$ – след репера. Приведенные условия отличаются коэффициентом перед вторым слагаемым. Отметим, что в криволинейной системе координат в ковариантную производную $\overset{\circ}{\nabla}_\mu$ необходимо включить также плоскую $\mathbb{SO}(3)$ связность, действующую на индексы i, j . Можно выписать также и другие возможные калибровочные условия, которые имеют то же линейное приближение. Вопрос о правильном выборе калибровки в настоящее время открыт. На данном этапе мы хотим только продемонстрировать, что систему координат необходимо зафиксировать, и что калибровочное условие зависит от коэффициента Пуассона, который является экспериментально наблюдаемой величиной.

Калибровочные условия (36.42), (36.43) представляют собой дифференциальные уравнения первого порядка и содержат определенный произвол в решениях. Поэтому для однозначного фиксирования решения необходимо дополнительно наложить граничные условия на репер для каждой конкретной задачи.

Если дефектов нет, то $T_{\mu\nu}^i = 0$, $R_{\mu\nu j}^i = 0$, и уравнения равновесия выполняются, т.к. эти условия удовлетворяют уравнениям Эйлера–Лагранжа для лагранжиана (36.35). В этом и только в этом случае можно ввести вектор смещения, и для него упругая калибровка сведется к уравнениям нелинейной теории упругости. При наличии дефектов поля смещений не существует, и упругая калибровка просто определяет репер.

При выборе функционала свободной энергии мы потребовали, чтобы условия $R_{\mu\nu}^{ij} = 0$ и $T_{\mu\nu}^i = 0$ удовлетворяли уравнениям Эйлера–Лагранжа. И это важно, т.к. в противном случае на вектор смещения появилось бы дополнительное условие

(уравнения Эйлера–Лагранжа) помимо уравнений теории упругости, следующих из упругой калибровки.

Подчеркнем еще раз важное обстоятельство. В геометрической теории дефектов мы предполагаем, что существует выделенная лабораторная система координат, в которой проводятся измерения. Эта система координат связана со средой без дефектов и упругих напряжений и соответствует плоскому евклидову пространству $\mathbb{R}^3$. Калибровочные условия (36.37), (36.38) и (36.42), (36.43) записаны именно в этом евклидовом пространстве $\mathbb{R}^3$, и содержат измеряемую величину – коэффициент Пуассона σ . Это обстоятельство существенно отличает геометрическую теорию дефектов от моделей гравитации, в которых все системы координат считаются равноправными.

Упругая калибровка есть уравнения нелинейной теории упругости, в которой тензор деформаций предполагается определенным через индуцированную метрику (36.9), а закон Гука сохраняется линейным. Обобщение на случай нелинейной зависимости тензора деформаций от напряжений очевидно.

Упругая калибровка используется для фиксирования диффеоморфизмов. Поскольку выражение для свободной энергии (36.35) инвариантно также относительно локальных $\mathbb{SO}(3)$ вращений, то их тоже необходимо фиксировать. С этой целью была предложена лоренцева калибровка для связности [296]

$$\delta^{\mu\nu} \partial_\mu \omega_{\nu j}{}^i = 0. \quad (36.44)$$

Калибровка (36.44) записана в лабораторной декартовой системе координат и имеет глубокий физический смысл. А именно, допустим, что дисклинаций нет ($R_{\mu\nu j}{}^i = 0$). Тогда $\mathbb{SO}(3)$ связность представляет собой чистую калибровку

$$\omega_{\mu j}{}^i = \partial_\mu S^{-1}{}_j{}^k S_k{}^i, \quad S_j{}^i \in \mathbb{SO}(3).$$

В этом случае лоренцева калибровка сведется к уравнениям главного кирального $\mathbb{SO}(3)$ поля

$$\delta^{\mu\nu} \partial_\mu (\partial_\nu S^{-1}{}_j{}^k S_k{}^i) = 0 \quad (36.45)$$

для спиновой структуры $\omega^{ij}(x)$, через которую выражается матрица вращений с помощью формулы (36.17). В этом случае уравнения (36.45) представляют собой систему нелинейных уравнений в частных производных второго порядка для спиновой структуры.

Замечание. Модели главного кирального поля (см., например, [299, 300, 301, 302, 303]), для различных групп и в разном числе измерений привлекают большое внимание в математической физике, поскольку допускают решения в виде топологических солитонов и находят широкое применение в физике. $\square$

Таким образом, лоренцева калибровка (36.44) означает следующее. При отсутствии дисклинаций уравнения равновесия тождественно удовлетворяются и существует поле спиновой структуры ω^{ij} , которое удовлетворяет уравнениям главного кирального поля. Тем самым мы предполагаем, что спиновая структура среды в отсутствие дефектов описывается моделью главного кирального $\mathbb{SO}(3)$ поля.

Модель главного кирального поля не является единственной, которую можно использовать для фиксирования локальных вращений. Для этой же цели можно использовать модель Скирма [304]. Уравнения Эйлера–Лагранжа этой модели нетрудно переписать в терминах $\mathbb{SO}(3)$ связности и использовать их в качестве калибровочного условия.

Существуют и другие модели для спиновых структур. Распределение магнитных моментов в ферромагнетиках и поле директора в жидких кристаллах описываются с помощью выражения для свободной энергии, зависящего непосредственно от векторного n поля [305, 297]. В последнее время большое внимание привлекает также модель Фаддеева для n поля [138]. Открытым остается вопрос о том, существуют ли такие калибровочные условия на $\mathbb{SO}(3)$ связность, которые в отсутствие дисклинаций приводили бы к этим моделям.

Таким образом, в геометрической теории дефектов мы ставим следующую задачу. Найти решение уравнений Эйлера–Лагранжа для свободной энергии (36.35), которое удовлетворяет упругой калибровке для репера и лоренцевой калибровке для $\mathbb{SO}(3)$ связности. В разделах (36.7.2) и (36.8.2) мы решим задачу для клиновой дислокации в рамках классической теории упругости и геометрической теории дефектов, а затем сравним полученные результаты.

36.6 Асимметричная теория упругости

В предыдущем разделе мы использовали теорию упругости и модель главного кривого $\mathbb{SO}(3)$ поля для фиксирования инвариантности свободной энергии (36.35) в геометрической теории дефектов. Это – не единственная возможность, т.к. для фиксирования калибровки можно применить и другие модели. В настоящем разделе мы покажем, как для фиксирования инвариантности относительно диффеоморфизмов и локальных вращений может быть использована другая модель – асимметричная теория упругости.

В начале прошлого века братья Коссера разработали теорию упругой среды, каждая точка которой характеризуется не только своим положением, но и ориентацией в пространстве [306]. То есть в каждой точке среды задан репер (рис. 36.8). С фи-

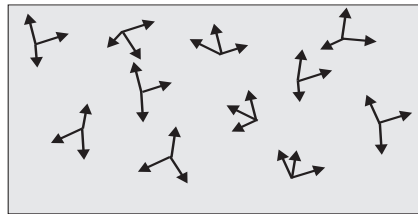


Рис. 36.8: Каждая точка среды Коссера характеризуется не только своим положением, но и ориентацией в пространстве.

зической точки зрения это означает, что каждый атом в кристаллической решетке является не точкой, а протяженным объектом, имеющим ориентацию. В этом случае тензор напряжений σ^{ij} уже не будет симметричным, и соответствующая теория называется асимметричной теорией упругости. Современное изложение этого подхода приведено в [307]. Покажем, что асимметричная теория упругости естественным образом вкладывается в геометрическую теорию дефектов.

Основными переменными в асимметричной теории упругости являются вектор смещений $u^i(x)$ и угол поворота $\omega^i(x)$. Направление псевдовектора ω^i совпадает с осью вращения элемента среды, а его длина равна углу поворота. Угол поворота дуален к полю спиновой структуры, $\omega_{ij} = \varepsilon_{ijk}\omega^k$, рассмотренной в разделе 36.3.

Среда Коссера характеризуется тензором напряжений $\sigma^{ij}(x)$ (плотность упругих сил, действующих на площадку с нормалью i в направлении j) и тензором мо-

ментных напряжений $\mu^{ij}(x)$ (плотность момента сил, действующих на площадку с нормалью i в направлении j). Среда Коссера находится в равновесии, если в каждой точке уравновешены силы и моменты сил [307]:

$$\partial_j \sigma^{ji} + f^i = 0, \quad (36.46)$$

$$\varepsilon^{ijk} \sigma_{jk} + \partial_j \mu^{ji} + m^i = 0, \quad (36.47)$$

где $f^i(x)$ и $m^i(x)$ – плотности неупругих внешних сил и моментов. Из уравнения (36.47) следует, что тензор упругих напряжений симметричен тогда и только тогда, когда выполнено условие $\partial_j \mu^{ji} + m^i = 0$.

Поле смещений и угол поворота однозначно определяют тензор деформаций $\epsilon_{ij}(x)$ и тензор изгиба-кручения $\kappa_{ij}(x)$

$$\begin{aligned} \epsilon_{ij} &:= \partial_i u_j - \omega_{ij}, \\ \kappa_{ij} &:= \partial_i \omega_j, \quad \omega_{ij} := \varepsilon_{ijk} \omega^k. \end{aligned} \quad (36.48)$$

В общем случае тензоры деформаций и изгиба-кручения никакой симметрии по индексам не имеют.

Закон Гука в среде Коссера заменяется двумя линейными соотношениями, связывающими тензоры напряжений и моментных напряжений с тензорами деформаций и изгиба-кручения:

$$\sigma_{ij} = 2\mu \epsilon_{\{ij\}} + 2\alpha \epsilon_{[ij]} + \lambda \delta_{ij} \epsilon_k^k, \quad (36.49)$$

$$\mu_{ij} = 2\gamma \kappa_{\{ij\}} + 2\epsilon \kappa_{[ij]} + \beta \delta_{ij} \kappa_k^k, \quad (36.50)$$

где μ, λ – постоянные Ламе, и $\alpha, \beta, \gamma, \epsilon$ – четыре новые упругие постоянные, характеризующие среду. Фигурные и квадратные скобки означают, соответственно, симметризацию и антисимметризацию индексов.

Асимметричная теория упругости сводится к симметричной теории упругости, рассмотренной в разделе 36.1, в случае, когда

$$\omega_{ij} = \frac{1}{2}(\partial_i u_j - \partial_j u_i). \quad (36.51)$$

Тогда тензор деформаций (36.48) симметричен и имеет тот же вид (36.4), что и ранее. При этом уравнение (36.49) переходит в закон Гука (36.3), а уравнение (36.46) – в закон Ньютона (36.2). Уравнение (36.47) вместе с (36.49) и (36.50) сводится к равенству

$$(\gamma + \epsilon) \varepsilon^{ijk} \Delta \partial_j u_k + m^i = 0.$$

Первый член равен нулю, как следствие уравнения (36.5). Таким образом, для спиновой структуры вида (36.51) и $m^i = 0$ мы возвращаемся к симметричной теории упругости.

Уравнения (36.46), (36.47), (36.49) и (36.50) вместе с граничными условиями определяют равновесное состояние среды Коссера. Покажем, как эта модель вкладывается в геометрическую теорию упругости. Прежде всего заметим, что в отсутствие дефектов ($T_{\mu\nu}^i = 0$, $R_{\mu\nu}^{ij} = 0$) существуют поля u^i и ω^{ij} . Тогда репер и $\mathbb{SO}(3)$ связность в линейном приближении определяются тензорами деформаций и изгиба-кручения:

$$e_\mu^i = \partial_\mu y^j S_j^i(\omega) \approx (\delta_\mu^j - \partial_\mu u^j)(\delta_j^i + \omega_j^i) \approx \delta_\mu^i - \epsilon_\mu^i, \quad (36.52)$$

$$\omega_\mu^{ij} \approx \partial_\mu \omega^{ij} = \varepsilon^{ijk} \kappa_{\mu k}. \quad (36.53)$$

Отметим, что соотношения (36.48) можно рассматривать как уравнения относительно вектора смещений и угла поворота при заданных $\epsilon_{ij}(x)$ и $\kappa_{ij}(x)$. Соответствующие условия интегрируемости были получены в [308]. Эти условия интегрируемости являются линейным приближением равенств $T_{\mu\nu}^i = 0$ и $R_{\mu\nu}^{ij} = 0$, определяющих отсутствие дефектов.

Если неупругие силы и моменты в среде отсутствуют ($f^i = 0$, $m^i = 0$), то асимметричная теория упругости сводится к уравнениям второго порядка на вектор смещения и поворота:

$$(\mu + \alpha)\Delta u^i + (\mu - \alpha + \lambda)\partial^i \partial_j u^j - 2\alpha \partial_j \omega^{ji} = 0, \quad (36.54)$$

$$(\gamma + \epsilon)\Delta \omega^i + (\gamma - \epsilon + \beta)\partial^i \partial_j \omega^j + 2\alpha \varepsilon^{ijk}(\partial_j u_k - \omega_{jk}) = 0. \quad (36.55)$$

Перепишем эти уравнения для репера и $\mathbb{SO}(3)$ связности:

$$(\mu + \alpha)\overset{\circ}{\nabla}^\mu e_\mu^i + (\mu - \alpha + \lambda)\overset{\circ}{\nabla}^i e^T - (\mu - \alpha)\omega_\mu^{\mu i} = 0, \quad (36.56)$$

$$\frac{1}{2}(\gamma + \epsilon)\varepsilon^{ijk}\overset{\circ}{\nabla}^\mu \omega_{\mu jk} + \frac{1}{2}(\gamma - \epsilon + \beta)\varepsilon^{\mu jk}\overset{\circ}{\nabla}^i \omega_{\mu jk} + 2\alpha \varepsilon^{i\mu j} e_{\mu j} = 0. \quad (36.57)$$

Конечно, это не единственные уравнения, которые в линейном приближении совпадают с уравнениями (36.54), (36.55). В настоящее время у нас нет аргументов для однозначного выбора. Полученные нелинейные уравнения асимметричной теории упругости можно использовать в качестве калибровочных условий в геометрической теории дефектов. При этом мы имеем шесть уравнений для фиксирования диффеоморфизмов (три параметра) и локальных $\mathbb{SO}(3)$ вращений (три параметра). Таким образом, асимметричная теория упругости естественным образом вкладывается в геометрическую теорию дефектов.

В разделе 36.5 рассмотрены упругая калибровка для репера и лоренцева калибровка для $\mathbb{SO}(3)$ связности. В этом случае при отсутствии дефектов спинные переменные среды не взаимодействуют с упругими напряжениями. В асимметричной теории упругости упругие напряжения непосредственно влияют на спиновую структуру, и наоборот.

36.7 Дислокации в теории упругости

В настоящем разделе мы опишем несколько простейших дислокаций в рамках теории упругости. Это сделано для того, чтобы в дальнейшем сравнить полученные ответы с результатами геометрической теории дефектов.

36.7.1 Винтовая дислокация

Винтовая дислокация изображена на рис.36.2b. В этом случае вектор Бюргера параллелен оси $z := x^3$ и направлен в ту же сторону $\mathbf{b} = (0, 0, b)$, $b > 0$. Исходя из симметрии задачи, будем решать задачу в цилиндрической системе координат r, φ, z . Решение ищем в виде

$$\{u_i\} = \{u_{\hat{r}} = 0, u_{\hat{\varphi}} = 0, u_{\hat{z}} = c\varphi\}, \quad c = \text{const}. \quad (36.58)$$

Линейная зависимость компоненты вектора смещений $u_{\hat{z}}$ от угла вытекает из вращательной симметрии задачи. Для винтовой дислокации необходимо решить уравнения равновесия (36.5) вместе с граничными условиями:

$$u_{\hat{z}}|_{\varphi=0} = 0, \quad u_{\hat{z}}|_{\varphi=2\pi} = b, \quad (36.59)$$

которые означают, что нижний берег разреза сдвинут вдоль оси z на вектор Бюргерса b .

Используя явные выражения для лапласиана и дивергенции в цилиндрических координатах (7.26)–(7.31), получаем, что для векторного поля смещений (36.58) справедливы равенства:

$$\Delta u_{\hat{r}} = 0, \quad \Delta u_{\hat{\varphi}} = 0, \quad \Delta u_{\hat{z}} = 0, \quad \nabla_i u^i = 0.$$

Поэтому уравнения равновесия выполнены.

Постоянная c находится из граничных условий (36.59): $c = b/2\pi$. Таким образом, векторное поле смещений для винтовой дислокации имеет только одну нетривиальную компоненту вдоль оси z :

$$u^{\hat{z}} = u_{\hat{z}} = \frac{b}{2\pi} \varphi. \quad (36.60)$$

Как видим, результат не зависит от упругих свойств среды.

Теперь вычислим индуцированную метрику в среде с винтовой дислокацией. Пусть

$$\{x^\mu\} = (r, \varphi, z) \in \mathbb{R}^3$$

– цилиндрическая система координат в среде после создания дислокации. Из определения вектора смещений (36.1) следует, что до создания дислокации каждая точка среды имела декартовы координаты

$$\{y^i\} = (x, y, z - b\varphi/(2\pi)) \in \mathbb{R}^3,$$

где

$$x := r \cos \varphi, \quad y := r \sin \varphi.$$

Теперь вычислим репер по формуле (36.11)

$$e_\mu^i := \partial_\mu y^i = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -r \sin \varphi & r \cos \varphi & -\frac{b}{2\pi} \\ 0 & 0 & 1 \end{pmatrix}. \quad (36.61)$$

Определитель репера равен $\det e_\mu^i = r$ как и в плоском случае.

Заметим, что вектор смещений (36.60) имеет скачок при $\varphi = 2\pi$. При его дифференцировании по углу φ возникает δ -функция, которую мы отбросили в соответствии с общим рецептом раздела 36.2. В результате получили репер, который является гладким всюду за исключением начала координат, где он вырожден.

Репер (36.61) определяет метрику винтовой дислокации

$$g_{\mu\nu} := e_\mu^i e_\nu^j \delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 + \frac{b^2}{4\pi^2} & -\frac{b}{2\pi} \\ 0 & -\frac{b}{2\pi} & 1 \end{pmatrix}. \quad (36.62)$$

Ее определитель равен $\det g_{\mu\nu} = r^2$, как и следовало ожидать.

Теперь вычислим компоненты тензора кручения. Здесь надо проявить осторожность, т.к. необходимо уловить δ -функцию на оси $z := x^3$. Проведем вычисления в декартовых координатах, потому что цилиндрические координаты не подходят, поскольку они вырождены на оси z .

Пусть $\mathbf{x} := (x, y) = (r, \varphi)$ – двумерный вектор в плоскости, перпендикулярной оси дислокации. Напомним, что фундаментальное решение G (функция Грина) уравнения Лапласа на плоскости

$$\Delta G = b\delta(\mathbf{x}),$$

где $\Delta := \partial_{xx}^2 + \partial_{yy}^2$ и $\delta(\mathbf{x})$ – двумерная δ -функция, имеет вид (см., например, [3])

$$G = \frac{b}{2\pi} \ln \frac{1}{r}. \quad (36.63)$$

Его частные производные имеют вид

$$G_x = -\frac{b}{2\pi} \frac{x}{r^2}, \quad G_y = -\frac{b}{2\pi} \frac{y}{r^2}.$$

Репер для винтовой дислокации в декартовой системе координат $\{x^\mu\} = (x, y, z)$ можно выразить через функцию Грина

$$e_\mu^i = \partial_\mu y^i = \begin{pmatrix} 1 & 0 & \frac{b}{2\pi} \frac{y}{r^2} \\ 0 & 1 & -\frac{b}{2\pi} \frac{x}{r^2} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -G_y \\ 0 & 1 & G_x \\ 0 & 0 & 1 \end{pmatrix}. \quad (36.64)$$

При нулевой $\mathbb{SO}(3)$ связности, $\omega_\mu^{ij} = 0$, что соответствует отсутствию дисклинаций, $R_{\mu\nu}^{ij} = 0$, тензор кручения (36.14) имеет вид

$$T_{\mu\nu}^i = \partial_\mu e_\nu^i - \partial_\nu e_\mu^i.$$

Прямые вычисления показывают, что для репера (36.64) все компоненты кручения равны нулю за исключением одной:

$$T_{xy}^z = G_{xx} + G_{yy} = \Delta G = b\delta(\mathbf{x}). \quad (36.65)$$

Его интегрирование по плоскости x, y дает вектор Бюргерса

$$b^{\hat{z}} = \int dx dy T_{xy}^{\hat{z}} = b,$$

что согласуется с общим определением (36.15).

Таким образом, наличие винтовой дислокации в упругой среде приводит к возникновению нетривиального репера (36.61), который, в свою очередь, порождает нетривиальное кручение (36.65). При этом мы предполагаем, что дисклинации отсутствуют, т.е. $R_{\mu\nu}^{ij} = 0$. Конечно, репер приводит к возникновению нетривиальной метрики (36.62). По этой метрике можно вычислить символы Кристоффеля $\tilde{\Gamma}_{\mu\nu}^\rho$ и соответствующий тензор кривизны $\tilde{R}_{\mu\nu\rho}^\sigma$ (в предположении нулевого кручения), который будет отличен от нуля. Однако это не тот тензор кривизны, который описывает дисклинации. Поэтому мы поместили его знаком тильды.

36.7.2 Клиновая дислокация

Под клиновой дислокацией мы понимаем бесконечную упругую среду, которая топологически совпадает с евклидовым пространством $\mathbb{R}^3$ и построена следующим образом. Берется бесконечная упругая среда без дефектов из которой вырезается бесконечный клин с углом $-2\pi\theta$, $\theta < 0$. Для определенности мы считаем, что острие

клина совпадает с осью $z := x^3$ – осью (ядром) дислокации (рис. 36.9). Затем края разреза симметрично сдвигаются и склеиваются. После этого среда под действием упругих сил приходит в равновесное состояние. Если клин из среды вырезается, то угол дефицита будем считать отрицательным: $-1 < \theta < 0$. При положительных θ клин вставляется. Таким образом, первоначально упругая среда занимает область, большую или меньшую евклидова пространства $\mathbb{R}^3$ в зависимости от знака угла дефицита θ , которая в цилиндрических координатах r, φ, z задается неравенствами:

$$0 < r < \infty, \quad 0 < \varphi < 2\pi\alpha, \quad -\infty < z < \infty, \quad \alpha := 1 + \theta. \quad (36.66)$$

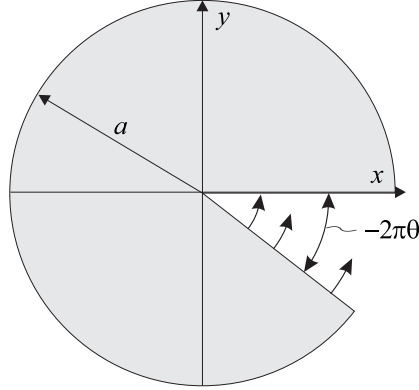


Рис. 36.9: Клиновая дислокация с углом дефицита $2\pi\theta$. При отрицательных и положительных θ клин, соответственно, вырезается и вставляется.

Отметим, что клиновую дислокацию часто называют дисклинацией. В нашем подходе такое название представляется неестественным, поскольку клиновая дислокация связана с нетривиальным кручением. Вдобавок, термин дисклинация используется для дефектов спиновой структуры.

Перейдем к математической постановке задачи для клиновой дислокации в рамках теории упругости. Чтобы избежать расходящихся выражений, которые возникают для бесконечной среды, будем считать, что клиновая дислокация представляет собой цилиндр конечного радиуса a . Эта задача обладает трансляционной инвариантностью вдоль оси z и вращательной инвариантностью в плоскости x, y . Поэтому будем использовать цилиндрическую систему координат r, φ, z . Пусть

$$u_i = (u_{\hat{r}}, u_{\hat{\varphi}}, u_{\hat{z}}) \quad (36.67)$$

– компоненты ковектора смещения относительно ортонормального базиса в цилиндрической системе координат. Этот ковектор в области (36.66) удовлетворяет уравнению равновесия, которое следует после подстановки (36.3) в уравнение (36.2),

$$(1 - 2\sigma)\Delta u_i + \overset{\circ}{\nabla}_i \overset{\circ}{\nabla}_j u^j = 0, \quad (36.68)$$

где $\overset{\circ}{\nabla}_i$ – ковариантная производная для плоской евклидовой метрики в рассматриваемой системе координат.

Выражения для дивергенции и лапласиана от ковектора смещения в цилиндрической системе координат были получены ранее (7.28)–(7.31). Исходя из симметрии задачи, будем искать решение уравнения (36.68) в виде

$$u_{\hat{r}} = u(r), \quad u_{\hat{\varphi}} = A(r)\varphi, \quad u_{\hat{z}} = 0, \quad (36.69)$$

где $u(r)$ и $A(r)$ – две неизвестные функции только от радиуса. Линейная зависимость угловой компоненты $u_{\hat{\varphi}}$ от угла вытекает из вращательной симметрии задачи. Поставим следующие граничные условия:

$$u_{\hat{r}}|_{r=0} = 0, \quad u_{\hat{\varphi}}|_{r=0} = 0, \quad u_{\hat{\varphi}}|_{\varphi=0} = 0, \quad u_{\hat{\varphi}}|_{\varphi=2\pi\alpha} = -2\pi\theta r, \quad \partial_r u_{\hat{r}}|_{r=a} = 0. \quad (36.70)$$

Первые четыре условия являются геометрическими и соответствуют процессу создания дислокации. Первые два условия – это условие отсутствия смещения оси дислокации. Третье и четвертое условия описывают склейку берегов разреза в линейном приближении. Последнее условие имеет простой физический смысл: отсутствие внешних сил на границе среды.

Неизвестную функцию $A(r)$ находим из предпоследнего граничного условия (36.70)

$$A(r) = -\frac{\theta}{1+\theta}r.$$

Теперь нужно решить уравнения равновесия (36.68). Явные выражения для лапласиана ковекторного поля и дивергенции в цилиндрической системе координат были получены ранее (7.26)–(7.31). Прямая подстановка векторного поля смещений (36.69) показывает, что φ и z компоненты уравнения равновесия (36.68) тождественно удовлетворяются, а радиальная компонента сводится к обыкновенному дифференциальному уравнению

$$\partial_r(r\partial_r u) - \frac{u}{r} = D, \quad D := -\frac{1-2\sigma}{1-\sigma} \frac{\theta}{1+\theta} = \text{const},$$

правая часть которого определяется углом дефицита и коэффициентом Пуассона. Общее решение этого уравнения зависит от двух постоянных интегрирования:

$$u = \frac{D}{2}r \ln r + c_1 r + \frac{c_2}{r}, \quad c_{1,2} = \text{const}.$$

Постоянная интегрирования $c_2 = 0$ в силу граничного условия в нуле. Константу c_1 находим из последнего граничного условия (36.70). В итоге получаем известное решение рассматриваемой задачи [298]

$$\begin{aligned} u_{\hat{r}} &= \frac{D}{2}r \ln \frac{r}{ea}, \\ u_{\hat{\varphi}} &= -\frac{\theta}{1+\theta}r\varphi. \end{aligned} \quad (36.71)$$

В выражении для $u_{\hat{r}}$ буква e обозначает основание натурального логарифма. Отметим, что радиальная компонента вектора смещений в пределе $a \rightarrow \infty$ расходится. Это значит, что для описания клиновой дислокации в рамках теории упругости необходимо рассматривать цилиндр конечного радиуса.

Линейная теория упругости применима в области малых относительных удлинений, которые для клиновой дислокации равны

$$\frac{du_{\hat{r}}}{dr} = -\frac{\theta}{1+\theta} \frac{1-2\sigma}{2(1-\sigma)} \ln \frac{r}{a}, \quad \frac{1}{r} \frac{du_{\hat{\varphi}}}{d\varphi} = -\frac{\theta}{1+\theta}.$$

Это значит, что мы вправе ожидать верных результатов для поля смещений при малых углах дефицита ($\theta \ll 1$) и вблизи края цилиндра ($r \sim a$).

Найдем метрику, индуцированную клиновой дислокацией, в линейном приближении по углу дефицита θ . Вычисления можно провести, воспользовавшись общей формулой (36.6) или известным выражением для вариации формы метрики (см. раздел 2.13)

$$\delta g_{\mu\nu} = -\overset{\circ}{\nabla}_{\mu} u_{\nu} - \overset{\circ}{\nabla}_{\nu} u_{\mu}. \quad (36.72)$$

После несложных вычислений по формуле (36.72), получим следующее выражение для двумерной части метрики в плоскости x, y :

$$dl^2 = \left(1 + \theta \frac{1-2\sigma}{1-\sigma} \ln \frac{r}{a}\right) dr^2 + r^2 \left(1 + \theta \frac{1-2\sigma}{1-\sigma} \ln \frac{r}{a} + \theta \frac{1}{1-\sigma}\right) d\varphi^2. \quad (36.73)$$

С этой метрикой мы сравним метрику, полученную путем решения трехмерных уравнений Эйнштейна, в разделе 36.8.2.

36.7.3 Краевая дислокация

В природе клиновые дислокации встречаются относительно редко, т.к. требуют добавления или удаления большого количества вещества, что связано со значительными энергетическими затратами. Тем не менее их изучение представляет большой интерес, т.к. другие прямолинейные дислокации можно представить в виде суперпозиции клиновых дислокаций. В этом смысле клиновые дислокации являются элементарными. Покажем это на примере краевой дислокации – одним из наиболее распространенных дефектов.

Краевая дислокация, ядро которой совпадает с осью z , изображена на рис. 36.10а. Она возникает в результате разреза среды по полуплоскости $y = 0, x > 0$, сдвига

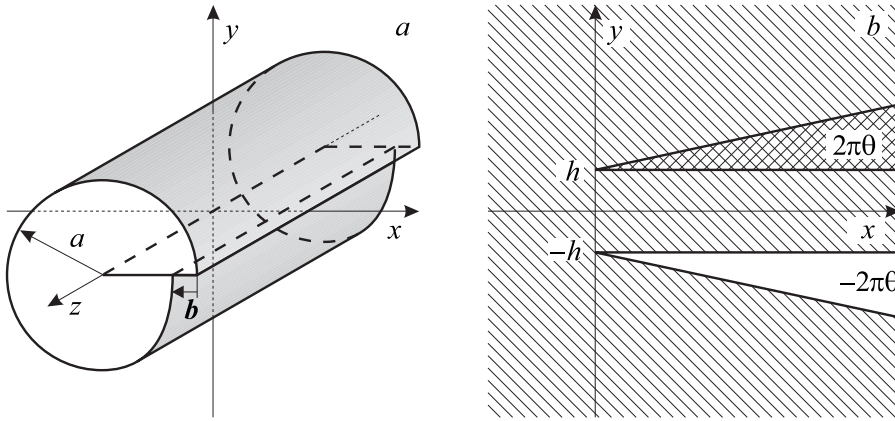


Рис. 36.10: Краевая дислокация с вектором Бюргерса $\mathbf{b}$, направленным к оси дислокации (а). Краевая дислокация, как диполь, состоящий из двух клиновых дислокаций с положительным и отрицательным углом дефицита (b).

нижнего берега разреза к оси z на постоянный (вдали от ядра дислокации) вектор Бюргерса $\mathbf{b}$ с последующей склейкой берегов разреза. Чтобы найти поле смещений краевой дислокации можно решить соответствующую краевую задачу для уравнения равновесия (36.68) [297], глава IV, §27, задача 4. Однако, зная явный вид вектора смещения для клиновой дислокации, мы поступим иначе.

Краевая дислокация представляет собой диполь, состоящий из двух близко расположенных параллельных клиновых дислокаций с положительным $2\pi\theta$ и отрицательным $-2\pi\theta$ углом дефицита, как показано на рис. 36.10b. Будем считать, что оси

первой и второй клиновой дислокации параллельны оси z и пересекают плоскость x, y , соответственно, в точках с координатами $(0, h)$ и $(0, -h)$. Расстояние между осями клиновых дислокаций равно $2h$. Тем самым мы заменили бесконечную среду на цилиндр конечного радиуса $a \gg h$. Ниже мы увидим, что окончательные ответы не зависят от a и могут быть продолжены на все пространство.

Из выражений для поля смещений (36.71) следует, что вдали от начала координат ($r \gg h$) поле смещений клиновых дислокаций с точностью до членов первого порядка малости по θ и h/r имеет вид

$$\begin{aligned} u_x^{(1)} &\approx -\theta \left[\frac{1-2\sigma}{2(1-\sigma)} x \ln \frac{r-h \sin \varphi}{eR} - (y-h) \left(\varphi - \frac{h \cos \varphi}{r} \right) \right], \\ u_y^{(1)} &\approx -\theta \left[\frac{1-2\sigma}{2(1-\sigma)} (y-h) \ln \frac{r-h \sin \varphi}{eR} + x \left(\varphi - \frac{h \cos \varphi}{r} \right) \right], \end{aligned} \quad (36.74)$$

$$\begin{aligned} u_x^{(2)} &\approx \theta \left[\frac{1-2\sigma}{2(1-\sigma)} x \ln \frac{r+h \sin \varphi}{eR} - (y+h) \left(\varphi + \frac{h \cos \varphi}{r} \right) \right], \\ u_y^{(2)} &\approx \theta \left[\frac{1-2\sigma}{2(1-\sigma)} (y+h) \ln \frac{r+h \sin \varphi}{eR} + x \left(\varphi + \frac{h \cos \varphi}{r} \right) \right]. \end{aligned} \quad (36.75)$$

Поскольку уравнения теории упругости линейны, то для нахождения поля смещений краевой дислокации достаточно сложить поля смещений (36.74) и (36.75). После несложных вычислений с точностью до смещения всей среды как целого на постоянный вектор вдоль оси y , получим поле смещений для краевой дислокации

$$\begin{aligned} u_x &= b \left[\operatorname{arctg} \frac{y}{x} + \frac{1}{2(1-\sigma)} \frac{xy}{x^2 + y^2} \right], \\ u_y &= -b \left[\frac{1-2\sigma}{2(1-\sigma)} \ln \frac{r}{ea} + \frac{1}{2(1-\sigma)} \frac{x^2}{x^2 + y^2} \right], \end{aligned} \quad (36.76)$$

где для модуля вектора Бюргерса введено обозначение

$$b := |\mathbf{b}| = -2h\theta.$$

Этот результат совпадает с выражением для поля смещений, полученного путем прямого решения уравнений теории упругости [297], глава IV, §27, задача 4. Тем самым мы доказали, что краевая дислокация является диполем, состоящим из двух параллельных клиновых дислокаций с углами дефицита противоположных знаков.

Найдем метрику, индуцированную краевой дислокацией. Используя формулу (36.72), в линейном приближении по θ и h/r получим нетривиальную двумерную часть метрики в плоскости x, y

$$dl^2 = \left(1 + \frac{1-2\sigma}{1-\sigma} \frac{b}{r} \sin \varphi \right) (dr^2 + r^2 d\varphi^2) - \frac{2b \cos \varphi}{1-\sigma} dr d\varphi. \quad (36.77)$$

Заметим, что индуцированная метрика для краевой дислокации не зависит от радиуса цилиндра a и, следовательно, может быть продолжена на всю плоскость x, y .

36.8 Линейные дислокации в геометрической теории дефектов

Теперь опишем простейшие дислокации в рамках геометрической теории дефектов и сравним полученные результатами с теми, которые были получены в рамках теории упругости в разделе 36.7.

36.8.1 Параллельные клиновые дислокации

В настоящем разделе описано произвольное распределение параллельных прямолинейных клиновых дислокаций в рамках геометрической теории дефектов. Решение этой задачи в теории упругости неизвестно.

При отсутствии дисклинаций, $R_{\mu\nu}{}^j = 0$, $\mathbb{SO}(3)$ связность является чистой калибровкой, и уравнения равновесия для $\mathbb{SO}(3)$ связности (36.29) тождественно удовлетворяются. Явный вид $\mathbb{SO}(3)$ связности в этом случае однозначно определяется полем спиновой структуры ω^{ij} . В силу лоренцевой калибровки (36.44) поле ω^{ij} удовлетворяет уравнениям главного кирального поля. Решение этой системы уравнений задает тривиальную $\mathbb{SO}(3)$ связность. Таким образом, при отсутствии дисклинаций задача сводится к решению уравнений Эйнштейна для репера в упругой калибровке и решению модели главного кирального поля для описания спиновой структуры. После этого можно вычислить тензор кручения по формуле (5.56), который определяет поверхностную плотность вектора Бюргера. Для простоты положим $\omega_\mu{}^{ij} = 0$.

Поскольку дисклинации отсутствуют (тензор кривизны равен нулю), то мы имеем пространство абсолютного параллелизма. Тогда вся геометрия определяется репером $e_\mu{}^i$, который однозначно задает тензор кручения (5.56) при нулевой $\mathbb{SO}(3)$ связности. Здесь мы предполагаем, что тривиальная $\mathbb{SO}(3)$ связность равна нулю. Репер $e_\mu{}^i$ удовлетворяет трехмерным уравнениям Эйнштейна с евклидовой сигнатурой метрики, которые следуют из выражения для свободной энергии (36.35) при $R_{\mu\nu}{}^j = 0$,

$$\tilde{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\tilde{R} = -\frac{1}{2}T_{\mu\nu}. \quad (36.78)$$

Здесь в правую часть уравнений Эйнштейна мы добавили источник дислокаций $T_{\mu\nu}$ (в гравитации, это – тензор энергии-импульса материи).

Отметим, что без источника дислокаций модель была бы тривиальной. Действительно, при $T_{\mu\nu} = 0$ из уравнений Эйнштейна (36.78) следует, что скалярная кривизна и тензор Риччи равны нулю: $\tilde{R} = 0$, $\tilde{R}_{\mu\nu} = 0$. Поскольку в трехмерном пространстве полный тензор кривизны взаимно однозначно определяется тензором Риччи (17.24), то отсюда следует, что без источников он также обращается в нуль. Обращение в нуль полного тензора кривизны и означает тривиальность модели, поскольку в этом случае дефектов просто нет. Аналогичное утверждение хорошо известно в гравитации. Обычно его формулируют следующим образом: “Трехмерная гравитация не описывает динамических, т.е. распространяющихся, степеней свободы”.

Для наших целей необходимо найти решение уравнений Эйнштейна (36.78) для произвольного числа клиновых дислокаций. Уравнения Эйнштейна представляют собой систему нелинейных дифференциальных уравнений в частных производных второго порядка. Точных решений, даже в трехмерном пространстве, известно не так много. Замечательное точное решение, описывающее произвольное статическое распределение точечных частиц хорошо известно в трехмерной гравитации для метрики лоренцевой сигнатуры $(+ - -)$ [309, 310, 311]. Получим аналог этого решения для евклидовой сигнатуры метрики и покажем, что в геометрической теории дефектов оно описывает произвольное распределение параллельных клиновых дислокаций. Тем самым сначала мы рассмотрим общий случай произвольного числа клиновых дислокаций, а затем подробно остановимся на интересующем нас примере одной клиновой дислокации. Это сделано намеренно, поскольку решение в более общем случае не вносит существенных усложнений. В то же время произвольное распределение клиновых дислокаций представляет гораздо больший интерес для приложений. Так, в

предыдущем разделе было показано, что краевая дислокация представляется диполем, состоящим из двух параллельных клиновых дислокаций разных знаков.

Рассмотрим упругую среду с произвольным распределением параллельных клиновых дислокаций. Выберем систему координат таким образом, чтобы ось $z := x^3$ была параллельна осям дислокаций, а оси $\{x^\alpha\} = \{x, y\}$, $\alpha = 1, 2$ были перпендикулярны оси z . Тогда метрика примет блочно диагональный вид

$$ds^2 = dl^2 + N^2 dz^2, \quad (36.79)$$

где

$$dl^2 := g_{\alpha\beta} dx^\alpha dx^\beta$$

– двумерная метрика на плоскости x, y . Благодаря трансляционной симметрии вдоль оси z двумерная метрика $g_{\alpha\beta}(x, y)$ и функция хода $N(x, y)$ не зависят от z . При этом мы не предполагаем никакой вращательной симметрии.

Отвлекаясь от физической аргументации, можно сказать проще. Рассмотрим блочно диагональную метрику вида (36.79), которая обладает трансляционной инвариантностью вдоль оси z . А затем покажем, что соответствующее решение уравнений Эйнштейна действительно описывает произвольное распределение параллельных клиновых дислокаций.

Тензор кривизны для метрики (36.79) имеет следующие компоненты:

$$\begin{aligned} \tilde{R}_{\alpha\beta\gamma}{}^\delta &= R_{\alpha\beta\gamma}^{(2) \delta}, & \tilde{R}_{\alpha z\gamma}{}^z &= \frac{1}{N} \nabla_\alpha \nabla_\gamma N, \\ \tilde{R}_{\alpha\beta\gamma}{}^z &= \tilde{R}_{\alpha z\gamma}{}^\delta = 0, \end{aligned}$$

где $R_{\alpha\beta\gamma}^{(2) \delta}$ – тензор кривизны для двумерной метрики $g_{\alpha\beta}$, и ∇_α – двумерная ковариантная производная с символами Кристоффеля также для двумерной метрики $g_{\alpha\beta}$. Тензор Риччи с скалярная кривизна имеют следующие компоненты:

$$\begin{aligned} \tilde{R}_{\alpha\beta} &= R_{\alpha\beta}^{(2)} + \frac{1}{N} \nabla_\alpha \nabla_\beta N, \\ \tilde{R}_{\alpha z} &= 0, \\ \tilde{R}_{zz} &= N \nabla^\alpha \nabla_\alpha N, \\ \tilde{R} &= R^{(2)} + \frac{2}{N} \nabla^\alpha \nabla_\alpha N, \end{aligned}$$

где $\nabla^\alpha := g^{\alpha\beta} \nabla_\beta$.

Выберем источник дислокаций в следующем виде

$$\begin{aligned} T_{zz} &= \frac{4\pi}{\sqrt{g^{(2)}}} \sum_{i=1}^N \theta_i \delta(\mathbf{r} - \mathbf{r}_i), \\ T_{\alpha\beta} &= T_{\alpha z} = T_{z\alpha} = 0, \end{aligned} \quad (36.80)$$

где $\delta(\mathbf{r} - \mathbf{r}_i) := \delta(x - x_i) \delta(y - y_i)$ – двумерная δ -функция на плоскости x, y с носителем в точке $\mathbf{r}_i(x_i, y_i)$. Появление множителя $g^{(2)} := \det g_{\alpha\beta}$ перед знаком суммы связано с тем, что δ -функция с точки зрения общих преобразований координат является не функцией, а скалярной плотностью. Ниже показано, что решение уравнений Эйнштейна с таким источником описывает N клиновых параллельных дислокаций

с углами дефицита θ_1 , которые пересекают плоскость x, y в точках (x_1, y_1) . В трехмерной гравитации такой источник соответствует частицам с массами $m_1 := 2\pi\theta_1$, покоящимся в точках $\mathbf{r}_1$.

Уравнения Эйнштейна (36.78) с источником (36.80) для блочно диагональной метрики (36.79) сводятся к четырем уравнениям:

$$\nabla_\alpha \nabla_\beta N - g_{\alpha\beta} \nabla^\gamma \nabla_\gamma N = 0, \quad (36.81)$$

$$-\frac{1}{2}N^3 R^{(2)} = -\frac{2\pi}{\sqrt{g^{(2)}}} \sum_{i=1} \theta_i \delta(\mathbf{r} - \mathbf{r}_i), \quad (36.82)$$

где $R^{(2)}$ – двумерная скалярная кривизна.

Метрика вида (36.79) все еще инвариантна относительно преобразований координат на плоскости x, y . Используя эту оставшуюся инвариантность, зафиксируем конформную калибровку на плоскости (локально это всегда возможно)

$$g_{\alpha\beta} = e^{2\phi} \delta_{\alpha\beta}, \quad (36.83)$$

где $\phi(x, y)$ – некоторая функция.

В конформной калибровке уравнение (36.81) принимает вид

$$\partial_\alpha \partial_\beta N = 0.$$

При постоянных граничных условиях для N на границе плоскости x, y это уравнение имеет единственное решение $N = \text{const}$. Изменив масштаб координаты z , можно положить $N = 1$ без ограничения общности. Тогда уравнение (36.82) сводится к уравнению Пуассона

$$\Delta \phi = 2\pi \sum_i \theta_i \delta(\mathbf{r} - \mathbf{r}_i), \quad (36.84)$$

которое имеет хорошо известное общее решение (см., например, [3], глава III, §11, раздел 7)

$$\phi = \sum_i \theta_i \ln |\mathbf{r} - \mathbf{r}_i| + \frac{1}{2} \ln C, \quad C = \text{const} > 0.$$

Следовательно, метрика в плоскости x, y имеет вид

$$dl^2 = C \prod_i |\mathbf{r} - \mathbf{r}_i|^{2\theta_i} (dr^2 + r^2 d\varphi^2), \quad 0 < r < \infty, \quad 0 < \varphi < 2\pi, \quad (36.85)$$

где полярные координаты r, φ покрывают всю плоскость $\mathbb{R}^2$ и не более того (что важно!). Любое решение уравнений Эйнштейна определено с точностью до выбора системы координат, поскольку уравнения ковариантны. Воспользуемся этим и положим $C = 1$, чего всегда можно добиться путем выбора масштаба радиальной координаты r .

Таким образом, метрика

$$ds^2 = \prod_i |\mathbf{r} - \mathbf{r}_i|^{2\theta_i} (dr^2 + r^2 d\varphi^2) + dz^2 \quad (36.86)$$

является точным решением нелинейных уравнений Эйнштейна, описывающим, как мы увидим, произвольное распределение параллельных клиновых дислокаций. Физический смысл данного решения будет ясен из дальнейшего рассмотрения.

Отметим, что переход к непрерывному распределению дислокаций в геометрическом подходе прост. Для этого вместо δ -образных источников в правую часть уравнений Эйнштейна нужно подставить непрерывное распределение источников дислокаций. Простейшие примеры будут рассмотрены в разделе 36.8.4.

В заключение данного раздела вычислим плотность энергии произвольного распределения клиновых дислокаций. По определению она равна интегралу от плотности свободной энергии

$$\begin{aligned} E &:= \int dx dy dz \sqrt{g} \tilde{R} = \int dz \int dx dy N^2 \det \sqrt{g^{(2)}} \left(R^{(2)} + \frac{2}{N} \nabla^\alpha \nabla_\alpha N \right) = \\ &= \int dz \int dx dy 4\pi \sum_i \theta_i \delta(\mathbf{r} - \mathbf{r}_i) = 4\pi \int dz \sum_i \theta_i, \end{aligned}$$

где мы воспользовались равенством $N = 1$ и уравнением равновесия (36.82). Таким образом, линейная плотность энергии клиновых дислокаций с точностью до постоянной равна сумме углов дефицита всех дислокаций.

36.8.2 Клиновая дислокация в геометрическом подходе

Чтобы показать, что метрика (36.85) действительно описывает произвольное распределение клиновых дислокаций, рассмотрим подробнее одну клиновую дислокацию с источником в начале координат,

$$T_{zz} := \frac{4\pi}{\sqrt{g^{(2)}}} \theta \delta(\mathbf{r}). \quad (36.87)$$

Соответствующая двумерная метрика (36.85) при $C = 1$ имеет вид

$$dl^2 = r^{2\theta} (dr^2 + r^2 d\varphi^2). \quad (36.88)$$

Перейдем в новую систему координат

$$\hat{r} := \frac{1}{\alpha} r^\alpha, \quad \hat{\varphi} := \alpha \varphi, \quad \alpha := 1 + \theta, \quad (36.89)$$

в которой метрика становится евклидовой

$$dl^2 = d\hat{r}^2 + \hat{r}^2 d\hat{\varphi}^2. \quad (36.90)$$

Однако теперь область изменения полярного угла отличается от 2π : $0 < \hat{\varphi} < 2\pi\alpha$, и покрывает плоскость x, y с вырезанным или добавленным клином $2\pi\theta$, что зависит от знака угла дефицита θ .

Поскольку в новой системе координат $\hat{r}, \hat{\varphi}$ метрика совпадает с евклидовой, то здесь мы имеем евклидову плоскость с вырезанным или добавленным клином, т.к. угол $\hat{\varphi}$ меняется в интервале $(0, 2\pi\alpha)$. Переход к координатам r, φ (36.89) обозначает склейку берегов образовавшегося клина, что соответствует конусу. Поэтому обе метрики (36.88) и (36.90) описывают один и тот же геометрический объект – коническую особенность. Очевидно, что тензоры кручения и кривизны равны нулю всюду, за исключением начала координат.

Возникновение конической особенности в точности совпадает с созданием клиновой дислокации в теории дефектов. Нетрудно убедиться в том, что общее решение

(36.85) описывает произвольное распределение конических особенностей с углами дефицита θ_1 , расположенными в точках $\mathbf{r}_1$. Следовательно, это решение описывает произвольное распределение параллельных клиновых дислокаций.

В дальнейшем нам понадобится еще один вид метрики конической особенности. Для этого совершим преобразование координат

$$f := \alpha \hat{r}, \quad \varphi := \frac{1}{\alpha} \hat{\varphi}. \quad (36.91)$$

Тогда метрика (36.90) примет вид

$$dl^2 = \frac{1}{\alpha^2} df^2 + f^2 d\varphi^2, \quad \alpha := 1 + \theta, \quad (36.92)$$

где $0 < f < \infty$ – радиальная координата и $0 < \varphi < 2\pi$ – полярный угол. Это – еще одна часто используемая форма метрики для конической особенности.

С качественной точки зрения создание клиновой дислокации совпадает с определением конической особенности. Однако существует количественное расхождение, поскольку метрика (36.92) зависит только от угла дефицита θ и не может совпадать с индуцированной метрикой (36.73), полученной в рамках теории упругости. Это расхождение связано с тем, что в теории упругости после вырезания клина и склейки берегов разреза (создания конической сингулярности) мы требуем, чтобы среда пришла в равновесие, т.е. вектор смещения в положении равновесия удовлетворял уравнениям равновесия. В то же время для конической особенности плоскость x, y после склейки может быть деформирована произвольным образом. С формальной точки зрения это проявляется в том, что индуцированная метрика (36.73), полученная в рамках теории упругости, явно зависит от коэффициента Пуассона, который отсутствует в теории гравитации.

Для решения этой проблемы в статье [295] была предложена упругая калибровка. Выберем упругую калибровку (36.43), как самую простую в случае клиновой дислокации. Эту задачу можно решать двумя способами. Во-первых, калибровочное условие можно подставить непосредственно в уравнения Эйнштейна. Во-вторых, можно найти решение в какой-то удобной системе координат, а затем найти такое преобразование координат, чтобы было выполнено калибровочное условие.

Поскольку точное решение для метрики известно (36.92), то проще следовать второму пути. Метрике (36.92) можно поставить в соответствие репер

$$e_{\hat{r}} = \frac{1}{\alpha}, \quad e_{\hat{\varphi}} = f.$$

Здесь шляпка над индексом означает, что он относится к ортонормальной системе координат, а индекс без шляпки является координатным. Компоненты этого репера являются квадратными корнями из соответствующих компонент метрики и поэтому имеют симметричное линейное приближение (36.41). Поскольку клиновая дислокация инвариантна относительно вращений в плоскости x, y , то совершим преобразование радиальной координаты $f \mapsto f(r)$. После преобразования компоненты репера примут вид

$$e_{\hat{r}} = \frac{f'}{\alpha}, \quad e_{\hat{\varphi}} = f, \quad (36.93)$$

где штрих обозначает дифференцирование по r . Репер, соответствующий евклидовой метрике в цилиндрической системе координат, выберем в виде

$$\overset{\circ}{e}_{\hat{r}} = 1, \quad \overset{\circ}{e}_{\hat{\varphi}} = r. \quad (36.94)$$

Ему соответствуют символы Кристоффеля $\overset{\circ}{\Gamma}_{\mu\nu}{}^\rho$ и $\mathbb{SO}(3)$ связность $\overset{\circ}{\omega}_{\mu i}{}^j$, которые определяют ковариантную производную. Выпишем только нетривиальные компоненты (см. раздел 7.2):

$$\begin{aligned}\overset{\circ}{\Gamma}_{r\varphi}{}^\varphi &= \overset{\circ}{\Gamma}_{\varphi r}{}^\varphi = \frac{1}{r}, & \overset{\circ}{\Gamma}_{\varphi\varphi}{}^r &= -r, \\ \overset{\circ}{\omega}_{\varphi\hat{r}}{}^{\hat{\varphi}} &= -\overset{\circ}{\omega}_{\varphi\hat{\varphi}}{}^{\hat{r}} = 1.\end{aligned}$$

Подстановка репера (36.93) в калибровочное условие (36.43) приводит к дифференциальному уравнению Эйлера на функцию перехода

$$\frac{f''}{\alpha} + \frac{f'}{\alpha r} - \frac{f}{r^2} + \frac{\sigma}{1-2\sigma} \left(\frac{f''}{\alpha} + \frac{f'}{r} - \frac{f}{r^2} \right) = 0. \quad (36.95)$$

Его общее решение зависит от двух постоянных $C_{1,2}$:

$$f = C_1 r^{\gamma_1} + C_2 r^{\gamma_2}, \quad (36.96)$$

где показатели степени $\gamma_{1,2}$ определяются квадратным уравнением

$$\gamma^2 + 2\gamma\theta b - \alpha = 0, \quad b := \frac{\sigma}{2(1-\sigma)},$$

которое при $\theta > -1$ имеет действительные корни разных знаков: положительный корень γ_1 и отрицательный корень γ_2 .

График функции $b(\sigma)$ приведен на рис. 36.11 слева. Напомним, что на коэффициент Пуассона имеются физические ограничения $-1 \leq \sigma \leq 1/2$ [297], глава I, §5. Поэтому значения b ограничены отрезком $-1/4 \leq b \leq 1/2$.

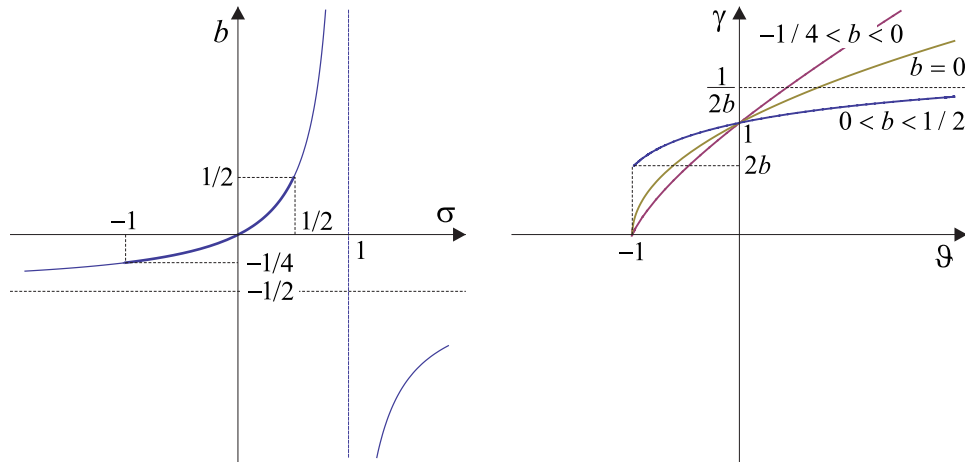


Рис. 36.11: Слева. График функции $b(\sigma)$. Справа. Графики функций $\gamma(\theta)$ для отрицательных, $-1/4 < b < 0$, нулевого, $b = 0$, и положительных значений, $0 < b < 1/2$.

Для фиксирования постоянных интегрирования наложим граничные условия на репер

$$e_r{}^{\hat{r}}|_{r=a} = 1, \quad e_\varphi{}^{\hat{\varphi}}|_{r=0} = 0. \quad (36.97)$$

Первое граничное условие соответствует последнему граничному условию на вектор смещения (36.70) (отсутствие внешних сил на поверхности цилиндра), а второе –

отсутствию угловой составляющей у тензора деформации в ядре дислокации. Уравнения (36.97) задают постоянные интегрирования

$$C_1 = \frac{\alpha}{\gamma_1 a^{\gamma_1 - 1}} \quad C_2 = 0. \quad (36.98)$$

Таким образом, полученный репер определяет метрику

$$dl^2 = \left(\frac{r}{a}\right)^{2\gamma-2} \left(dr^2 + \frac{\alpha^2 r^2}{\gamma^2} d\varphi^2\right), \quad (36.99)$$

где мы, для простоты, отбросили у положительного корня γ_1 индекс:

$$\gamma := -\theta b + \sqrt{\theta^2 b^2 + 1 + \theta} > 0.$$

Это и есть решение поставленной задачи. Найденное решение справедливо при всех углах дефицита θ и всех $0 < r < a$. Полученная метрика зависит от трех постоянных: θ , σ и a . Зависимость от угла дефицита θ появилась благодаря тому, что он входит в правую часть уравнений Эйнштейна (36.78). Зависимость от коэффициента Пуассона возникла вследствие упругой калибровки (36.43), и, наконец, зависимость от радиуса цилиндра следует из граничного условия (36.97).

На рис. 36.11 справа приведены графики функций $\gamma(\theta)$ для отрицательных, $-1/4 < b < 0$, нулевого, $b = 0$, и положительных значений, $0 < b < 1/2$. Для положительных значений b график функции имеет асимптоту $\gamma = 1/(2b)$ при $\theta \rightarrow \infty$.

Если клиновая дислокация отсутствует, то $\theta = 0$, $\alpha = 1$, $\gamma = 1$, и метрика (36.99) переходит в евклидову метрику $dl^2 = dr^2 + r^2 d\varphi^2$, что и следовало ожидать.

Если выполнено равенство

$$\gamma = \alpha \quad \Leftrightarrow \quad \frac{\theta(1 + \theta)}{1 - \sigma} = 0, \quad (36.100)$$

то метрика клиновой дислокации (36.99) примет вид

$$dl^2 = \left(\frac{r}{a}\right)^{2(\alpha-1)} (dr^2 + r^2 d\varphi^2).$$

Эта метрика отличается от метрики (36.88) только на постоянный множитель, который всегда можно абсорбировать за счет растяжки радиальной координаты. Поскольку на коэффициент Пуассона существует ограничение $-1 \leq \sigma \leq 1/2$, то второе уравнение (36.100) имеет два решения: $\theta = 0$ и $\theta = -1$. В первом случае дислокация отсутствует, а второй случай является предельным и соответствует удалению всего вещества, что также не представляет интереса.

Сравним метрику (36.99), полученную в рамках геометрического подхода, с индуцированной метрикой из теории упругости (36.73). Во-первых, она имеет более простой вид. Во-вторых, в линейном приближении по θ справедливо равенство

$$\gamma \approx 1 + \theta \frac{1 - 2\sigma}{2(1 - \sigma)}, \quad (36.101)$$

и метрика (36.99), как нетрудно проверить, действительно совпадает с метрикой (36.73), полученной в рамках теории упругости. Мы видим, что индуцированная метрика (36.73) является только линейным приближением для метрики, полученной в геометрической теории дефектов, которая, к тому же, имеет более простой вид.

Вне рамок теории возмущений мы наблюдаем существенные различия. В частности, метрика (36.73) сингулярна в начале координат, в то время как метрика (36.99), найденная вне рамок теории возмущений, регулярна.

Тензоры напряжений и деформаций связаны между собой законом Гука (36.3). Поскольку тензор деформаций является линейным приближением для индуцированной метрики, то возникает экспериментальная возможность проверки формулы (36.99). Для этого необходимо измерить поле напряжений для одиночной клиновой дислокации. Тем самым геометрическая теория дефектов может быть экспериментально подтверждена или опровергнута.

Задача восстановления поля смещений по заданной метрике сводится к решению дифференциальных уравнений (36.6), в правую часть которых необходимо подставить метрику (36.99), с граничными условиями (36.70). На этой задаче мы останавливаться не будем. Отметим, что в геометрической теории дефектов сложный этап нахождения вектора смещений там, где он существует, просто отсутствует и не является необходимым.

36.8.3 Конформные отображения

Для упрощения формул все вычисления настоящего раздела проведены в конформной, а не в упругой калибровке для двумерной части метрики. Поэтому результаты не зависят от коэффициента Пуассона.

В рамках геометрической теории дефектов можно решить задачу описания произвольного числа параллельных клиновых дислокаций с помощью теории функций комплексного переменного. Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовой системой координат x, y, Z . Пусть произвольное число N параллельных клиновых дислокаций с углами дефицита θ_i , $i = 1, \dots, N$, расположены параллельно оси Z . Тем самым задача обладает трансляционной симметрией вдоль оси Z и, по существу, двумерна. Введем на плоскости x, y комплексные координаты $z := x + iy$. Пусть дислокации пересекают плоскость x, y в точках z_i . Тогда нетривиальная двумерная часть метрики (36.86) на плоскости комплексного переменного z примет вид

$$dl^2 = dzd\bar{z} \prod_i [(z - z_i)(\bar{z} - \bar{z}_i)]^{\theta_i}, \quad (36.102)$$

где черта обозначает комплексное сопряжение.

Процесс создания клиновых дислокаций можно описать, если найти такое преобразование координат, где метрика (36.102) станет евклидовой. Это преобразование является конформным отображением комплексной плоскости $z \mapsto w(z)$. Нетрудно проверить, что евклидова метрика

$$dl^2 = dw d\bar{w}, \quad (36.103)$$

после преобразования

$$w := u + iv = e^{i\phi} \int_{z_0}^z d\zeta \prod_i (\zeta - z_i)^{\theta_i} + C. \quad (36.104)$$

где z_0 – некоторая фиксированная точка, $z_0 \neq z_i$, и $\phi \in \mathbb{R}$, $C \in \mathbb{C}$ – произвольные числа, переходит в метрику (36.102). Изменение точки z_0 приводит к переопределению постоянной C . Поэтому, без ограничения общности, положим $z_0 = 0$. Тогда точка $z = 0$ отобразится в точку $w = C$. Контур интегрирования в (36.104) не должен

пересекать и содержать петли вокруг точек z_i . При $\theta_i > -1$ интеграл в правой части сходится. Ниже мы проанализируем этот интеграл в нескольких простых случаях.

Если все клиновые дислокации пересекают плоскость z в точках, лежащих на оси x , т.е. $\text{im } z_i = 0$, то интеграл (36.104) хорошо известен в теории функции комплексного переменного и называется *интегралом Кристоффеля–Шварца* (см. раздел ??). Он конформно отображает верхнюю полуплоскость $z > 0$ на многоугольник, одна вершина которого лежит в бесконечности, а остальные – в точках w_i . Каждая точка z_i отображается в вершину w_i с внутренним углом $\pi(1 + \theta_i)$. В общем случае, когда $\text{im } z_i \neq 0$ для некоторых дислокаций, конформное отображение (36.104) является более сложным.

Полный угол дефицита (или заряд), нормированный на 2π , равен

$$\Theta := \sum_{i=1}^N \theta_i. \quad (36.105)$$

Из физических соображений вытекает, что $\Theta > -1$, потому что нельзя удалить больше вещества, чем все $\mathbb{R}^3$. Полный вектор Бюргерса (дипольный момент) и квадрупольный момент, нормированные на 2π , имеют вид

$$B := \sum_{i=1}^N \theta_i z_i, \quad (36.106)$$

$$M := \frac{1}{2} \sum_{i,j=1}^N (\theta_i \theta_j - \theta_i \delta_{ij}) z_i z_j = \frac{1}{2} B^2 - \frac{1}{2} \sum_i \theta_i z_i^2. \quad (36.107)$$

В данном определении полный вектор Бюргерса и квадрупольный момент рассматриваются как комплексные числа. В разделе 36.7.3 мы определили вектор Бюргерса другим образом. Как вектор, он имеет такую же абсолютную величину $|B|$, но отличается от вектора, соединяющего начало координат комплексной плоскости, с точкой B .

Интеграл (36.104) имеет разные асимптотики при $z \rightarrow \infty$ в зависимости от полного угла дефицита. Выпишем три первых члена разложения для возможных значений Θ :

$$w \approx e^{i\phi} \left(\frac{1}{1+\Theta} z^{1+\Theta} - \frac{B}{\Theta} z^\Theta + \frac{M}{\Theta-1} z^{\Theta-1} \right) + C, \quad \Theta \neq 0, 1, \quad (36.108)$$

$$w \approx e^{i\phi} \left(z - B \ln z - \frac{M}{z} \right) + C, \quad \Theta = 0, \quad (36.109)$$

$$w \approx e^{i\phi} \left(\frac{1}{2} z^2 - Bz + M \ln z \right) + C, \quad \Theta = 1. \quad (36.110)$$

При $\Theta \neq 0, 1, 2, \dots$ логарифмических слагаемых в разложении нет. В случае $\Theta = 0$ логарифм появляется при векторе Бюргерса. Квадрупольное слагаемое имеет логарифмический вид для $\Theta = 1$. Значения $\Theta = 2, 3, \dots$ определяют положение слагаемого с логарифмом в разложении $w(z)$ в высших членах разложения, которые отброшены.

Разложения (36.108)–(36.110) показывают, что любая экстремаль, уходящая в бесконечность в плоскости z также уходит в бесконечность в плоскости w . В таких случаях экстремали, очевидно, являются полными. Если экстремаль не уходит в бесконечность, то она также полна, т.к. либо замкнута, либо может быть продолжена до

бесконечного значения канонического параметра внутри ограниченной области (при этом предполагается, что экстремали продолжаются через коническую особенность естественным образом).

Поведение функции $w(z)$ на больших расстояниях в главном приближении определяется первыми слагаемыми в разложениях (36.108)–(36.110). В этом случае двумерная часть метрики принимает вид

$$dl^2 = dw d\bar{w} = (z\bar{z})^\Theta dz d\bar{z},$$

что совпадает с метрикой (36.88) для отдельной клиновой дислокации, записанной в комплексных координатах. Поэтому в дальнейшем мы рассмотрим детально только дипольное и квадрупольное приближение.

Интеграл (36.104) задает конформное отображение между комплексной плоскостью z (сечение среды, содержащей произвольное число клиновых дислокаций) и плоскостью w (сечение среды с евклидовой метрикой до создания дислокаций). Плоскость w имеет разрезы, наглядно показывающие процесс создания дислокаций. Ниже мы рассмотрим несколько простейших случаев, описывающих небольшое число клиновых дислокаций.

Интеграл Кристоффеля–Шварца зависит от двух постоянных ϕ и C . Для определенности, предположим, что точка z_1 всегда совпадает с началом координат плоскости z , и ни одна из точек z_i , $i = 2, 3, \dots$, не лежит на отрицательной части вещественной оси. Тогда постоянные ϕ и C можно выбрать таким образом, чтобы образ w_1 совпал с началом координат плоскости w , и отрицательная часть оси x отображалась в отрицательную часть вещественной оси u .

Интеграл Кристоффеля–Шварца можно взять в элементарных функциях только в некоторых специальных случаях. Приведем несколько примеров, которые имеют простую геометрическую и физическую интерпретацию.

Клиновая дислокация

Простейший пример отображения Кристоффеля–Шварца (36.104) дает клиновая дислокация, описанная в разделе 36.8.2. В этом случае мы имеем только одну ось дислокации, расположенную в начале координат

$$z_1 = 0, \quad \theta_1 =: \theta.$$

Интеграл Кристоффеля–Шварца (36.104) для клиновой дислокации явно берется

$$w = \frac{e^{-i\pi\theta}}{1+\theta} z^{1+\theta}, \quad (36.111)$$

где мы положили для наглядности $\phi := -\pi\theta$ и $C = 0$. Эта функция осуществляет конформное отображение верхней полуплоскости $\text{im } z > 0$ в двуугольник с вершинами в точках $w_1 = 0$ и $w_2 = \infty$. При этом точки $z_1 = 0$ и $z_2 = \infty$ отображаются в точки $w_1 = 0$ и $w_2 = \infty$, соответственно.

Функция $z^{1+\theta}$ голоморфна, например, в комплексной плоскости z с разрезом вдоль неотрицательной части вещественной полуоси, $\text{re } z \geq 0$. Рассмотрим куда отображается вещественная ось z при движении от $-\infty$ до ∞ . Из формулы (36.111) следует равенство

$$\arg w = -\pi\theta + (1+\theta) \arg z.$$

Отсюда вытекает, что отрицательная часть вещественной оси $\arg z = \pi$ отображается в луч, выходящий под углом π , т.е. совпадает с отрицательной полуосью $\arg w = \pi$. Положительная часть вещественной оси $\arg z = 0$ отображается в луч на плоскости w , выходящий из начала координат под углом $-\pi\theta$ к оси ординат. Как уже отмечалось, функция $z^{1+\theta}$ голоморфна во всей z -плоскости с разрезом, например, вдоль неотрицательной части вещественной оси. Поэтому конформное отображение (36.111) аналитически продолжается в нижнюю полуплоскость через луч $\arg z = \pi$. В целом, функция (36.111) отображает плоскость z с разрезом вдоль неотрицательной вещественной полуоси в плоскость w с вырезанным клином, угол которого равен $2\pi|\theta|$. Соответствующее конформное для отрицательных углов дефицита, $-1 < \theta < 0$, показано на рис. 36.12. При $-1 < \theta < 0$ клин вырезается, а при $0 < \theta < \infty$ клин

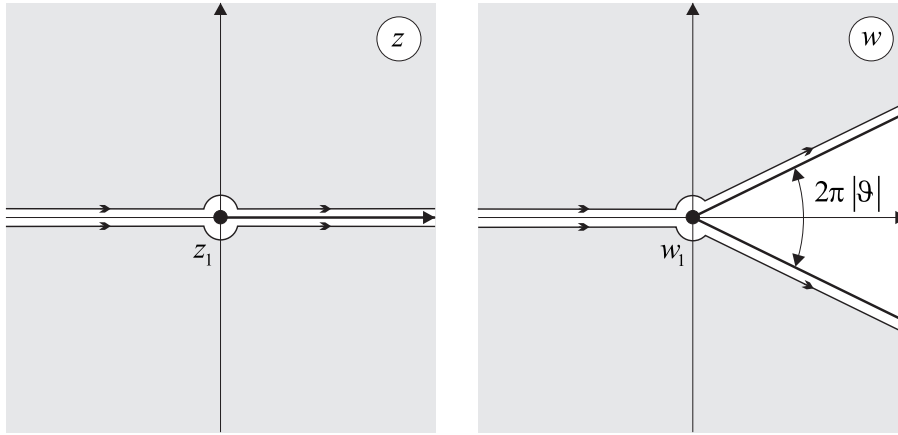


Рис. 36.12: Конформное отображение Кристоффеля–Шварца для клиновой дислокации углом дефицита $0 < \theta < 0$.

вставляется.

С математической точки зрения клиновая дислокация является конической особенностью на евклидовой плоскости с углом дефицита $2\pi\theta$.

Этот пример конформного отображения явно следует процессу создания клиновой дислокации. Мы берем бесконечную упругую среду (плоскость w), вырезаем клин с углом $2\pi|\theta|$ для отрицательных $\theta < 0$ и отождествляем точки разреза, соответствующие одной и той же точке в плоскости z . Для положительных углов дефицита $\theta > 0$ необходимо сделать разрез в плоскости w , раздвинуть берега разреза и вставить клин вещества без упругих напряжений.

Краевая дислокация

Диполь, состоящий из двух клиновых дислокаций с углами дефицита противоположного знака, см. раздел 36.7.3, характеризуется постоянным вектором Бюргерса $\mathbf{B}$, который мы определим немного позже. Для того, чтобы построить соответствующее конформное отображение, выберем следующее положение клиновых дислокаций:

$$\begin{aligned} z_1 &= 0, & z_2 &= h, \\ \theta_1 &= -\theta, & \theta_2 &= \theta, \end{aligned} \quad (36.112)$$

где $\operatorname{im} h := 0$. Для определенности положим $0 < \theta < 1$. Тогда интеграл Кристоффеля–Шварца (36.104) принимает вид

$$w = \int_0^z d\zeta \left(\frac{\zeta - h}{\zeta} \right)^\theta. \quad (36.113)$$

Здесь постоянные ϕ и C выбраны таким образом, что $w_1(z_1) = 0$, и отрицательная полуось x отображается в отрицательную полуось u .

Посмотрим, куда отображается вещественная ось переменного z . Нетрудно проверить, что луч $\arg z = \pi$ отображается в луч $\arg w = \pi$. Кроме того, при интегрировании вдоль вещественной оси отрезок $[0, h]$ отображается в луч $\arg w = \pi\theta$. Здесь следует заметить, что на отрезке $\zeta \in [0, h]$ выполнено равенство

$$(\zeta - h)^\theta = e^{i\pi\theta}(h - \zeta)^\theta, \quad h - \zeta > 0.$$

Положение второй вершины в верхней полуплоскости w (см. рис. 36.13) выражается

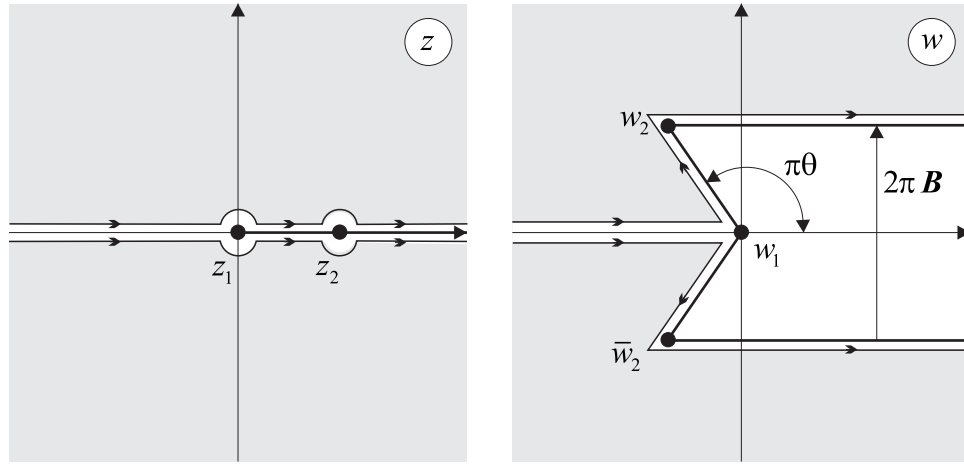


Рис. 36.13: Конформное отображение Кристоффеля–Шварца для диполя из двух клиновых дислокаций, которые представляют одну краевую дислокацию

через гамма-функцию Эйлера (см. пример ??):

$$w_2 = \int_0^h d\zeta \left(\frac{\zeta - h}{\zeta} \right)^\theta = h e^{i\pi\theta} \int_0^1 dx \left(\frac{1-x}{x} \right)^\theta = h e^{i\pi\theta} \Gamma(1-\theta)\Gamma(1+\theta). \quad (36.114)$$

При дальнейшем движении точки z вдоль вещественной оси от h до ∞ точка w движется вдоль прямой, параллельной вещественной оси. Таким образом построено отображение верхней полуплоскости на треугольник в плоскости w , одна из вершин которого расположена в бесконечности. Поскольку соответствующий интеграл голоморфен на отрицательном луче вещественной оси $\arg z = \pi$, то его можно продолжить в нижнюю полуплоскость. В целом, интеграл Кристоффеля–Шварца (36.113) отображает всю комплексную плоскость z с разрезом вдоль неотрицательной части вещественной оси на комплексную плоскость w , из которой удалена полоса, как показано на рис. 36.13 для $1/2 < \theta < 1$.

Определим вектор Бюргерса, как вектор, соединяющий две точки w_- и w_+ , которые соответствуют нижнему и верхнему берегам разреза для одной точки на плоскости z . Для краевой дислокации с $x > z_2$ – это постоянный вектор. Длина нормированного вектора равна

$$2\pi|\mathbf{B}| := w_+ - w_- = 2\operatorname{im} w_2. \quad (36.115)$$

Это определение пригодно для произвольного распределения клиновых дислокаций. В случае краевой дислокации будем считать, что вектор Бюргерса имеет начало в точке w_- и конец – в w_+ .

В общем случае разрез можно определить как ломаную, соединяющую последовательность точек $z_1, \dots, z_N, \infty$, включающую бесконечно удаленную точку. Конечно, соответствующий вектор Бюргерса может быть непостоянным.

Используя выражение (36.114) и свойства гамма-функций, мы получаем вектор Бюргерса для краевой дислокации:

$$2\pi|\mathbf{B}| = 2h\Gamma(1-\theta)\Gamma(1+\theta)\sin(\pi\theta) = 2\pi h\theta.$$

Этот замечательный результат показывает, что нормированный вектор Бюргерса равен произведению $h\theta$. Заметим, что это точный результат для диполя из двух клиновых дислокаций, и он совпадает с определением (36.106) в разложении интеграла Кристоффеля–Шварца на больших расстояниях для произвольного распределения клиновых дислокаций.

Теперь получим двумерную часть метрики краевой дислокации в конформной калибровке (36.83) на больших расстояниях $r \gg h$. Из уравнения (36.113) вытекает равенство дифференциалов

$$dw = \left(\frac{z-h}{z} \right)^\theta dz.$$

Поэтому метрика на больших расстояниях принимает вид

$$dl^2 = dw d\bar{w} = \left(\frac{(z-h)(\bar{z}-h)}{z\bar{z}} \right)^\theta dz d\bar{z} \approx \left(1 - \frac{2B \cos \theta}{r} \right) (dr^2 + r^2 d\varphi^2), \quad (36.116)$$

где $B = h\theta$ – модуль нормированного вектора Бюргерса, и мы учли только первую поправку. Эта метрика отличается от метрики (36.77) для краевой дислокации, полученной в рамках теории упругости, потому что записана в конформной, а не в упругой калибровке. Кроме того вектор Бюргерса в рассматриваемом случае параллелен оси y , а в предыдущем – оси x .

Сравним конформное отображение для диполя из двух клиновых дислокаций с конформным отображением (36.109) в дипольном приближении, возникающем при $\Theta = 0$ и $B \neq 0$,

$$w = z - B \ln z + i\pi B = z - B(\ln|z| + i \arg z) + i\pi B, \quad (36.117)$$

где, для определенности, мы зафиксировали постоянные и для $\ln z$ выбрали голоморфную ветвь $0 < \arg z < 2\pi$.

Чтобы понять как устроено конформное отображение (36.117), посмотрим, во что отображается вещественная прямая $\operatorname{im} z = 0$. При движении вдоль луча $\arg z = \pi$ от $-\infty$ до 0 точка на плоскости w также движется вдоль вещественной оси от $-\infty$ до ∞ . Продолжая движение точки z от 0 до B ее образ движется вдоль прямой, параллельной вещественной оси w , из бесконечности до точки

$$w_2 := B - B \ln B + i\pi B.$$

При дальнейшем движении точки z вдоль вещественной оси от B до ∞ точка w делает разворот и снова уходит в бесконечность. Соответствующее конформное отображение показано на рис. 36.14. Поскольку функция (36.117) голоморфна на луче

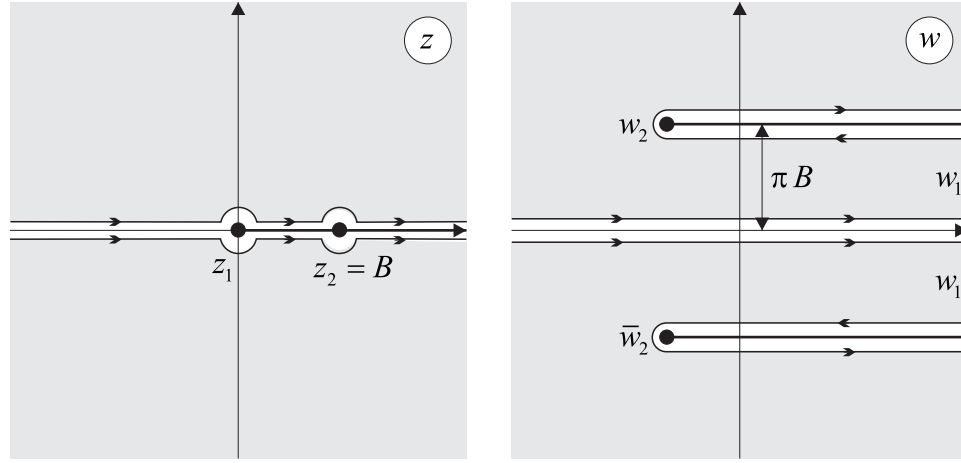


Рис. 36.14: Конформное отображение в дипольном приближении (краевая дислокация).

$\arg z = \pi$, то ее можно продолжить на всю плоскость z с разрезом вдоль неотрицательной части вещественной оси, как показано на рисунке. Начало координат $z_1 = 0$ отображается в бесконечно удаленную точку $w_1(z_1) = \infty$. Это конформное отображение имеет менее прозрачный физический смысл, зато является более универсальным в том смысле, что не зависит от деталей распределения клиновых дислокаций на малых расстояниях. Сравнивая данное конформное отображение с конформным отображением на рис. 36.13, мы видим, что оно получается при стремлении точки w_1 к бесконечности.

Квадрупольная дислокация

Чистое квадрупольное приближение можно реализовать с помощью четырех клиновых дислокаций:

$$\begin{array}{llll} z_1 = 0, & z_2 = h & z_3 = l & z_4 = l + h, \\ \theta_1 = -\theta, & \theta_2 = \theta, & \theta_3 = \theta & \theta_4 = -\theta, \end{array}$$

где h, l и θ – некоторые положительные постоянные. Для определенности, предположим, что $l > h$ и $0 < \theta < 1$. Для такого распределения клиновых дислокаций только квадрупольный момент отличен от нуля:

$$\Theta = 0, \quad B = 0, \quad M = \theta lh.$$

Соответствующее конформное отображение задается интегралом Кристоффеля–Шварца

$$w = \int_0^z d\zeta \left[\frac{(\zeta - h)(\zeta - l)}{\zeta(\zeta - l - h)} \right]^\theta, \quad (36.118)$$

где мы положили $\phi = 0$ и $C = 0$. Его нельзя вычислить в элементарных функциях, но можно проанализировать качественно. Расположение вершин многоугольника в

плоскости w задается сходящимися интегралами:

$$\begin{aligned} w_2 &= h e^{i\pi\theta} \int_0^1 dx \left[\frac{(1-x)(p-x)}{x(p+1-x)} \right]^\theta, \\ w_3 &= w_2 + (l-h) \int_0^1 dx \left[\frac{x(1-x)}{(q+x)(q+1-x)} \right]^\theta, \\ w_4 &= w_3 + h e^{-i\pi\theta} \int_0^1 dx \left[\frac{(1-x)(p-x)}{x(p+1-x)} \right]^\theta, \end{aligned}$$

где

$$p := \frac{l}{h} > 1, \quad q := \frac{h}{l-h} > 1.$$

При этом $w(z_1) = 0$. Отсюда следует, что четвертая вершина лежит на вещественной оси: $\operatorname{im} w_4 = 0$. Данное конформное отображение продолжается в нижнюю полуплоскость через луч $\arg z = \pi$. Соответствующее конформное отображение показано на рис. 36.15. Оно имеет ясную физическую интерпретацию: квадрупольная дислокация

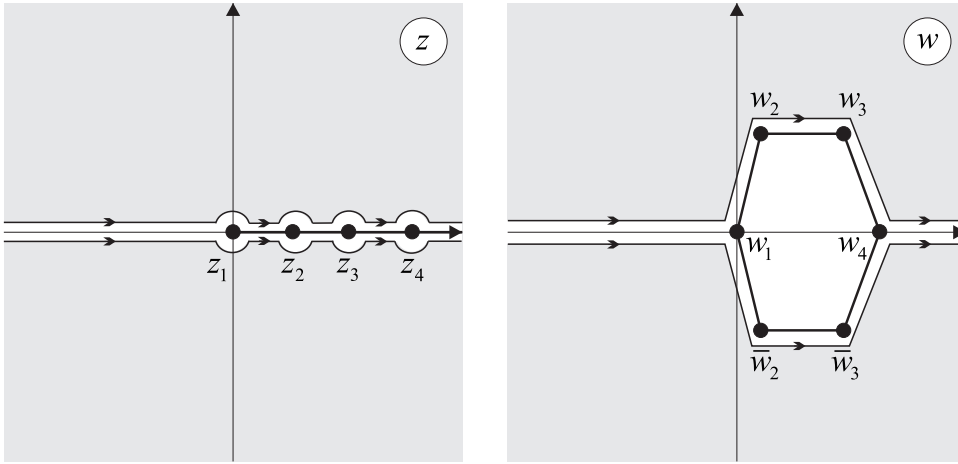


Рис. 36.15: Конформное отображение для четырех клиновых дислокаций, представляющих квадрупольную дислокацию.

возникает после удаления шестиугольной призмы из бесконечной среды и склейки точек разреза, лежащих на одной вертикали.

Теперь исследуем квадрупольное приближение в интеграле Кристоффеля–Шварца в общем случае для произвольного распределения вершин. Пусть $\Theta = 0$, $B = 0$, $\operatorname{im} M = 0$, $M > 0$. Положим $\phi = 0$ и $C = 0$. Тогда конформное отображение (36.109) для произвольного распределения клиновых дислокаций в квадрупольном приближении принимает вид

$$w = z - \frac{M}{z}. \quad (36.119)$$

Чтобы описать данное конформное отображение, посмотрим во что переходят окружности $z = r e^{i\varphi}$ на плоскости w :

$$w = r e^{i\varphi} - \frac{M}{r} e^{-i\varphi} = \left(r - \frac{M}{r} \right) \cos \varphi + i \left(r + \frac{M}{r} \right) \sin \varphi, \quad r = \text{const.}$$

Перепишем данное уравнение для вещественной и мнимой части

$$\begin{aligned} u &= \left(r - \frac{M}{r}\right) \cos \varphi, \\ v &= \left(r + \frac{M}{r}\right) \sin \varphi, \end{aligned} \quad \Rightarrow \quad \left(\frac{u}{a}\right)^2 + \left(\frac{v}{b}\right)^2 = 1,$$

где введены постоянные

$$a := r - \frac{M}{r}, \quad b := r + \frac{M}{r}.$$

Отсюда следует, что конформное отображение (36.119) переводит окружности в эл-

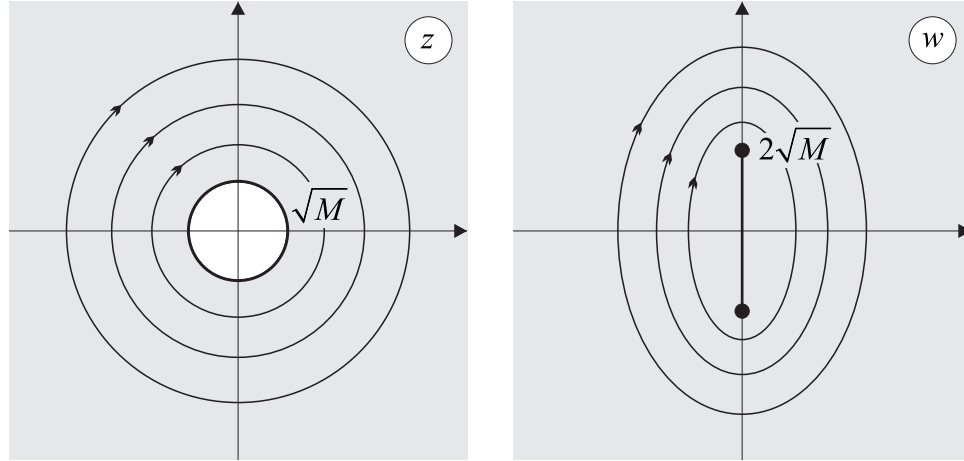


Рис. 36.16: Конформное отображение для квадрупольного приближения.

липсы, как показано на рис. 36.16. Окружности, радиус которых превышает $\sqrt{M}$ отображаются в эллипсы с сохранением ориентации, т.е. проходятся в одну сторону, и внешняя часть комплексной плоскости $|z| > \sqrt{M}$ отображается на всю плоскость w с разрезом вдоль мнимой оси. Внутренние окружности $0 < |z| < \sqrt{M}$ также отображаются в эллипсы, которые тоже покрывают всю плоскость w . Мы видим, что функция (36.119) отображает всю плоскость z на риманову поверхность, состоящую из двух экземпляров плоскости w , которые сшиты вдоль разреза. Читатель без труда может представить себе создание квадрупольной дислокации.

36.8.4 Непрерывное распределение линейных дислокаций

Одним из основных преимуществ геометрического подхода является возможность описания непрерывного распределения дислокаций и дисклинаций. Для простоты формул мы будем использовать не упругую калибровку, а конформную калибровку для двумерной части метрики (36.83).

Клиновые дислокации

В качестве первого примера рассмотрим вращательно симметричное распределение параллельных клиновых дислокаций (см. раздел 36.8.1). Предположим, что они однородно распределены внутри диска радиуса a , перпендикулярного линиям дислокаций, со следующим источником

$$T_{zz} := \frac{2\rho}{\sqrt{g^{(2)}}}, \quad (36.120)$$

где мы выписали единственную нетривиальную компоненту и

$$\rho(r) := \begin{cases} q, & r \leq a, \\ 0, & r > a. \end{cases}$$

Нормированный полный угол дефицита для такого распределения равен

$$\Theta = \frac{1}{2}qa^2.$$

Трехмерные уравнения Эйнштейна (36.78) с источником (36.120) анализируются так же, как и для параллельных клиновых дислокаций в разделе 36.8.1. В результате для конформного множителя ϕ двумерной части метрики получаем единственное уравнение

$$\Delta\phi = \rho(r).$$

Общее решение данного уравнения имеет вид

$$\phi(r) = \frac{1}{2\pi} \int d\vec{s} \rho(\vec{s}) \ln|\vec{s} - \vec{r}|.$$

Интеграл в правой части берется:

$$\phi(r) = \begin{cases} \frac{q}{2}a^2 \ln a - \frac{qa^2}{4} + \frac{qr^2}{4}, & r \leq a, \\ \frac{q}{2}a^2 \ln r, & r > a. \end{cases} \quad (36.121)$$

Вне дислокаций интервал принимает вид

$$dl^2 = r^{qa^2}(dr^2 + r^2d\varphi^2). \quad (36.122)$$

Это значит, что рассматриваемое непрерывное распределение клиновых дислокаций снаружи имеет тот же вид, что и одна клиновая дислокация (36.88) с углом дефицита Θ . Поэтому траектории фононов на больших расстояниях выглядят так же, как для одной клиновой дислокации. Этот результат трудно предугадать заранее, т.к. уравнения Эйнштейна нелинейны, и принципа суперпозиции для решений в общем случае нет. Заметим, что метрика (36.122) определяет точное решение трехмерных уравнений Эйнштейна (после добавления dz^2).

Краевые дислокации

В разделе 36.7.3 мы доказали, что краевая дислокация представляет собой диполь, состоящий из двух клиновых дислокаций противоположного знака. Для определенности выберем расположение дислокаций в том виде (36.112), как это было сделано при анализе конформных отображений. Тогда уравнения Эйнштейна сведутся к уравнению Пуассона для конформного множителя

$$\Delta\phi = 2\pi\theta\delta(\mathbf{r} - \mathbf{h}) - 2\pi\theta\delta(\mathbf{r}), \quad \theta > 0, \quad (36.123)$$

где мы ввели двумерный вектор $\mathbf{h} := (h, 0)$, и необходимо рассмотреть предел $h \rightarrow 0$ при условии $2\pi|\mathbf{B}| = 2\pi h\theta = \text{const}$. Для вычисления предела преобразуем правую

часть уравнения. Пусть $f(\mathbf{r}) \in \mathcal{D}'(\mathbb{R}^2)$ – пробная функция [3]. Тогда справедливы равенства:

$$\int d\mathbf{r} f(\mathbf{r}) [\delta(\mathbf{r} - \mathbf{h}) - \delta(\mathbf{r})] = f(\mathbf{h}) - f(0) \approx (\mathbf{h} \nabla f)|_{\mathbf{r}=0} = - \int d\mathbf{r} f(\mathbf{r}) (\mathbf{h} \nabla \delta(\mathbf{r})),$$

где ∇ – двумерный градиент. Поэтому в пределе $h \rightarrow 0$ уравнение Пуассона (36.123) принимает вид

$$\Delta \phi = -2\pi\theta(\mathbf{h} \nabla \delta(\mathbf{r})). \quad (36.124)$$

Таким образом, получено уравнение для одной краевой дислокации, расположенной в начале координат.

Если мы имеем непрерывное распределение краевых дислокаций, то после замены δ -функции на некоторое распределение уравнение (36.124) примет вид

$$\Delta \phi = -2\pi \nabla \beta, \quad (36.125)$$

где $\beta(\mathbf{r})$ – плотность распределения краевых дислокаций. В каждой точке вектор β отличается от вектора Бюргерса соответствующей краевой дислокации направлением, а их модули равны. Для краевой дислокации (36.112) вектор $\mathbf{h}$ параллелен вектору ∂_x , а вектор Бюргерса (36.115) – вектору ∂_y , т.е. они перпендикулярны.

Нетрудно проверить, что к уравнению Пуассона (36.125) приводит следующий источник в трехмерных уравнениях Эйнштейна

$$T_{zz} = -\frac{4\pi \nabla \beta}{\sqrt{g^{(2)}}},$$

остальные компоненты которого равны нулю.

Рассмотрим простейший случай. Пусть задано распределение $\beta(\mathbf{r}) := (\beta_x(\mathbf{r}), \beta_y(\mathbf{r}))$. Предположим, что векторы Бюргерса всех краевых дислокаций направлены вдоль оси y . Тогда $\beta_y = 0$, и мы выберем однородное распределение краевых дислокаций внутри круга радиуса a :

$$\beta_x = \begin{cases} \beta, & r \leq a, \\ 0, & r > a. \end{cases} \quad \beta = \text{const},$$

Для такого распределения

$$\nabla \beta = -\beta \cos \varphi \delta(a - r),$$

где мы продифференцировали ступеньку по радиусу и получили δ -функцию. Тогда решение уравнения (36.125) примет вид

$$\begin{aligned} \phi &= - \int_0^\infty dr' r' \int_0^{2\pi} d\varphi' [-\beta \cos \varphi' \delta(a - r')] \frac{1}{2} \ln[r^2 + r'^2 - 2rr' \cos(\varphi' - \varphi)] = \\ &= \begin{cases} -\pi\beta r \cos \varphi, & r \leq a, \\ -\frac{\pi\beta a^2 \cos \varphi}{r}, & r > a. \end{cases} \end{aligned} \quad (36.126)$$

На больших расстояниях от распределения дислокаций, $r \gg a$, пространство описывается асимптотически плоской метрикой

$$dl^2 = e^{-\frac{2\pi\beta a^2 \cos \varphi}{r}} (dr^2 + r^2 d\varphi^2) \approx \left(1 - \frac{2B \cos \varphi}{r}\right) (dr^2 + r^2 d\varphi^2), \quad (36.127)$$

где

$$B := \beta \pi a^2. \quad (36.128)$$

Мы видим, что модуль полного вектора Бюргерса равен плотности β , умноженной на площадь круга πa^2 , что интуитивно ясно. Метрика (36.127) совпадает с метрикой для краевой дислокации (36.116), полученной при рассмотрении конформных отображений. Таким образом, на больших расстояниях в первом приближении непрерывное распределение краевых дислокаций ведет себя так же, как одна краевая дислокация с полным вектором Бюргерса (36.128).

Принцип суперпозиции в данном случае работает только для первого члена разложения линейного элемента. Остальные поправки отличают непрерывное распределение краевых дислокаций от одной краевой дислокации. Вторая поправка легко находится в комплексных координатах. Действительно, конформный множитель (36.126) вне распределения дислокаций приводит к метрике

$$dl^2 = e^{-\frac{2B \cos \varphi}{r}} dz d\bar{z} = e^{-B(\frac{1}{z} + \frac{1}{\bar{z}})} dz d\bar{z}.$$

Ее можно записать в плоском виде $dl^2 = dw d\bar{w}$, где

$$w := \int^z d\zeta e^{-\frac{B}{\zeta}} = \int^z d\zeta \left(1 - \frac{B}{\zeta} + \frac{B^2}{2\zeta^2} - \dots \right) \approx z - B \ln z - \frac{B^2}{2z}.$$

Это разложение дает квадрупольную поправку $M = B^2/2$ к непрерывному распределению краевых дислокаций, что следует из разложения (36.109).

36.8.5 Разрезание и склеивание

Рассмотрим евклидово пространство $\mathbb{R}^3$, которое моделирует упругую среду без дефектов в равновесном состоянии. Произведем произвольные разрезы, деформируем среду каким-либо образом и затем произведем склейку, добавив или удалив часть вещества в случае необходимости. В результате возникнет дефект упругой среды – одна или несколько дислокаций. Этот процесс можно описать в явном виде с помощью координатных преобразований. Если задано преобразование координат, то можно вычислить индуцированный репер и метрику, соответствующие данной дислокации. Вне разрезов тензоры кручения и кривизны данного репера (метрики) равны нулю, т.к. в этих частях среды имеются только упругие деформации, которые описывают диффеоморфизмы. Следовательно, вне разрезов уравнения равновесия (36.5) удовлетворены с нулевой правой частью. Они также удовлетворяются на поверхностях склейки, поскольку по предположению склейка репера производится достаточно гладко. Таким образом, источники дислокаций в правой части уравнений равновесия могут возникнуть только на краях разрезов: на линиях (линейные дислокации) и в точках (точечные дислокации). Поскольку размерность носителей источников меньше трех, то они должны иметь δ -образный характер.

Это верно для отдельных дислокаций или их конечного числа. Для непрерывного распределения дислокаций нетривиальная правая часть уравнений равновесия может возникнуть на поверхностях или в объеме.

Ниже с помощью координатных преобразований мы опишем пример создания комбинированной клиновой и винтовой дислокации без анализа особенности на линии дислокации.

Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовой X, Y, Z и цилиндрической R, Φ, Z системами координат. Евклидова метрика имеет вид

$$ds^2 = dX^2 + dY^2 + dZ^2 = dR^2 + R^2 d\Phi^2 + dZ^2. \quad (36.129)$$

Совершим преобразование координат $R, \Phi, Z \mapsto f, \varphi, z$, заданное следующими формулами:

$$\begin{aligned} R &:= \frac{f}{\alpha}, \\ \Phi &:= \alpha\varphi, \\ Z &:= z - c\varphi, \quad \alpha > 0, \quad c = \text{const}, \end{aligned} \quad (36.130)$$

где радиальная координата обозначена буквой f , имея в виду дальнейшее преобразование координат $f \mapsto r = r(f)$. В новых координатах евклидова метрика примет вид

$$ds^2 = \frac{1}{\alpha^2} df^2 + (f^2 + c^2) d\varphi^2 + dz^2 - 2c d\varphi dz. \quad (36.131)$$

Эта метрика описывает дислокацию и, по предположению, определена во всем евклидовом пространстве за исключением оси z :

$$0 < f < \infty, \quad 0 < \varphi < 2\pi, \quad -\infty < z < \infty. \quad (36.132)$$

Данной области определения конечных координат отвечает следующая область определения исходных координат в евклидовом пространстве:

$$0 < R < \infty, \quad 0 < \Phi < 2\pi\alpha, \quad -\infty < z < \infty. \quad (36.133)$$

Обратим внимание на область определения угловой координаты. При $0 < \alpha < 1$ областью (36.133) является не все, а только часть евклидова пространства. Если $\alpha > 1$, то область определения выходит за пределы евклидова пространства $\mathbb{R}^3$.

Теперь можно описать процесс склейки. У нас есть отображение

$$y = (R, \Phi, Z) \mapsto x = (f, \phi, z).$$

В исходном пространстве, точки которого обозначены координатами y , мы отождествляем те точки среды, которые отображаются в одну точку x многообразия с дислокациями. В рассматриваемом случае – это точки, лежащие на полуплоскостях $\Phi = 0$ и $\Phi = 2\pi\alpha$, что следует из области определения (36.133). Следовательно, в исходном евклидовом пространстве мы делаем разрез по полуплоскости X, Z , край которой совпадает с осью Z , раздвигаем берега разреза на угол $2\pi\theta$ при $\theta < 0$, где $\alpha := 1 + \theta$, вставляем в образовавшуюся полость клин того же вещества. Кроме того перед склейкой точки нижнего берега разреза сдвигаются параллельно вдоль оси Z на вектор Бюргерса $b := 2\pi c$, что следует из формул преобразования координат (36.130). Описанный процесс создания комбинированной клиновой и винтовой дислокации показан на рис. 36.17.

Таким образом, при $c = 0$ метрика (36.131) описывает клиновую дислокацию, а при $\alpha = 1$ – винтовую (36.62). Строго говоря, полный вектор Бюргерса для комбинированной дислокации будет переменным. У него есть постоянная компонента $b := 2\pi c$ вдоль оси z , соответствующая винтовой дислокации, и переменная компонента $-2\pi\theta r$, при $\theta \ll 1$, которая соответствует клиновой дислокации.

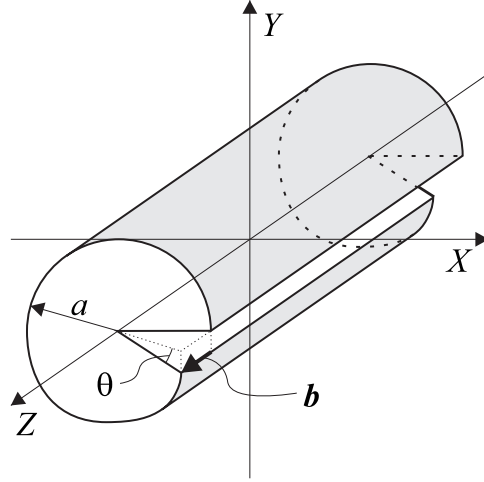


Рис. 36.17: Комбинированная клиновая и винтовая дислокация с углом дефицита θ и вектором Бюргерса $\mathbf{b}$.

Мы построили метрику, описывающую комбинированную клиновую и винтовую дислокацию, с помощью процедуры разрезания и склеивания. Для того, чтобы удовлетворить уравнениям теории упругости, необходимо переписать полученную метрику в упругой калибровке, рассмотренной в разделе 36.5. Метрике (36.131) поставим в соответствие репер

$$e_\mu^i = \begin{pmatrix} \frac{1}{\alpha} & 0 & 0 \\ 0 & f & -c \\ 0 & 0 & 1 \end{pmatrix}, \quad (36.134)$$

для которого линейное приближение по вектору смещения имеет вид (36.41). Теперь сделаем замену радиальной координаты $f \mapsto r$ так, чтобы удовлетворить упругой калибровке. После подстановки репера (36.134) в калибровочное условие (36.43) для функции $f(r)$ возникает уравнение Эйлера (36.95) такое же, что и для клиновой дислокации. Его решение при граничных условиях (36.97) имеет тот же вид

$$f = \frac{\alpha}{\gamma a^{\gamma-1}} r^\gamma, \quad (36.135)$$

где

$$\gamma := -\theta \hat{b} + \sqrt{\theta^2 \hat{b}^2 + 1 + \theta}, \quad \hat{b} := \frac{\sigma}{2(1 - \sigma)},$$

где постоянную $\hat{b}$ мы поместили шляпкой, чтобы отличать ее от модуля вектора Бюргерса. В итоге получаем метрику в упругой калибровке

$$ds^2 = \left(\frac{r}{a}\right)^{2\gamma-2} dr^2 + \left[\left(\frac{r}{a}\right)^{2\gamma-2} \frac{\alpha^2 r^2}{\gamma^2} + c^2\right] d\varphi^2 + dz^2 - 2c d\varphi dz, \quad (36.136)$$

которая описывает комбинированную клиновую и винтовую дислокацию.

Комбинированную клиновую и винтовую дислокацию нетрудно описать в рамках теории упругости. Фактически, все вычисления уже проделаны в разделах 36.7.1 и 36.7.2, где были по отдельности рассмотрены винтовая и клиновая дислокации. Чтобы найти векторное поле смещений в рассматриваемом случае, достаточно просто сложить смещения для клиновой и винтовой дислокаций, т.к. мы рассматриваем линейную теорию упругости.

Напомним, что смещения для клиновой дислокации равны (36.71)

$$u_{\hat{r}}^{\text{wedge}} = -\theta \frac{1-2\sigma}{2(1-\sigma)} r \ln \frac{r}{ea}, \quad u_{\hat{\varphi}}^{\text{wedge}} = -\theta r \varphi, \quad u_{\hat{z}}^{\text{wedge}} = 0. \quad (36.137)$$

где мы удержали только линейные по θ слагаемые, которых достаточно в линейном приближении.

Для винтовой дислокации смещения имеют вид (36.60)

$$u_{\hat{r}}^{\text{screw}} = 0, \quad u_{\hat{\varphi}}^{\text{screw}} = 0, \quad u_{\hat{z}}^{\text{screw}} = c\varphi = \frac{b}{2\pi}\varphi.$$

Полное векторное поле смещений равно сумме

$$u = u^{\text{wedge}} + u^{\text{screw}}$$

и приводит к индуцированной метрике

$$ds^2 = \left(1 + \theta \frac{1-2\sigma}{1-\sigma} \ln \frac{r}{a}\right) dr^2 + \left[r^2 \left(1 + \theta \frac{1-2\sigma}{1-\sigma} \ln \frac{r}{a} + \theta \frac{1}{1-\sigma}\right) + c^2\right] d\varphi^2 + dz^2 - 2c d\varphi dz, \quad (36.138)$$

где мы, опять же, удержали только линейные по θ слагаемые.

Метрика для комбинированной клиновой и винтовой дислокации (36.138) применима только для малых углов дефицита, $\theta \ll 1$, и вблизи края цилиндра, $r \sim a$, где малы относительные смещения. Точная метрика (36.136), полученная в рамках геометрической теории дефектов, имеет более простой вид, применима для всех углов дефицита $\theta > -1$ и для всех значений радиуса $0 < r < a$. Нетрудно проверить, что в линейном приближении по θ точная метрика воспроизводит метрику (36.138), полученную в рамках теории упругости. Вычисления здесь аналогичны тем, которые были описаны для клиновой дислокации в разделе 36.8.2. Это говорит о преимуществе геометрического подхода.

36.9 Рассеяние фононов на клиновых дислокациях

Фононами в твердом теле называются частицы, возникающие при вторичном квантовании волновых уравнений для упругих колебаний. Поэтому, строго говоря, задача о рассеянии фононов на дислокациях является квантовомеханической. Ниже мы рассмотрим только классические аспекты этой задачи.

Допустим, что некоторый волновой процесс, например, сжатие среды с дефектами, описывается вещественной или комплекснозначной функцией $w(t, x)$. Как мы установили в предыдущих разделах, наличие дислокаций приводит к возникновению нетривиальной метрики $g_{\alpha\beta}$. Поэтому волновое уравнение для функции $w(t, x)$ в среде со статическим распределением дислокаций имеет вид

$$g^{\alpha\beta} \tilde{\nabla}_{\alpha} \tilde{\nabla}_{\beta} w = 0, \quad (36.139)$$

где $g^{\alpha\beta}$ – метрика, обратная к метрике

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & -g_{\mu\nu} \end{pmatrix} \quad (36.140)$$

и $\tilde{\nabla}$ – ковариантная производная с символами Кристоффеля. Выше мы нумеруем четырехмерные координаты греческими буквами из начала алфавита $\{x^{\alpha}\} :=$

$\{x^0, x^1, x^2, x^3\}$, а буквами из середины греческого алфавита – только пространственные координаты $\{x^\mu\} := \{x^1, x^2, x^3\}$. По четырехмерной метрике (36.140) можно вычислить символы Кристоффеля (4.20), которые определяют систему нелинейных обыкновенных уравнений для экстремалей $x^\alpha(\tau)$ (линий экстремальной длины, которые в римановой геометрии совпадают с геодезическими, см. раздел 16.2). Для метрики блочно диагонального вида (36.140) эти уравнения расщепляются:

$$\ddot{x}^0 = 0, \quad (36.141)$$

$$\ddot{x}^\mu = -\tilde{\Gamma}_{\nu\rho}^{\mu} \dot{x}^\nu \dot{x}^\rho, \quad (36.142)$$

где точки обозначают дифференцирование по каноническому параметру и $\tilde{\Gamma}_{\nu\rho}^{\mu}$ – трехмерные символы Кристоффеля, построенные по трехмерной метрике $g_{\mu\nu}$, которая, напомним, для статического распределения дефектов зависит только от пространственных координат. Пусть $\{x^\alpha(\tau)\}$ – произвольная экстремаль в четырехмерном пространстве-времени. Тогда для метрики (36.140) ее естественная проекция на пространство $\{x^\alpha(\tau)\} \mapsto \{0, x^\mu(\tau)\}$ также является экстремалью, но уже для пространственной части метрики $g_{\mu\nu}$.

Уравнения для экстремалей (36.141), (36.142) инвариантны относительно линейных преобразований канонического параметра τ . Поэтому из уравнения (36.141) следует, что, не ограничивая общности, канонический параметр можно отождествить с временем: $\tau = ct = x^0$, где c – скорость распространения колебаний.

При анализе асимптотик решений волнового уравнения (36.139) полезны представления о фронтах волны и лучах, как это делается в геометрической оптике [169]. Мы не будем останавливаться на математических аспектах этого подхода, который нетривиален и сложен [312], и дадим лишь физическое описание. В эйкональном (высокочастотном) приближении (см. раздел 16.7) фононы распространяются вдоль лучей, совпадающих с нулевыми экстремалью для четырехмерной метрики $g_{\alpha\beta}$. Форма лучей, которые мы отождествляем с траекториями фононов, определяется трехмерной метрикой $g_{\mu\nu}$. Это значит, что в эйкональном приближении траектории поперечных и продольных фононов (см. раздел 36.11) в среде с дефектами одинаковы и определяются уравнениями (36.142). Отличие сводится к тому, что скорости распространения поперечных и продольных фононов различны и равны, соответственно, c_τ и c_L .

Заметим, что экстремали $x(t)$ являются бихарактеристиками волнового уравнения (36.139) (см. раздел 16.6).

Другой подход к анализу экстремалей основан на принципе наименьшего действия. В разделе 16.5 было показано, что уравнения (36.142) следуют из лагранжиана

$$L = \frac{1}{2} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu, \quad (36.143)$$

который равен кинетической энергии точечной частицы с единичной массой, движущейся по риманову многообразию. Конечно, при этом мы отождествили канонический параметр с временной координатой. Мы видим, что и фононы, и точечные частицы движутся вдоль одних и тех же траекторий. Разница сводится только к скорости движения. Хотя член с потенциальной энергией в лагранжиане (36.143) отсутствует, движение частиц нетривиально, т.к. метрика явно зависит от точки многообразия. Уравнение (36.142) представляет собой второй закон Ньютона, в котором сила квадратична по скорости. В связи с этим движение фононов в присутствии дислокаций сильно отличается от ньютонова движения частиц в потенциальном поле. Например, в дальнейшем мы покажем наличие нового типа замкнутых траекторий.

В случаях, когда симметрий задачи недостаточно для нахождения достаточного числа интегралов движения, на помощь приходит уравнение Гамильтона–Якоби (см. раздел 16.5). Лагранжиан (36.143) приводит к гамильтониану

$$H = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu,$$

где $g^{\mu\nu}$ – обратная метрика и $p_\mu := \partial L / \partial \dot{x}^\mu$ – канонический импульс. Тогда уравнение Гамильтона–Якоби для функции действия $S(x, t)$ принимает вид

$$\frac{\partial S}{\partial t} + \frac{1}{2} g^{\mu\nu} \frac{\partial S}{\partial x^\mu} \frac{\partial S}{\partial x^\nu} = 0. \quad (36.144)$$

Для укороченной функции действия (см. раздел 21.8)

$$W(x) := S(x, t) + Et, \quad E = \text{const},$$

уравнение Гамильтона–Якоби упрощается

$$\frac{1}{2} g^{\mu\nu} \frac{\partial W}{\partial x^\mu} \frac{\partial W}{\partial x^\nu} = E. \quad (36.145)$$

Здесь E обозначает полную (в рассматриваемом случае кинетическую) энергию фона. Это уравнение решено в разделе 36.9.3 для произвольного распределения N параллельных клиновых дислокаций.

36.9.1 Рассеяние на одной клиновой дислокации

При наличии одной клиновой дислокации индуцированная метрика упругой среды в цилиндрических координатах r, φ, z в упругой калибровке имеет вид

$$g_{\mu\nu} = \begin{pmatrix} \left(\frac{r}{a}\right)^{2\gamma-2} & 0 & 0 \\ 0 & \frac{\alpha^2 r^2}{\gamma^2} \left(\frac{r}{a}\right)^{2\gamma-2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (36.146)$$

где нетривиальная часть метрики в плоскости r, φ была получена ранее (36.99). Обратная метрика также является блочно диагональной

$$g^{\mu\nu} = \begin{pmatrix} \left(\frac{a}{r}\right)^{2\gamma-2} & 0 & 0 \\ 0 & \frac{\gamma^2}{\alpha^2 r^2} \left(\frac{a}{r}\right)^{2\gamma-2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Символы Кристоффеля для метрики (36.146) вычисляются по формулам (4.20). В результате получаем, что только четыре символа Кристоффеля отличны от нуля:

$$\tilde{\Gamma}_{rr}^r = \frac{\gamma-1}{r}, \quad \tilde{\Gamma}_{\varphi\varphi}^r = -\frac{\alpha^2 r}{\gamma}, \quad \tilde{\Gamma}_{r\varphi}^\varphi = \tilde{\Gamma}_{\varphi r}^\varphi = \frac{\gamma}{r}.$$

В эйкональном приближении фононы распространяются вдоль экстремалей $x^\mu(t)$ (здесь мы положили скорость фононов равной единице, $c = 1$), определяемых уравнениями (36.142). В рассматриваемом случае эти уравнения принимают вид

$$\ddot{r} = -\frac{\gamma-1}{r}\dot{r}^2 + \frac{\alpha^2}{\gamma}r\dot{\varphi}^2, \quad (36.147)$$

$$\ddot{\varphi} = -\frac{2\gamma}{r}\dot{r}\dot{\varphi}, \quad (36.148)$$

$$\ddot{z} = 0, \quad (36.149)$$

где точка обозначает дифференцирование по времени t . Из последнего уравнения следует, что фононы движутся вдоль оси z с постоянной скоростью, что соответствует трансляционной инвариантности метрики вдоль z . Это значит, что рассеяние фононов на клиновой дислокации, как и следовало ожидать, сводится к двумерной задаче в плоскости r, φ .

Систему уравнений для $r(t)$ и $\varphi(t)$ (36.147) и (36.148) можно явно проинтегрировать. С этой целью найдем два первых интеграла. Во-первых, для любой метрики уравнения для экстремалей имеют интеграл

$$g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = \text{const.}$$

Отсюда следует равенство

$$r^{2\gamma-2}\dot{r}^2 + \frac{\alpha^2}{\gamma^2}r^{2\gamma}\dot{\varphi}^2 = C_0 = \text{const} > 0. \quad (36.150)$$

Во-вторых, инвариантность метрики относительно вращений вокруг оси z приводит к существованию дополнительного интеграла. Он строится следующим образом. Инвариантности метрики соответствует вектор Киллинга, который в цилиндрических координатах имеет простой вид $K = \partial_\varphi$. Непосредственной проверкой можно доказать, что

$$g_{\mu\nu}K^\mu\dot{x}^\nu = \text{const.}$$

В рассматриваемом случае это приводит к закону сохранения

$$r^{2\gamma}\dot{\varphi} = C_1 = \text{const.} \quad (36.151)$$

Исследуем форму экстремалей $r = r(\varphi)$. Из уравнений (36.150), (36.151) можно найти первые производные

$$\dot{r} = \pm r^{-2\gamma+1} \sqrt{C_0 r^{2\gamma} - \frac{\alpha^2}{\gamma^2} C_1^2}, \quad (36.152)$$

$$\dot{\varphi} = C_1 r^{-2\gamma}. \quad (36.153)$$

Допустимые значения радиальной координаты r , при которых подкоренное выражение неотрицательно, будут найдены позже. Отсюда следует уравнение, определяющее форму нерадиальных ($C_1 \neq 0$) экстремалей,

$$\frac{dr}{d\varphi} = \frac{\dot{r}}{\dot{\varphi}} = \pm r \sqrt{\frac{C_0}{C_1^2} r^{2\gamma} - \frac{\alpha^2}{\gamma^2}}. \quad (36.154)$$

Это уравнение легко интегрируется, и в результате получаем явную формулу для экстремалей:

$$\left(\frac{r}{r_m}\right)^{2\gamma} \sin^2[\alpha(\varphi + \varphi_0)] = 1, \quad (36.155)$$

где

$$r_m := \left(\frac{C_1 \alpha}{\sqrt{C_0} \gamma}\right)^{1/\gamma} = \text{const} > 0, \quad \varphi_0 = \text{const}.$$

Постоянная r_m положительна и определяет минимальное расстояние, на которое экстремаль подходит к ядру дислокации, т.е. $r \geq r_m$. Только при этих значениях r подкоренное выражение в (36.154) неотрицательно. Постоянная интегрирования φ_0 произвольна и соответствует инвариантности задачи относительно вращений вокруг оси дислокации.

Уравнения для экстремалей (36.147), (36.148) имеют также вырожденные решения

$$\frac{1}{\gamma} r^\gamma = \pm \sqrt{C_0} (t + t_0), \quad \varphi = \text{const}, \quad t_0 = \text{const}, \quad (36.156)$$

которые соответствуют радиальному движению фононов. Такие траектории неустойчивы в том смысле, что вблизи них нет нерадиальных экстремалей.

Отметим, что экстремали в виде окружности ($r = \text{const}$) отсутствуют как следствие уравнения (36.147), хотя интегралы движения (36.150), (36.151) и допускают такое решение. Это связано с тем, что при получении первого интеграла (36.150) уравнение (36.147) умножается на $\dot{r}$.

Проанализируем форму нерадиальных экстремалей (36.155). Для каждой экстремали радиус r сначала убывает от бесконечности до минимального значения r_m , а затем возрастает от r_m до бесконечности. При этом, не ограничивая общности, можно считать, что аргумент синуса в (36.155) меняется 0 до π . Это следует из того, что в среде до создания дислокаций экстремальными являются прямые линии. Таким образом, получаем область изменения полярного угла

$$0 \leq \varphi + \varphi_0 \leq \frac{\pi}{\alpha}. \quad (36.157)$$

Это значит, что экстремаль приходит из бесконечности под углом $-\varphi_0$ и уходит на бесконечность под углом $\pi/\alpha - \varphi_0$. Это соответствует углу рассеяния

$$\chi := \pi - \frac{\pi}{\alpha} = \frac{\pi\theta}{1 + \theta}. \quad (36.158)$$

Отметим, что угол рассеяния фононов зависит только от угла дефицита θ и не зависит от упругих свойств среды. Угол рассеяния имеет простую физическую интерпретацию. Для положительных θ среда разрезается и раздвигается. В образовавшуюся полость вставляется клин из той же среды без упругих напряжений. Затем происходит склейка и клин сжимается под действием упругих сил. Коэффициент сжатия среды для каждой окружности с центром в начале координат из соображений симметрии равен $1/(1 + \theta)$. Поэтому угол рассеяния равен половине угла дефицита, умноженному на коэффициент сжатия:

$$\chi = \frac{2\pi\theta}{2} \times \frac{1}{1 + \theta}.$$

При $\theta = 0$ дислокация отсутствует и угол рассеяния, как и следовало ожидать, равен нулю.

При положительных углах дефицита угол рассеяния положителен, что соответствует отталкиванию фононов от клиновой дислокации. Соответствующие экстремали показаны на рис. 36.18*a*, и они имеют асимптоты при $r \rightarrow \infty$. Отметим, что

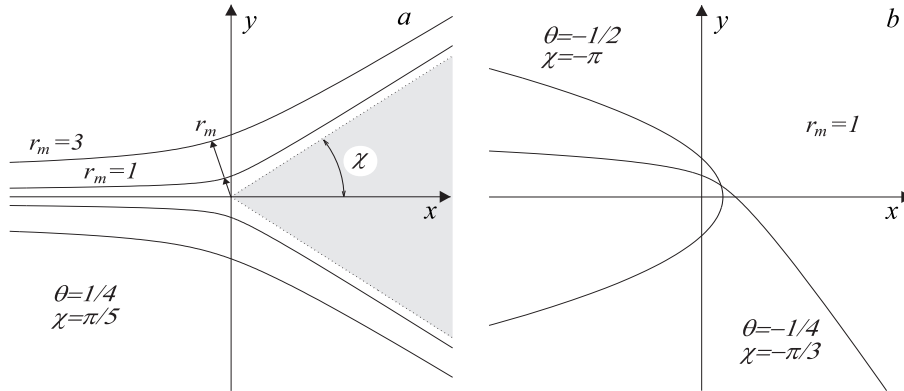


Рис. 36.18: (а) Экстремали для клиновой дислокации с положительным ($\theta > 0$) углом дефицита. Показаны две экстремали и их отражение относительно оси x для одного значения $\theta > 0$, но разных r_m . (б) Экстремали для клиновой дислокации с отрицательным $-1/2 \leq \theta < 0$ углом дефицита. Показаны две экстремали для одинаковых значений r_m , но разных углах дефицита $\theta < 0$.

при положительных углах дефицита не каждые две точки плоскости r, φ можно соединить экстремалью, т.е. справа от клиновой дислокации имеется зона, в которую фононы, падающие слева не попадают вовсе (она показана темным).

Все экстремали на рисунках в данном разделе построены численно для $\gamma = \alpha$. Значения угла дефицита θ , угла рассеяния χ и минимального расстояния до оси дислокации r_m указаны на рисунках. Для определенности мы выбрали $\varphi_0 = \pi$, что соответствует падению фононов слева.

При отрицательных углах дефицита угол рассеяния отрицателен и определяется той же формулой (36.158). Это соответствует притяжению фононов к оси дислокации. На рис. 36.18*b* показаны две экстремали с одинаковым параметром r_m , но для двух дислокаций с разными углами дефицита. При $-1/2 < \theta < 0$ угол рассеяния меняется в пределах от 0 до $-\pi$ (см. рис. 36.18*b*). При $\theta = -1/2$ угол рассеяния равен $-\pi$, т.е. фонон возвращается, сделав один облет оси дислокации. Отметим, что при отрицательных углах дефицита фононы не имеют асимптот при $r \rightarrow \infty$, т.е. при $\theta = -1/2$ фононы падают из бесконечности $x \rightarrow -\infty, y \rightarrow +\infty$ и возвращаются в бесконечность $x \rightarrow -\infty, y \rightarrow -\infty$.

Если угол дефицита достаточно мал ($-1 < \theta < -1/2$) то фонон делает один или несколько оборотов вокруг дислокации, а затем уходит в бесконечность. Примеры таких траекторий показаны на рисунках 36.19–36.21.

Рассмотрим асимптотику нерадиальных экстремалей при $r \rightarrow \infty$. В соответствии с уравнением (36.152) вдали от ядра дислокации

$$\dot{r} \approx \pm \sqrt{C_0} r^{-\gamma+1} \quad \Rightarrow \quad r \sim t^{1/\gamma}.$$

Отсюда следует, что зависимость радиуса от времени такая же, как и для радиальных экстремалей (36.156). Поскольку $\gamma > 0$, то бесконечное значение r соответствует бесконечному значению времени t . Это значит что плоскость r, φ с заданной метрикой (36.146) полна при $r \rightarrow \infty$. Начало координат (ось дислокации) является сингулярной

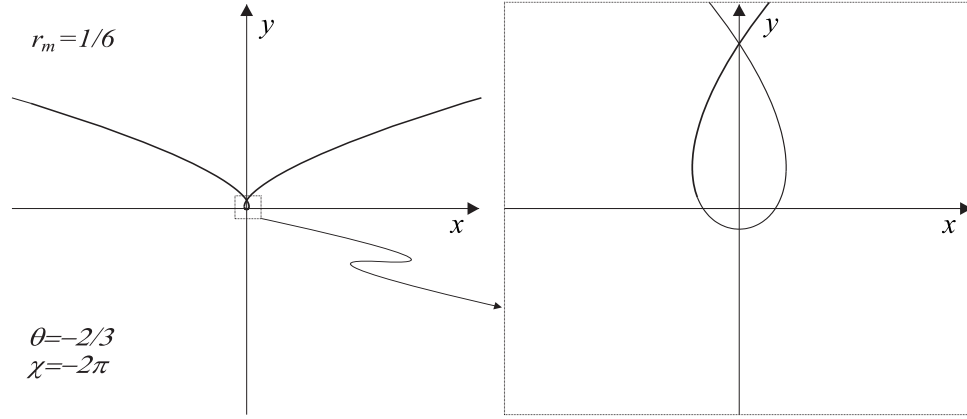


Рис. 36.19: При $\theta = -2/3$ экстремаль делает один оборот вокруг дислокации, а затем идет вперед в первоначальном направлении. Справа показана увеличенная часть траектории из квадрата вблизи оси дислокации.

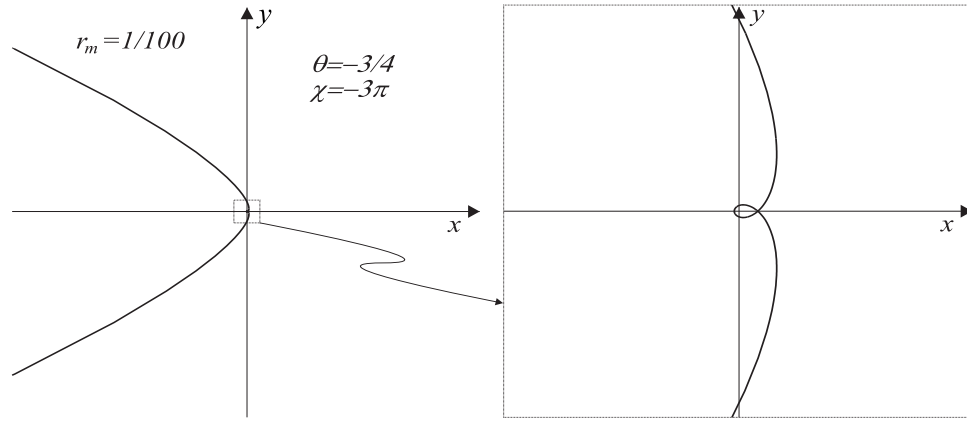


Рис. 36.20: При $\theta = -3/4$ экстремаль делает полтора оборота вокруг дислокации, а затем возвращается назад. Справа показана увеличенная часть траектории из квадрата вблизи оси дислокации.

точкой (коническая особенность). В нее попадают только радиальные экстремали при конечном значении времени.

Интегралы движения (36.150), (36.151) имеют простой физический смысл. Уравнения для экстремалей (36.142) следуют из вариационного принципа для лагранжиана (36.143), описывающего движение свободной безмассовой точечной частицы по риманову многообразию в нетривиальной метрикой $g_{\mu\nu}(x)$. При этом метрика рассматривается как заданное внешнее поле, по которому варьирование не проводится.

Энергия, соответствующая лагранжиану (36.143), равна

$$E = \frac{1}{2} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = \frac{1}{2} \left(\frac{r}{a} \right)^{2\gamma-2} \dot{r}^2 + \frac{1}{2} \frac{\alpha^2}{\gamma^2} \left(\frac{r}{a} \right)^{2\gamma-2} r^2 \dot{\varphi}^2 + \frac{1}{2} \dot{z}^2.$$

Если метрика, как в нашем случае, не зависит от времени явно, то энергия сохраняется ($E = \text{const}$) и для движения в плоскости r, φ ее численное значение пропорционально интегралу движения C_0 .

Для клиновой дислокации метрика не зависит от полярного угла φ , и лагранжиан инвариантен относительно вращений: $\varphi \mapsto \varphi + \text{const}$. Согласно первой теореме Нетер

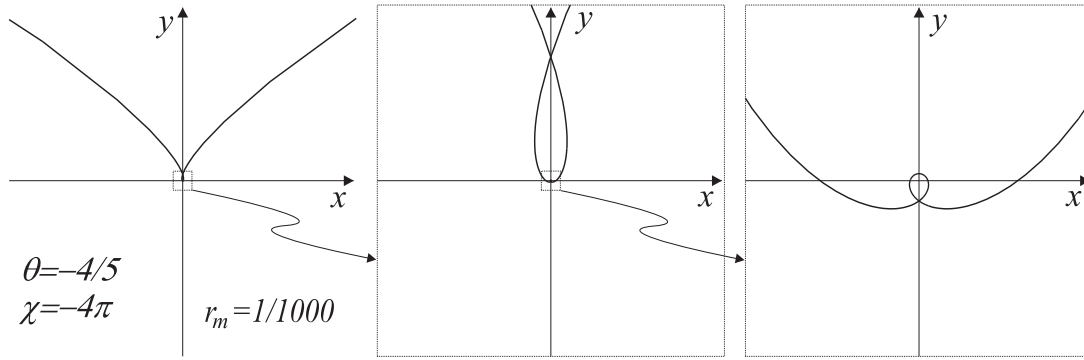


Рис. 36.21: При $\theta = -4/5$ экстремаль делает два оборота вокруг дислокации, а затем идет вперед в первоначальном направлении. Справа показана последовательно увеличенная часть траектории из двух квадратов вблизи оси дислокации.

(см. раздел 20.2) этой инвариантности соответствует сохранение углового момента

$$J = -\frac{\alpha^2}{\gamma^2} \left(\frac{r}{a}\right)^{2\gamma-2} r^2 \dot{\varphi} = \text{const.}$$

Отсюда следует, что постоянная интегрирования C_1 пропорциональна угловому моменту.

Отметим, что поведение экстремалей качественно отличается от траекторий точечных частиц, двигающихся в плоском пространстве с евклидовой метрикой $\delta_{\mu\nu}$ во внешнем потенциальном поле $U(x)$. Нетрудно доказать, что траектории точечных частиц, описываемых лагранжианом

$$L = \frac{1}{2} \delta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu - U$$

не могут совпадать с экстремальными (36.142) ни для какой функции $U(x)$.

36.9.2 Космические струны и множественные образы

Метрика для клиновой дислокации с отрицательным углом дефицита, $\theta < 0$, совпадает с пространственной частью метрики космической струны [313]. Интересно отметить, что метрика космической струны была найдена совершенно другим путем, как решение четырехмерных уравнений Эйнштейна в линейном приближении. Тем не менее метрика, по существу, одна и та же, и, следовательно, экстремали, найденные в предыдущем разделе, описывают также распространение световых лучей в пространстве-времени при наличии космической струны. Хорошо известным следствием наличия космической струны является возникновение двойного изображения звезды, находящейся за космической струной, см. рис. 36.22 слева. В настоящем разделе получена общая формула для углового разделения образов. Первоначально угловое разделение образов было получено для малых углов в системе координат, где экстремали являются прямыми линиями (т.е. до того, как создана дислокация). Мы получим общую формулу справедливую для всех значений углов и в системе координат, связанной с дислокацией.

Предположим, что звезда, струна и наблюдатель находятся на одной прямой (оси абсцисс x). Пусть струна находится в начале координат, а звезда и наблюдатель находятся на расстояниях l и d от струны, соответственно, как показано на рис. 36.22.

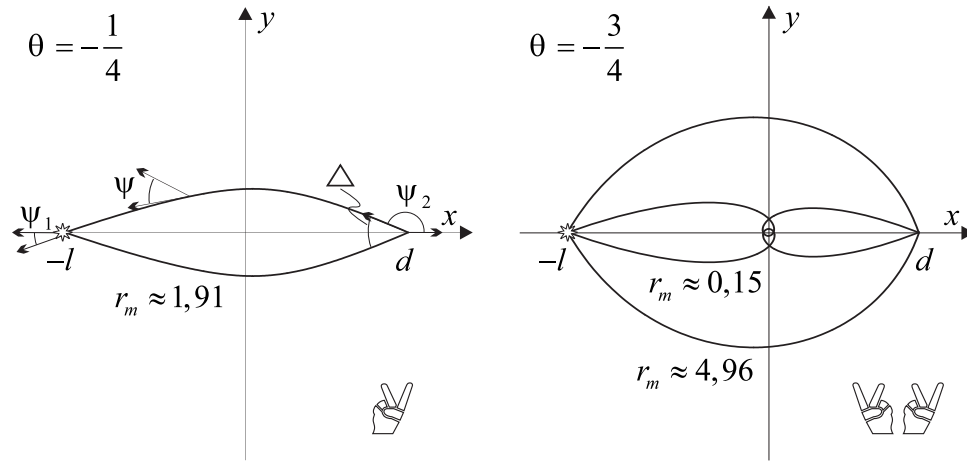


Рис. 36.22: Двойной образ звезды, находящейся за космической струной с небольшим, по модулю, углом дефицита (слева). Если угол дефицита близок к -2π , тогда звезда может иметь произвольное четное число образов. Справа показана ситуация, когда звезда имеет четыре образа, два из которых созданы лучами света, делающими один полный оборот вокруг космической струны.

Обозначим полярные координаты на плоскости x, y через r, φ . Угол ψ между касательным вектором к лучу света и радиальным направлением определяется следующим равенством

$$\operatorname{tg} \psi = r \frac{d\varphi}{dr}, \quad (36.159)$$

что ясно из рисунка. Правую часть равенства можно вычислить, используя уравнение (36.155) для формы экстремали, которое перепишем в виде

$$\Phi := \left(\frac{r}{r_m} \right)^{2\gamma} \sin^2 [\alpha(\varphi + \varphi_0)] - 1 = 0. \quad (36.160)$$

Используя правило дифференцирования сложной функции, получим выражение для производной

$$\frac{dr}{d\varphi} = - \frac{\partial \Phi / \partial \varphi}{\partial \Phi / \partial r} = - \frac{\alpha}{\gamma} r \frac{\cos [\alpha(\varphi + \varphi_0)]}{\sin [\alpha(\varphi + \varphi_0)]}$$

Теперь равенство (36.159) примет вид

$$\operatorname{tg} \psi = - \frac{\gamma}{\alpha} \operatorname{tg} [\alpha(\varphi + \varphi_0)]. \quad (36.161)$$

В дальнейшем полученное соотношение будет использовано для определения углового разделения образов звезды для наблюдателя.

Уравнение экстремали (36.160) зависит от двух параметров: r_m и φ_0 . Эти параметры можно выразить через расстояния l от звезды до космической струны и d от космической струны до наблюдателя. Для заданной экстремали положения звезды и наблюдателя определяются следующими уравнениями:

$$\begin{aligned} \left(\frac{l}{r_m} \right)^{2\gamma} \sin^2 \phi_1 &= 1, \\ \left(\frac{d}{r_m} \right)^{2\gamma} \sin^2 \phi_2 &= 1, \end{aligned} \quad (36.162)$$

где введены обозначения

$$\phi_1 := \alpha(\pi + 2\pi m + \phi_0), \quad \phi_2 := \alpha\phi_0. \quad (36.163)$$

Выше целое число $m = 0, 1, 2, \dots$ обозначает число оборотов, которое совершает луч света вокруг космической струны по пути от звезды до наблюдателя. Разность углов ϕ_1 и ϕ_2 должна быть меньше π , т.к. в противном случае звезда и наблюдатель пересекались бы разными ветвями экстремали. Это следует из формулы (36.157). Поэтому неравенство

$$\alpha\pi(1 + 2m) < \pi \quad (36.164)$$

ограничивает максимальное число обходов луча света вокруг космической струны. Формулы (36.162) неявно выражают постоянные r_m и φ_0 через l и d .

Угловое разделение образов равно $\Delta := 2(\pi - \psi_2)$, что ясно из рис. 36.22 слева. Поэтому из формулы (36.161) вытекает равенство

$$\operatorname{tg} \left(\pi - \frac{\Delta}{2} \right) = -\frac{\gamma}{\alpha} \operatorname{tg} \phi_2 \quad \Leftrightarrow \quad \operatorname{tg} \phi_2 = \frac{\alpha}{\gamma} \operatorname{tg} \frac{\Delta}{2}. \quad (36.165)$$

Из равенств (36.162) следует уравнение

$$d^\gamma |\sin \phi_2| = l^\gamma |\sin [\phi_2 + \alpha\pi(1 + 2m)]|, \quad (36.166)$$

где угол ϕ_2 связан с угловым разделением образов Δ равенством (36.165). Таким образом, мы получили общую неявную формулу, которая позволяет определить угловое разделение образов звезды по параметрам космической струны γ и α , а также расстояниям l и d .

Для малых отрицательных углов дефицита, $-\theta \ll 1$, справедливы равенства:

$$\alpha := 1 + \theta, \quad \gamma \approx 1 + \theta \frac{1 - 2\sigma}{2(1 - \sigma)},$$

где мы воспользовались формулой (36.101). Из неравенства (36.164) следует $m = 0$, т.е. наблюдатель видит только два образа звезды: один сверху, а другой снизу космической струны. В дальнейшем мы увидим, что при этом угловое разделение мало и имеет порядок $\Delta \sim \theta$. Поэтому из равенства (36.165) следует, что угол $\phi_2 \sim \theta$, а для постоянных α и γ можно использовать нулевой приближение $\alpha \approx \gamma \approx 1$. Таким образом, для малых углов дефицита уравнение (36.166) принимает вид

$$d \sin \frac{\Delta}{2} = -l \sin \left(\frac{\Delta}{2} + \pi\theta \right).$$

Поскольку Δ и θ имеют один порядок малости, то отсюда вытекает выражение для углового разделения образов звезды при малых углах дефицита космической струны

$$\Delta = \frac{l}{d+l} 2\pi|\theta|. \quad (36.167)$$

Это – хорошо известная формула [314]. Совпадение ответов является нетривиальным результатом, т.к. угловые разделения получены в разных системах координат. В разделе 36.8.3 было показано, что эти системы координат связаны конформным преобразованием, которое сохраняет углы.

Отметим, что в линейном приближении угловое разделение образов (36.167) не зависит от упругих свойств пространства.

В общем случае, если угол дефицита достаточно мал, то наблюдатель видит только два образа звезды. Допустим, что угол дефицита близок к -2π . Тогда неравенство (36.164) будет выполнено для всех $m = 0, 1, \dots, M$, где M – некоторое максимальное значение. Поэтому наблюдатель будет видеть $2(1 + M)$ образов звезды: два образа, созданные лучами света, которые не делают ни одного оборота, два образа для лучей света, делающими один оборот и т.д. На рис. 36.22 справа показана звезда с четырьмя образами для $\theta = -3/4$. Для такого угла дефицита каждая экстремаль делает один полный оборот вокруг космической струны перед тем, как уйти в бесконечность. Однако две экстремали достигают наблюдателя до того, как сделают оборот. Они характеризуются большим значением параметра r_m .

36.9.3 Общий подход к анализу параллельных клиновых дислокаций

Прямое интегрирование уравнений для экстремалей для метрики (36.102) представляет значительные трудности, потому что в плоскости x, y нет вектора Киллинга. Тем не менее описать рассеяние фононов можно для асимптотических форм метрики.

В общем случае интеграл (36.104) задает решение уравнения Гамильтона–Якоби для экстремалей (36.144)

$$\frac{\partial S}{\partial t} + \frac{1}{2} g^{\mu\nu} \frac{\partial S}{\partial x^\mu} \frac{\partial S}{\partial x^\nu} = 0, \quad (36.168)$$

где

$$S := \sqrt{2E} \sqrt{(w - c)(\bar{w} - \bar{c})} - Et, \quad c := a + ib = \text{const},$$

и E – энергия частицы (фонона). Для евклидовой метрики (36.103) экстремальными являются прямые линии,

$$(u - a) \sin \gamma = (v - b) \cos \gamma, \quad (36.169)$$

или

$$(w - c) e^{-i\gamma} = (\bar{w} - \bar{c}) e^{i\gamma}, \quad (36.170)$$

проходящие через точку (a, b) под углом γ . Для того, чтобы найти траекторию фонона в исходной декартовой системе координат, уравнение (36.169) надо просто переписать в системе координат x, y . В общем случае это нельзя сделать в явном виде, т.к. интеграл (36.104) не берется. Однако можно проанализировать асимптотику траекторий на больших расстояниях (задача рассеяния) $|z| \gg |z_i|$. Ниже мы проанализируем рассеяние фононов для дипольного и квадрупольного приближения.

Дипольное приближение

Рассмотрим произвольное распределение клиновых дислокаций. Если полный угол дефицита равен нулю, $\Theta = 0$, а полный вектор Бюргерса отличен от нуля, $B \neq 0$, тогда поведение экстремалей на больших расстояниях определяется дипольным слагаемым в разложении (36.109). Двумерный интервал в плоскости z , соответствующий конформному отображению (36.117), в полярных координатах r, φ имеет вид

$$dl^2 = \left(1 - B \frac{2 \cos \varphi}{r}\right) (dr^2 + r^2 d\varphi^2) \quad (36.171)$$

и является асимптотически плоским. Экстремали (36.169) определяются равенством

$$r \sin(\varphi - \gamma) + B(\ln r \sin \gamma + \varphi \cos \gamma) - a \sin \gamma + b \cos \gamma = 0, \quad (36.172)$$

где координаты $w = (u, v)$ выражены через $z = (r, \varphi)$ с помощью равенства (36.117). В дипольном приближении рассеяние определяется вектором Бюргерса, и, следовательно, траектории фононов зависят от угла, под которым они уходят в бесконечность. Мы рассмотрим два случая. При $\gamma = 0$ траектории на больших расстояниях перпендикулярны вектору Бюргерса. В этом случае уравнение (36.172) сводится к равенству

$$\frac{y}{x} = -\operatorname{tg} \frac{y+b}{B},$$

и экстремали параметризуются одной постоянной b . На рис. 36.23 экстремали показаны для $b = 0$ и $b > 0$. Экстремали для отрицательных $b < 0$ получаются из тех, что изображены на рисунке справа, отражением $y \mapsto -y$. Экстремали имеют горизонтальные асимптоты на бесконечности, которые сдвинуты на вектор Бюргерса πB , $B > 0$. Таким образом, рассеяние сводится к сдвигу траекторий фононов на вектор Бюргерса и задержки по времени. Отметим наличие возвращающихся траекторий для фононов, падающих справа. Если фононы падают слева, то рассеяние назад отсутствует. Это значит, что дислокация невидима для фононов, падающих слева и перпендикулярно вектору Бюргерса.

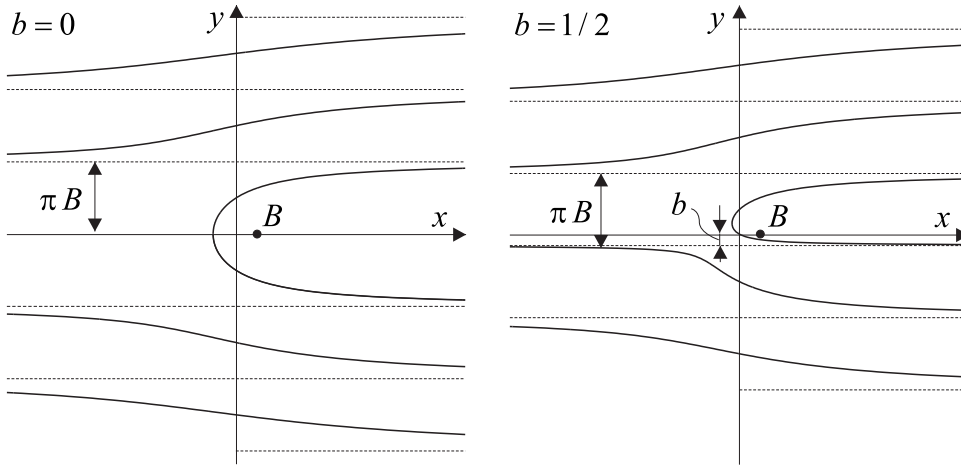


Рис. 36.23: Экстремали для краевой дислокации с $b = 0$ и $b = 1/2$, которые перпендикулярны вектору Бюргерса на бесконечности. Для всех $b \neq 0$ возвращающиеся экстремали проходят через линию дислокации.

Экстремали, параллельные вектору Бюргерса в бесконечности, имеют другое поведение и показаны на рис. 36.24. Они соответствуют значению $\gamma = \pi/2$ и определены уравнением

$$x^2 + y^2 = \exp \left(2 \frac{x+a}{B} \right). \quad (36.173)$$

В этом случае экстремали параметризуются одной произвольной постоянной a . На бесконечности экстремали не имеют асимптот, и поэтому удаленный наблюдатель может почувствовать наличие дислокации, т.к. экстремали не являются прямыми линиями. Рассеяние фононов также сопровождается задержкой по времени. Для $a = a_0$, где

$$a_0 := B \ln B,$$

экстремаль имеет точку самопересечения, расположенную в точке $(B, 0)$ комплексной плоскости z . Если $a > a_0$, то экстремали имеют по одной ветви. При $a < a_0$ решение уравнения (36.173) содержат две ветви. Одна ветвь начинается и заканчивается в бесконечности, а другая является замкнутой и окружает линию дислокации. На самом деле эта ветвь может не быть замкнутой, если скорость фонона вдоль оси x^3 отлична от нуля. В таком случае она будет представлять собой спираль.

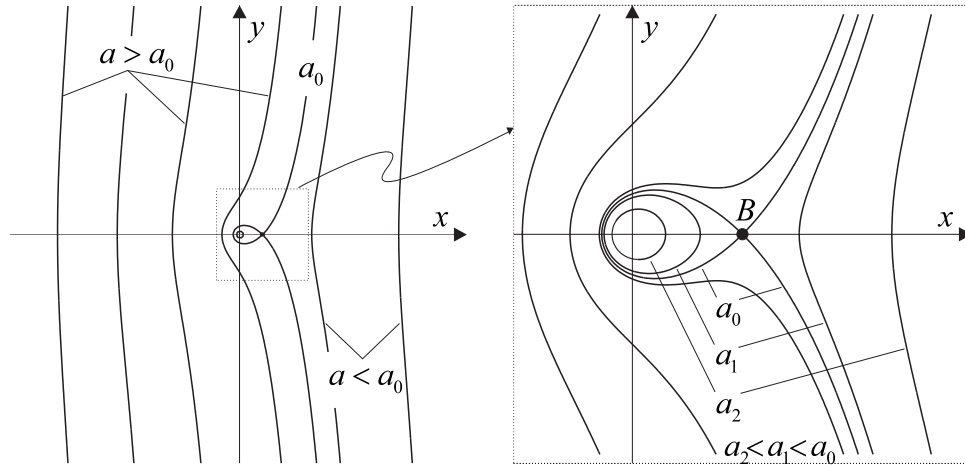


Рис. 36.24: Экстремали для краевой дислокации, которые параллельны вектору Бюргерса на бесконечности. Справа показана увеличенная центральная часть. Для $a < a_0$ экстремали имеют по две ветви.

Напомним, что экстремали построены для дипольного приближения, и поэтому можно ожидать качественного согласия только вдали от ядра дислокации.

Квадрупольное приближение

В квадрупольном приближении $\Theta = 0$, $B = 0$, $M \neq 0$, и конформное отображение (36.109) на больших расстояниях принимает вид

$$w = z - \frac{M}{z}. \quad (36.174)$$

Без потери общности положим $\text{im } M = 0$ и предположим, что $M > 0$. Для отрицательных значений M всю картину рассеяния нужно повернуть на угол π . Двумерная часть метрики для конформного отображения (36.174) в полярных координатах r, φ равна

$$dl^2 = \left(1 + 2M \frac{\cos(2\varphi)}{r^2}\right) (dr^2 + r^2 d\varphi^2). \quad (36.175)$$

Она стремится к евклидовой метрике на бесконечности $r \rightarrow \infty$ быстрее, чем в дипольном приближении. В квадрупольном приближении уравнение для экстремалей (36.169) сводится к равенству

$$r \sin(\varphi - \gamma) + \frac{M}{r} \sin(\varphi + \gamma) - a \sin \gamma + b \cos \gamma = 0. \quad (36.176)$$

В бесконечности это уравнение определяет прямые линии, и поэтому все экстремали имеют асимптоты.

Рассмотрим два случая. Экстремали, параллельные оси x на бесконечности, соответствуют значению $\gamma = 0$ и определяются уравнением

$$x^2 + y^2 = -\frac{My}{y+b}, \quad (36.177)$$

где $b \neq 0$ – произвольный параметр. Для отрицательных значений $b < 0$ все экстремали расположены в верхней полуплоскости и показаны на рис. 36.25 слева. Экстремали для положительных $b > 0$ получаются из экстремалей для $b < 0$ путем отражения $y \mapsto -y$. Для $-2\sqrt{M} < b < 0$ экстремаль имеет только одну ветвь, идущую от $x = -\infty$ до $x = +\infty$ и касающуюся оси x в начале координат. Если $b = -2\sqrt{M}$, то экстремаль имеет точку самопересечения с координатами $x = 0$, $y = \sqrt{M}$. При $b < -2\sqrt{M}$ экстремаль имеет две ветви. Одна бесконечная ветвь идет от $x = -\infty$ до $x = +\infty$ и не касается оси абсцисс. Вторая ветвь является замкнутой и проходит через начало координат. Все бесконечные экстремали имеют одинаковые асимптоты $y = -b$ слева и справа при $x \rightarrow \pm\infty$, и все рассеяние сводится только к задержке во времени.

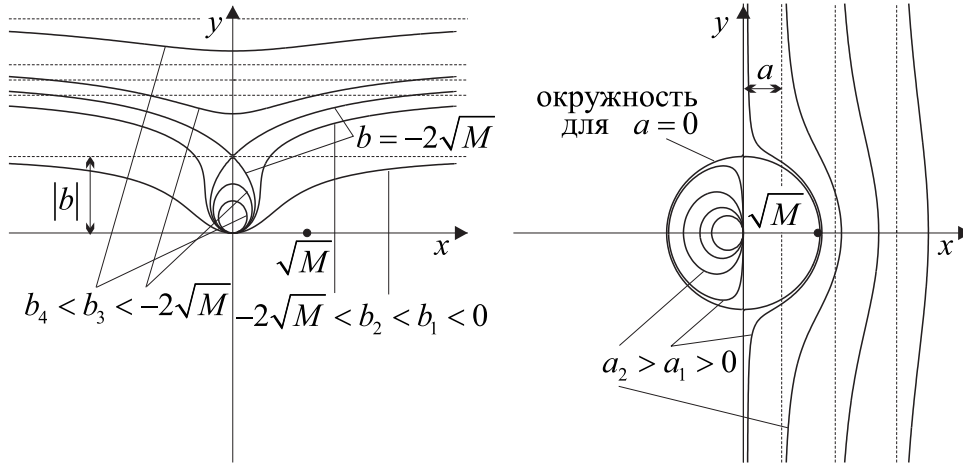


Рис. 36.25: Экстремали для квадрупольного приближения. Слева показаны экстремали, параллельные оси абсцисс x , а справа – параллельные оси ординат y в бесконечности. Каждая экстремаль при $b < -2\sqrt{M}$ имеет две ветви.

Экстремали, параллельные оси ординат y в бесконечности, соответствуют значению $\gamma = \pi/2$ и определяются уравнением.

$$x^2 + y^2 = \frac{Mx}{x-a}.$$

Оно отличается от уравнения (36.177) знаком правой части и заменой $x \leftrightarrow y$. Здесь постоянная M также предполагается положительной. Для $a = 0$ экстремаль представляет собой окружность радиуса $\sqrt{M}$. Для положительных значений $a > 0$ экстремали показаны на рис. 36.25 справа. Экстремали для отрицательных $a < 0$ получаются отражением $x \mapsto -x$. Каждая экстремаль имеет две ветви. Одна ветвь является бесконечной и идет от $y = -\infty$ до $y = +\infty$ с одинаковой асимптотикой $x = a$. Вторая ветвь замкнута, лежит внутри окружности на другой полуплоскости и проходит через начало координат. В этом случае рассеяние также сводится только к задержке во времени.

36.10 Цилиндрическая дислокация

В настоящем разделе мы рассмотрим новый тип дислокаций – цилиндрические, у которых δ -образная особенность сосредоточена на поверхности цилиндра. Этот вид дислокаций интересен тем, что индуцированная метрика, полученная в рамках теории упругости, в точности совпадает с метрикой в геометрической теории дефектов, где она является точным решением уравнений Эйнштейна. Кроме того, будет показано, что цилиндрическая дислокация может иметь физические приложения, т.к. представляет собой континуальную модель двустенных нанотрубок.

В трехмерной гравитации цилиндрические дислокации соответствуют статическим замкнутым струнам [315].

36.10.1 Цилиндрическая дислокация в теории упругости

Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ (бесконечную однородную и изотропную упругую среду или эфир в общей теории относительности) с декартовой системой координат $x^i, y^i, i = 1, 2, 3$. Евклидова метрика, как всегда, обозначена через $\delta_{ij} := \text{diag}(+, +, +)$. Основной переменной в теории упругости является векторное поле смещений $u^i(x)$, $x \in \mathbb{R}^3$, которое характеризует смещение точек среды (36.1). В отсутствии внешних сил законы Ньютона (36.2) и Гука (36.3) сводятся к системе уравнений второго порядка (36.5) с некоторыми граничными условиями, которые определяются рассматриваемой задачей.

Поставим задачу для цилиндрической дислокации, показанной на рис. 36.26*a*. Эта

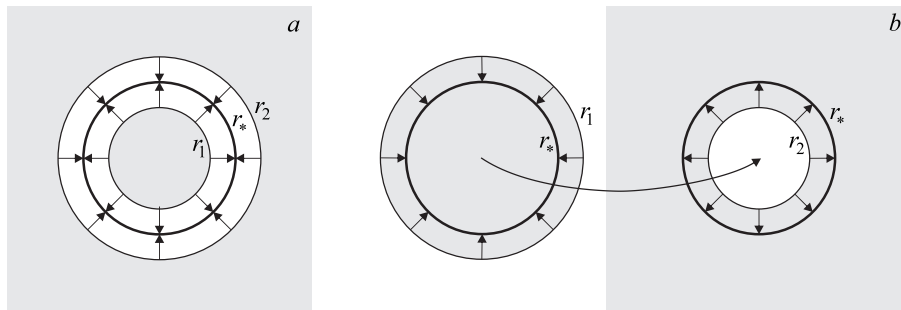


Рис. 36.26: Отрицательная (а) и положительная (b) цилиндрические дислокации.

дислокация возникает в результате следующей процедуры. Мы вырезаем толстую трубку вещества между двумя параллельными цилиндрами радиусов r_1 и r_2 ($r_1 < r_2$), осью которых является ось $z := x^3$, сдвигаем цилиндры симметрично друг к другу и затем склеиваем их. После установления равновесия поверхность склейки также представляет собой цилиндр радиуса r_* , который необходимо найти, благодаря вращательной и трансляционной симметрии.

По нашему соглашению рис. 36.26*a* изображает отрицательную цилиндрическую дислокацию, т.к. часть вещества удалена, $r_1 < r_2$. Эту процедуру можно инвертировать, добавив часть вещества к $\mathbb{R}^3$ как показано на рис. 36.26*b*. В этом случае цилиндрическая дислокация будет положительной, $r_1 > r_2$.

Описание цилиндрической дислокации в рамках теории упругости естественно провести в цилиндрической системе координат r, φ, z . Обозначим компоненты поля смещений относительно ортонормального базиса через $u^{\hat{r}}, u^{\hat{\varphi}}$ и $u^{\hat{z}}$. В рассматриваемом случае две компоненты равны нулю, $u^{\hat{\varphi}} = 0, u^{\hat{z}} = 0$, благодаря симметрии

задачи, а радиальная компонента зависит только от радиуса $u^{\hat{r}}(r)$. Для простоты индекс этой компоненты писать не будем, $u^{\hat{r}}(r) := u(r)$.

Граничные условия для цилиндрической дислокации имеют вид

$$u|_{r=0} = 0, \quad u|_{r=\infty} = 0, \quad \left. \frac{du_{\text{in}}}{dr} \right|_{r=r_*} = \left. \frac{du_{\text{ex}}}{dr} \right|_{r=r_*}. \quad (36.178)$$

Первые два условия являются геометрическими, а третье означает равенство упругих сил на поверхности склейки. Буквы “in” и “ex” обозначают векторное поле смещений, соответственно, на внутренней (internal) и внешней (external) частях среды.

Кроме этого граничные условия необходимо дополнить условием сплошности среды

$$r_1 + u_{\text{in}}(r_*) = r_2 + u_{\text{ex}}(r_*) \quad \Leftrightarrow \quad l := r_2 - r_1 = u_{\text{in}}(r_*) - u_{\text{ex}}(r_*), \quad (36.179)$$

где l – расстояние между исходными цилиндрами (толщина удаленного вещества).

В рассматриваемом случае уравнения теории упругости (36.5) легко решаются. Лапласиан от радиальной компоненты (только он отличен от нуля) и дивергенция в цилиндрических координатах имеют вид (см. раздел 7.2)

$$\begin{aligned} \Delta u_{\hat{r}} &= \frac{1}{r} \partial_r (r \partial_r u_{\hat{r}}) + \frac{1}{r^2} \partial_\varphi^2 u_{\hat{r}} + \partial_z^2 u_{\hat{r}} - \frac{1}{r^2} u_{\hat{r}} - \frac{2}{r^2} \partial_\varphi u_{\hat{\varphi}}, \\ \partial_i u^i &= \frac{1}{r} \partial_r (r u^{\hat{r}}) + \frac{1}{r} \partial_\varphi u^{\hat{\varphi}} + \partial_z u^{\hat{z}}, \end{aligned}$$

где индексы опущены с помощью евклидовой метрики: $u_{\hat{r}} = u^{\hat{r}}$, $u_{\hat{\varphi}} = u^{\hat{\varphi}}$ и $u_{\hat{z}} = u^{\hat{z}}$. Угловая φ и z компоненты уравнений (36.5) тождественно удовлетворяются, а радиальная компонента сводится к уравнению

$$\partial_r \left[\frac{1}{r} \partial_r (r u) \right] = 0, \quad (36.180)$$

которое имеет общее решение

$$u = ar - \frac{b}{r}, \quad a, b = \text{const},$$

зависящее от двух постоянных интегрирования a и b . Из первых двух граничных условий (36.178) вытекает, что решения внутри и снаружи поверхности склейки равны

$$\begin{aligned} u_{\text{in}} &= ar, \quad a > 0, \\ u_{\text{ex}} &= -\frac{b}{r}, \quad b > 0. \end{aligned} \quad (36.181)$$

Знаки постоянных интегрирования соответствуют отрицательной цилиндрической дислокации, изображенной на рис. 36.26а. Для положительной цилиндрической дислокации, рис. 36.26б, постоянные имеют противоположные знаки: $a < 0$ и $b < 0$. Третье граничное условие (36.178) определяет радиус цилиндра, вдоль которого проходит склейка,

$$r_*^2 = \frac{b}{a}. \quad (36.182)$$

После простых алгебраических вычислений с учетом условия сплошности (36.179) постоянные интегрирования можно выразить через радиусы:

$$a = \frac{r_2 - r_1}{r_2 + r_1} = \frac{l}{2r_*}, \quad b = \frac{r_2^2 - r_1^2}{4} = \frac{lr_*}{2}, \quad (36.183)$$

где

$$l = r_2 - r_1, \quad r_* = \frac{r_2 + r_1}{2}$$

– толщина удаленной трубы вещества и радиус поверхности склеивания. Первое выражение (36.183) накладывает ограничение на область значений постоянной интегрирования $0 < |a| < 1$, т.к. $r_1 > 0$ и $r_2 > 0$. Для отрицательной и положительной цилиндрической дислокации, соответственно, $l > 0$ и $l < 0$. Мы видим, что поверхность склейки находится в точности посередине радиусов r_1 и r_2 . Итак, уравнение (36.181) с постоянными интегрирования (36.183) дает полное решение для цилиндрической дислокации в рамках линейной теории упругости. Мы ожидаем, что оно справедливо при малых относительных деформациях: $l/r_1 \ll 1$, $l/r_2 \ll 1$.

Отметим, что результат, полученный в рамках теории упругости, не зависит от коэффициента Пуассона. В этом смысле цилиндрическая дислокация является чисто геометрическим дефектом, который не зависит от упругих свойств среды.

Теперь вычислим геометрические характеристики риманова многообразия, соответствующего цилиндрическому дефекту. С геометрической точки зрения упругие деформации (36.1) представляют собой диффеоморфизм областей евклидова пространства. Исходная среда описывается декартовыми координатами y^i и евклидовой метрикой δ_{ij} . Возврат отображения $x \mapsto y$ индуцирует нетривиальную метрику на евклидовом пространстве $\mathbb{R}^3$, соответствующем цилиндрической дислокации. В декартовых координатах она имеет вид

$$g_{ij}(x) = \frac{\partial y^k}{\partial x^i} \frac{\partial y^l}{\partial x^j} \delta_{kl}.$$

Мы используем криволинейные цилиндрические координаты и поэтому изменим обозначения. Обозначим индексы относительно криволинейной системы координат греческими буквами x^μ , $\mu = 1, 2, 3$. Тогда индуцированная метрика для цилиндрической дислокации равна

$$g_{\mu\nu}(x) = \frac{\partial y^\rho}{\partial x^\mu} \frac{\partial y^\sigma}{\partial x^\nu} \overset{\circ}{g}_{\rho\sigma} \quad (36.184)$$

где $\overset{\circ}{g}_{\rho\sigma}$ – евклидова метрика, записанная в цилиндрических координатах. Обозначим цилиндрические координаты точки среды до того как была создана дислокация через $\{y, \varphi, z\}$, где буква y без индекса обозначает радиальную координату. Поскольку при создании дислокации координаты φ и z фиксированной точки среды не меняются, то диффеоморфизм описывается одной функцией от радиуса: $y = r - u(r)$, где

$$u(r) := \begin{cases} ar, & r < r_*, \\ \frac{b}{r}, & r > r_*. \end{cases} \quad (36.185)$$

Эта функция имеет скачок $ar_* + b/r_* = l$ при переходе через разрез и поэтому не является непрерывной. В связи с этим при вычислении геометрических объектов необходимо проявлять осторожность. С этой целью введем функцию

$$v(r) := \begin{cases} a, & r \leq r_*, \\ \frac{b}{r^2}, & r \geq r_*, \end{cases} \quad (36.186)$$

которая непрерывна на поверхности склейки. Графики функций u и v приведены на рис. 36.27

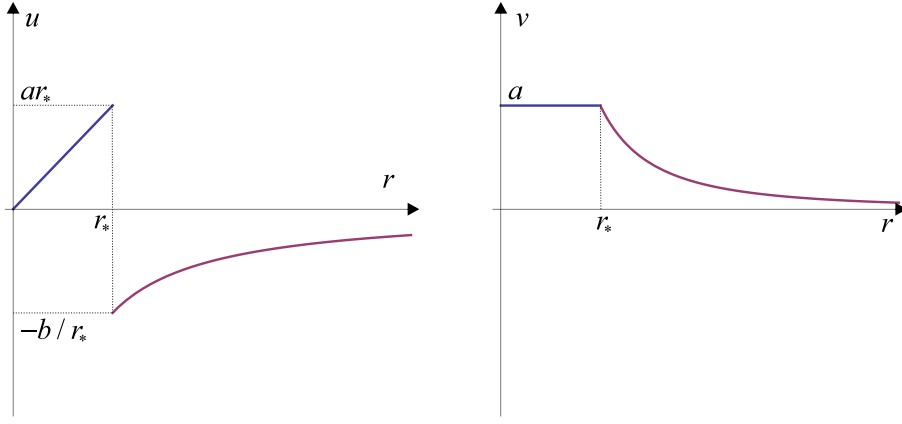


Рис. 36.27: Графики функций u и v . Функция u имеет скачок l на поверхности склейки при r_* , а функция v непрерывна, но имеет скачок производной.

Функция v отличается от производной по радиусу r от векторного поля смещений $u(r)$, определенного равенством (36.185), на δ -функцию:

$$u' = v - l\delta(r - r_*). \quad (36.187)$$

Индукцированная метрика вне поверхности склейки задана выражением (36.184) и неопределена на поверхности склейки. Поэтому мы *определяем* метрику цилиндрической дислокации следующим образом

$$ds^2 := (1 - v)^2 dr^2 + (r - u)^2 d\varphi^2 + dz^2. \quad (36.188)$$

Ей соответствует элемент объема

$$\sqrt{g} = (1 - v)(r - u), \quad \text{где} \quad g := \det g_{\mu\nu}.$$

Метрика (36.188) отличается от формальной подстановки $y = r - u(r)$ в евклидову метрику $ds^2 = dy^2 + y^2 d\varphi^2 + dz^2$ на квадрат δ -функции в g_{rr} компоненте. Эта процедура необходима в геометрической теории дефектов, т.к. в противном случае вектор Бюргерса нельзя выразить через поверхностный интеграл (см. раздел 36.2). Таким образом, радиальная компонента метрики $g_{rr}(r) = (1 - v)^2$ цилиндрической дислокации является непрерывной функцией, а угловая компонента $g_{\varphi\varphi} = (r - u)^2$ имеет скачок на поверхности склейки.

Итак, мы решили уравнения линейной теории упругости (36.5) с граничными условиями (36.178), (36.179), которые описывают цилиндрическую дислокацию. Вектор смещения (36.185) и, следовательно, индуцированная метрика (36.188) не зависят от упругих свойств среды (коэффициента Пуассона). Это демонстрирует универсальный характер цилиндрических дислокаций.

Теперь вычислим геометрические характеристики цилиндрической дислокации. Компоненты метрики являются недифференцируемыми функциями, и, следовательно, вычисления, включающие дифференцирование и умножение в общем случае не определены. Поэтому мы проведем вычисления таким образом как будто все компоненты метрики являются достаточно гладкими функциями, а затем убедимся, что все неопределенные слагаемые в конечном ответе благополучно сокращаются. Это означает, что какую бы регуляризацию мы не выбрали, конечный ответ от нее не зависит.

Сначала вычислим символы Кристоффеля. Только четыре их компоненты отличны от нуля:

$$\begin{aligned}\tilde{\Gamma}_{rrr} &= -v'(1-v), \\ \tilde{\Gamma}_{r\varphi\varphi} &= \tilde{\Gamma}_{\varphi r\varphi} = [1-v+l\delta(r-r_*)](r-u), \\ \tilde{\Gamma}_{\varphi\varphi r} &= -[1-v+l\delta(r-r_*)](r-u),\end{aligned}$$

где использовано соотношение (36.187). Отличные от нуля компоненты символов Кристоффеля с одним верхним индексом имеют вид

$$\begin{aligned}\tilde{\Gamma}_{rr}{}^r &= -\frac{v'}{1-v}, \\ \tilde{\Gamma}_{r\varphi}{}^\varphi &= \tilde{\Gamma}_{\varphi r}{}^\varphi = \frac{1-v+l\delta(r-r_*)}{r-u}, \\ \tilde{\Gamma}_{\varphi\varphi}{}^r &= -\frac{[1-v+l\delta(r-r_*)](r-u)}{(1-v)^2}.\end{aligned}$$

Заметим, что если бы мы не отбросили δ -функцию в g_{rr} компоненте метрики (36.188), то на нее пришлось бы делить, что является неопределенной операцией.

Из компонент тензора кривизны только одна отлична от нуля

$$\begin{aligned}\tilde{R}_{r\varphi r\varphi} &= \partial_r \tilde{\Gamma}_{\varphi r\varphi} + \tilde{\Gamma}_{rr}{}^r \tilde{\Gamma}_{\varphi\varphi r} - \tilde{\Gamma}_{\varphi r}{}^\varphi \tilde{\Gamma}_{r\varphi\varphi} = \\ &= l(r-u) \left[\delta'(r-r_*) + \frac{v'}{1-v} \delta(r-r_*) \right],\end{aligned}\quad (36.189)$$

где

$$\delta'(r-r_*) := \partial_r \delta(r-r_*)$$

– производная δ -функции. Неопределенные слагаемые с квадратами $\delta^2(r-r_*)$ в тензоре кривизны благополучно сократились. Нетривиальная компонента тензора кривизны (36.189) все еще остается неопределенной, потому что коэффициент при δ' и δ -функции имеет разрыв. Мы видим, что тензор кривизны (36.189) равен нулю всюду за исключением поверхности склейки, как и должно быть и с математической, и с физической точек зрения (36.189).

Еще раз отметим важное обстоятельство. Предписание отбросить δ -функцию из g_{rr} компоненты метрики для цилиндрической дислокации (36.188), которое было продиктовано физическими соображениями (равенство сил на обоих берегах разреза), определяет правильный путь проведения математических вычислений.

У тензора Риччи две нетривиальных компоненты:

$$\begin{aligned}\tilde{R}_{rr} &= \frac{l}{r-u} \left[\delta'(r-r_*) + \frac{v'}{1-v} \delta(r-r_*) \right], \\ \tilde{R}_{\varphi\varphi} &= \frac{l(r-u)}{(1-v)^2} \left[\delta'(r-r_*) + \frac{v'}{1-v} \delta(r-r_*) \right].\end{aligned}$$

Скалярная кривизна равна

$$\tilde{R} = \frac{2l}{(r-u)(1-v)^2} \left[\delta'(r-r_*) + \frac{v'}{1-v} \delta(r-r_*) \right].$$

Уравнения Эйнштейна,

$$\sqrt{g} \left(\tilde{R}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \tilde{R} \right) = -\frac{1}{2} T_{\mu\nu}, \quad (36.190)$$

тождественно удовлетворяются за исключением zz компоненты, которая принимает следующий вид

$$2l \left(\frac{1}{1-v} \delta'(r-r_*) + \frac{v'}{(1-v)^2} \delta(r-r_*) \right) = T_{zz}, \quad (36.191)$$

где T_{zz} – источник цилиндрической дислокации (аналог тензора энергии-импульса полей материи в общей теории относительности). Используя тождество

$$f(r)\delta'(r-r_*) = f(r_*)\delta'(r-r_*) - \partial_r f \delta(r-r_*),$$

справедливое для произвольной дифференцируемой функции $f \in C^1(\mathbb{R}_+)$ на вещественной положительной полупрямой, сингулярную часть можно переписать в виде

$$T_{zz} = \frac{2l}{1-v(r_*)} \delta'(r-r_*) = \frac{4lr_*}{2r_*-l} \delta'(r-r_*). \quad (36.192)$$

Мы видим, что все неопределенные слагаемые в уравнениях Эйнштейна сократились. В общем случае, если компоненты метрики не являются непрерывными функциями, компоненты тензора кривизны содержат квадраты δ -функций и произведения δ -функций на скачки и, следовательно, математически не определены. Поэтому сокращение всех неопределенных слагаемых в уравнениях Эйнштейна является исключительным случаем. Несмотря на то, что символы Кристоффеля и компоненты тензора кривизны являются неопределенными, тензор энергии-импульса корректно определен. Такая же ситуация имеет место для распределения клиновых дислокаций с множителем $\sqrt{g}$ для определения тензора энергии-импульса (36.190). Множитель $\sqrt{g}$ возникает из-за того, что δ -функция является не функцией, а скалярной плотностью.

В уравнениях Эйнштейна (36.190) мы переопределили тензор энергии-импульса (источник) по сравнению с тем, как он был определен в общей теории относительности (23.2). В настоящем разделе это не тензор, а тензорная плотность. Это существенно при работе с обобщенными функциями.

36.10.2 Цилиндрическая дислокация в геометрической теории дефектов

В предыдущем разделе мы описали цилиндрическую дислокацию и доказали, что индуцированная метрика является точным решением уравнений Эйнштейна. Однако остался открытым вопрос о том, что уравнения геометрической теории дефектов имеют еще какие-либо решения для цилиндрической дислокации при том же источнике. Поэтому в настоящем разделе мы опишем цилиндрическую дислокацию без привлечения теории упругости. С этой целью мы решим уравнения равновесия, следующие из вариационного принципа для действия (36.36) в системе координат, определенной упругой (36.43) и лоренцевой (36.44) калибровочными условиями.

Пусть цилиндрическая дислокация описывается источником $T_{\mu\nu}$, явный вид которого будет определен позже.

Цилиндрическая дислокация соответствует упругой среде без спиновой структуры и поэтому без дисклинаций. В этом случае тензор кривизны равен нулю $R_{\mu\nu}{}^{ij} = 0$ (но не тензор $\tilde{R}_{\mu\nu}{}^{ij}$!), уравнения равновесия, полученные вариацией действия (36.36) по связности $\omega_\mu{}^{ij}$ выполнены, и мы можем положить $R^\Lambda{}_{\mu\nu} = 0$ в действии (36.36).

После этого действие варьируется по реперу e_μ^i , и возникают уравнения Эйнштейна (36.190).

Для цилиндрической дислокации метрика имеет два вектора Киллинга, ∂_z и ∂_φ в цилиндрической системе координат, которые соответствуют инвариантности относительно трансляций вдоль оси z и вращениям в плоскости x, y . Для решения уравнений Эйнштейна выберем метрику в диагональном виде

$$g_{\mu\nu} = \begin{pmatrix} A^2 & 0 & 0 \\ 0 & B^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (36.193)$$

где $A(r)$ и $B(r)$ – две неизвестные положительные функции от радиуса. Хотя это не самый общий вид метрики, совместимой с симметрией задачи, его будет достаточно для наших целей. Соответствующий репер также можно выбрать в диагональном виде:

$$e_{\hat{r}} = A, \quad e_{\hat{\varphi}} = B, \quad e_{\hat{z}} = 1, \quad (36.194)$$

где координатные индексы обозначены $\{\mu\} = (r, \varphi, z)$, а индексы относительно ортонормального базиса помечены шляпкой: $\{i\} = (\hat{r}, \hat{\varphi}, \hat{z})$. Мера объема имеет вид

$$\sqrt{g} = AB.$$

Если бы упругая калибровка (36.43) не использовалась, то оставшуюся свободу в выборе радиальной координаты можно было бы использовать для дальнейшего упрощения вида метрики. Например, можно было бы положить $A = 1$.

Последующие вычисления будут проведены так, как будто функции A и B являются достаточно гладкими, а полученные в дальнейшем разрывы будут оправданы тем, что все неопределенные слагаемые в уравнениях для A и B сократятся.

Символы Кристоффеля для метрики (36.193) имеют четыре нетривиальные компоненты:

$$\tilde{\Gamma}_{rr}^r = \frac{A'}{A}, \quad \tilde{\Gamma}_{r\varphi}^\varphi = \tilde{\Gamma}_{\varphi r}^\varphi = \frac{B'}{B}, \quad \tilde{\Gamma}_{\varphi\varphi}^r = -\frac{BB'}{A^2},$$

где штрих обозначает производную по радиусу. Тензор кривизны имеет только одну отличную от нуля компоненту

$$\tilde{R}_{r\varphi r\varphi} = BB'' - \frac{A'BB'}{A}.$$

Нетривиальные компоненты тензора Риччи и скалярная кривизна имеют вид

$$\begin{aligned} \tilde{R}_{rr} &= \frac{B''}{B} - \frac{A'B'}{AB}, \\ \tilde{R}_{\varphi\varphi} &= \frac{BB''}{A^2} - \frac{A'BB'}{A^3}, \\ \tilde{R} &= \frac{2}{AB} \left(\frac{B''}{A} - \frac{A'B'}{A^2} \right). \end{aligned}$$

Мы полагаем, что источник имеет только одну отличную от нуля компоненту

$$T_{zz} := L\delta'(r - r_*), \quad (36.195)$$

где L и r_* – две постоянных, которые характеризуют силу и положение цилиндрической дислокации. Этот вид источника продиктован результатами, полученными в

рамках теории упругости (36.192). В общем случае вид источника может быть произволен, и ему соответствует свой репер и метрика. Это будет решением другой задачи, возможно, с непрерывным распределением параллельных дислокаций с вращательной симметрией.

Нетрудно проверить, что rr , $\varphi\varphi$, а также недиагональные компоненты уравнений Эйнштейна (36.190) для метрики (36.193) тождественно удовлетворяются. Компонента zz сводится к обыкновенному дифференциальному уравнению

$$\left(\frac{B'}{A}\right)' = \frac{1}{2}L\delta'(r - r_*). \quad (36.196)$$

Это – *линейное* неоднородное дифференциальное уравнение по отношению к комбинации компонент метрики B'/A . Поэтому рассмотрение сингулярного источника (36.195) оправдано. Данное уравнение определено для положительных непрерывных $A(r)$ и имеет общее решение

$$B' = \frac{1}{2}A(r_*)L\delta(r - r_*) + c_1A, \quad c_1 = \text{const}, \quad (36.197)$$

где c_1 – постоянная интегрирования. Полученное решение можно еще раз проинтегрировать

$$B = \frac{1}{2}A(r_*)L\theta(r - r_*) + c_1 \int_0^r ds A(s) + c_2, \quad c_2 = \text{const}, \quad (36.198)$$

где θ – функция ступеньки:

$$\theta(r - r_*) := \begin{cases} 0, & r \leq r_*, \\ 1, & r > r_*. \end{cases} \quad (36.199)$$

Таким образом, мы нашли решение уравнений Эйнштейна для цилиндрической дислокации. Оно зависит от одной произвольной непрерывной положительной функции $A(r)$ и двух постоянных интегрирования $c_{1,2}$. Произвольная функция A в решении отражает наличие остаточной свободы в выборе радиальной координаты и определяется упругой калибровкой. По существу, уравнения Эйнштейна с источником (36.195) определяют только скачок компоненты репера $e_\varphi^{\hat{\varphi}}$. Произвольная функция $e_r^{\hat{r}} = A(r)$ описывает свободу в выборе радиальной координаты.

Теперь мы наложим упругую калибровку (36.43). В цилиндрических координатах плоский репер можно выбрать диагональным с компонентами

$$\overset{\circ}{e}_r^{\hat{r}} = 1, \quad \overset{\circ}{e}_\varphi^{\hat{\varphi}} = r, \quad \overset{\circ}{e}_z^{\hat{z}} = 1.$$

Он определяет плоские символы Кристоффеля $\overset{\circ}{\Gamma}_{\mu\nu}^\rho$ и $\mathbb{SO}(3)$ связность $\overset{\circ}{\omega}_{\mu i}^j$, которые имеют следующие нетривиальные компоненты (см. раздел 7.2):

$$\begin{aligned} \overset{\circ}{\Gamma}_{r\varphi}^\varphi &= \overset{\circ}{\Gamma}_{\varphi r}^\varphi = \frac{1}{r}, & \overset{\circ}{\Gamma}_{\varphi\varphi}^r &= -r, \\ \overset{\circ}{\omega}_{\varphi\hat{r}}^{\hat{\varphi}} &= -\overset{\circ}{\omega}_{\varphi\hat{\varphi}}^{\hat{r}} = 1. \end{aligned}$$

Подстановка репера (36.194) в калибровочное условие (36.43) дает дифференциальное уравнение Эйлера для функции $A(r)$

$$A' + \frac{A}{r} - \frac{B}{r^2} + \frac{\sigma}{1 - 2\sigma} \left(A' + \frac{B'}{r} - \frac{B}{r^2} \right) = 0, \quad (36.200)$$

где функция $B(r)$ задана равенством (36.198).

Мы ищем классическое решение данного уравнения внутри поверхности склейки, A_{in} , и вне ее, A_{ex} , со следующими “асимптотически свободными” граничными условиями:

$$B_{\text{in}}|_{r=0} = 0, \quad 0 < A_{\text{in}}|_{r=0} < \infty, \quad B_{\text{ex}}|_{r \rightarrow \infty} = r. \quad (36.201)$$

Граничные условия в нуле $r = 0$ такие же, как и для евклидовой метрики. Кроме того, на поверхности склейки мы накладываем граничные условия:

$$A_{\text{in}}|_{r=r_*} = A_{\text{ex}}|_{r=r_*}, \quad B_{\text{in}}|_{r=r_*} + \frac{LA(r_*)}{2} = B_{\text{ex}}|_{r=r_*}. \quad (36.202)$$

Первое условие сшивки обеспечивает равенство нормальных упругих сил. Второе условие является следствием уравнения (36.198) и дает скачок $e_\varphi^{\hat{\varphi}}$ компоненты репера. Условия сшивки (36.201) и (36.202) аналогичны граничным условиям (36.178) в теории упругости.

Подстановка выражения для B из равенства (36.198) в первое условие сшивки (36.201) определяет одну из постоянных интегрирования: $c_2 = 0$.

Уравнение (36.200) проще решить относительно B вместо A . Внутри и снаружи поверхности склейки $B' = c_1 A$ как следствие уравнения (36.197). Тогда уравнение (36.200) сводится к условию

$$\frac{B''}{c_1} + \frac{B'}{c_1 r} - \frac{B}{r^2} + \frac{\sigma}{1-2\sigma} \left(\frac{B''}{c_1} + \frac{B'}{r} - \frac{B}{r^2} \right) = 0. \quad (36.203)$$

Это уравнение Эйлера совпадает с уравнением для клиновой дислокации (36.95), где роль угла дефицита конической особенности играет разность $c_1 - 1$. Оно имеет общее решение

$$B = D_1 r^{\gamma_1} + D_2 r^{\gamma_2}, \quad D_{1,2} = \text{const}, \quad (36.204)$$

зависящее от двух постоянных интегрирования $D_{1,2}$, и где постоянные $\gamma_{1,2}$ являются корнями квадратного уравнения

$$\gamma^2 + \frac{(c_1 - 1)\sigma}{1 - \sigma} \gamma - c_1 = 0.$$

При $c_1 > 0$ это уравнение имеет два вещественных корня противоположных знаков: положительный корень γ_1 и отрицательный γ_2 .

В области интегрирования мы имеем равенства $D_2 = 0$ и $\gamma_1 = 1$ как следствие первых двух условий (36.201), при этом равенство $\gamma_1 = 1$ эквивалентно условию $c_1 = 1$ для $\sigma \neq 1/2$. Следовательно, решение уравнений Эйнштейна во внутренней области имеет вид

$$B_{\text{in}} = D_1 r, \quad A_{\text{in}} = D_1. \quad (36.205)$$

Оно зависит от одной произвольной постоянной D_1 . Итак, два первых граничных условия (36.201) фиксируют две постоянных интегрирования уравнений Эйнштейна, $c_1 = 1$ и $c_2 = 0$, которые одинаковы и для внутренней, и для внешней области, и одну постоянную интегрирования в упругой калибровке, $D_2 = 0$.

Для того, чтобы уменьшить число индексов, обозначим постоянные интегрирования в равенстве (36.204) во внешней области новыми буквами:

$$B_{\text{ex}} = E_1 r^{\gamma_1} + E_2 r^{\gamma_2}, \quad E_{1,2} = \text{const}.$$

Постоянная интегрирования $c_1 = 1$ уже зафиксирована во внутренней области, и поэтому $\gamma_1 = 1$ и $\gamma_2 = -1$. Третье асимптотическое условие (36.201) Определяет $E_1 = 1$ и дает решение во внешней области:

$$B_{\text{ex}} = r + \frac{E_2}{r}, \quad A_{\text{ex}} = 1 - \frac{E_2}{r^2}, \quad (36.206)$$

которое также зависит от произвольной постоянной E_2 .

Для того, чтобы выразить постоянные D_1 и E_2 через параметры задачи L и r_* , характеризующие источник, используем условия сшивки решений. Первое условие (36.202) связывает постоянные D_1 и E_2 . Обозначив $D_1 := 1 - a$ и $E_2 := b$, получаем уравнение (36.182). Наконец, второе условие сшивки (36.202) приводит к равенству

$$L = 2 \frac{B_{\text{ex}}(r_*) - B_{\text{in}}(r_*)}{A(r_*)} = \frac{4lr_*}{2r_* - l},$$

где $l := 2\sqrt{ab}$, совпадающее с коэффициентом при δ -функции в источнике (36.195).

Таким образом, мы решили задачу для цилиндрической дислокации в рамках геометрической теории дефектов. Поле репера определено формулой (36.194), где функции A и B заданы равенствами (36.205) и (36.206). Метрика (36.193) в точности совпадает с индуцированной метрикой (36.188), полученной в рамках теории упругости. Важно отметить, что мы получили метрику для цилиндрической дислокации, решив уравнения Эйнштейна в упругой калибровке, ничего не зная о векторном поле смещений. Если нужно, то вектор смещения можно восстановить во внутренней и внешней области путем решения уравнения

$$\frac{\partial y^i}{\partial x^\mu} = e_\mu^i$$

с соответствующими граничными условиями, где функции y^i определены равенством (36.1). Это уравнение имеет решение, т.к. тензор кривизны равен нулю, $\bar{R}_{\mu\nu\rho\sigma} = 0$.

С геометрической точки зрения, нам просто нужно найти такую систему координат, в которой метрика становится евклидовой. Это легко сделать непосредственно. Существенная двумерная часть метрики во внутренней и внешней областях имеет вид

$$dl_{\text{in}}^2 = (1 - a)^2 dr^2 + (1 - a)^2 r^2 d\varphi^2, \\ dl_{\text{ex}}^2 = \left(1 - \frac{b}{r}\right)^2 dr^2 + \left(r + \frac{b}{r}\right)^2 d\varphi^2.$$

Вводя новые координаты $y := (1 - a)r$ во внутренней и $y := r + b/r$ во внешней областях, метрика приводится к евклидову виду $dl^2 = dy^2 + y^2 d\varphi^2$ в обеих областях. После этого сразу находится вектор смещения (36.185), используя определение (36.1).

Таким образом, мы решили задачу для цилиндрической дислокации в рамках теории упругости и геометрической теории дефектов. Результаты совпадают, хотя это совпадение не является автоматическим. Например, метрика для клиновой дислокации, полученная в рамках геометрического подхода существенно отличается от индуцированной метрики, полученной в рамках теории упругости (см. раздел 36.8.2), и воспроизводит последнюю только в линейном приближении. Это происходит потому что упругая калибровка (36.43) сводится к уравнениям нелинейной теории упругости для вектора смещений.

Возникает естественный вопрос: “Зачем использовать относительно сложный геометрический подход, если теория упругости также работает?” Ответ прост. Теория упругости работает для описания отдельных дефектов или их небольшого числа. Если число дефектов велико, то граничные условия настолько усложняются, что нет никакой надежды на решение данной задачи в рамках теории упругости. Например, мы не знаем решение задачи для произвольного распределения параллельных клиновых дислокаций в рамках теории упругости. В то время как эта же задача имеет простое решение в геометрическом подходе (36.86). Существует еще одно важное оправдание. Допустим, что имеется непрерывное распределение дефектов. Тогда эту задачу невозможно даже поставить в рамках теории упругости, т.к. векторного поля смещений просто не существует. При этом геометрический подход дает простую постановку: мы имеем хорошо определенные выражения для свободной энергии (36.36) и калибровочных условий (36.43), (36.44). Единственное отличие по сравнению с отдельными дефектами состоит в том, что надо рассматривать непрерывный источник дефектов $T_{\mu\nu}$ (тензор энергии-импульса).

Поверхностная плотность вектора Бюргерса задается компонентами тензора кручения (36.14). Простые вычисления показывают, что тензор кручения для цилиндрической дислокации имеет только одну отличную от нуля компоненту

$$T_{r\varphi}^{\hat{\varphi}} = -T_{\varphi r}^{\hat{\varphi}} = 1 - v + l\delta(r - r_*), \quad (36.207)$$

которая сингулярна на поверхности склейки, и нетривиальна во внутренней и внешней областях. Проекция вектора Бюргерса на ось x задается интегралом и равна нулю,

$$b^x = - \int_{\mathbb{R}^2} dx^\mu \wedge dx^\nu T_{\mu\nu}^{\hat{\varphi}} \sin \varphi = \int_0^\infty dr r \int_0^{2\pi} d\varphi T_{r\varphi}^{\hat{\varphi}} \sin \varphi = 0.$$

Аналогично, его проекция на произвольную прямую, пересекающую начало координат, также равна нулю. Следовательно, вектор Бюргерса цилиндрической дислокации равен нулю.

36.10.3 Цилиндрическая дислокация в общей теории относительности

Основным результатом раздела 36.10.1 является пространственная метрика (36.188), описывающая цилиндрическую дислокацию. Мы доказали, что она удовлетворяет трехмерным уравнениям Эйнштейна (36.190) с источником (36.195). Обобщение этой метрики на четырехмерный случай проводится непосредственно.

Предположим, что цилиндрическая дислокация не движется, т.е. метрика статична

$$ds^2 = dt^2 - (1 - v)^2 dr^2 - (r - u)^2 d\varphi^2 - dz^2, \quad (36.208)$$

где функции v и u определены равенствами (36.186) и (36.185). Компонента $g_{\varphi\varphi}$ является разрывной, а g_{rr} — непрерывна, но имеет скачок производной. Нетрудно проверить, что все компоненты символов Кристоффеля и кривизны для метрики (36.208), имеющие хотя бы один временной индекс тождественно равны нулю. Поэтому эта метрика удовлетворяет уравнениям Эйнштейна

$$\sqrt{|g|} \left(\tilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \tilde{R} \right) = -\frac{1}{2} T_{\alpha\beta}, \quad \alpha, \beta = 0, 1, 2, 3, \quad (36.209)$$

где греческие индексы из начала алфавита пробегают все значения индексов от 0 до 3. Тензор энергии-импульса материи имеет только две исчезающие компоненты:

$$T_{00} = -T_{zz} = L\delta'(r - r_*). \quad (36.210)$$

Замечательным обстоятельством является тот факт, что все неопределенные слагаемые в левой части уравнения (36.209) для метрики (36.208) сокращаются. Итак, метрика (36.208) удовлетворяет четырехмерным уравнениям Эйнштейна, и ее физический смысл ясен из предыдущего рассмотрения: она описывает цилиндрическую дислокацию.

36.10.4 Коническая цилиндрическая дислокация

В предыдущих разделах мы описали цилиндрическую дислокацию в рамках теории упругости и геометрической теории дефектов. Метрика в обоих случаях оказалась идентичной. Ниже мы опишем другой тип цилиндрических дислокаций, которые названы коническими цилиндрическими дислокациями, т.к. они имеют такую же асимптотику на бесконечности как и конические особенности. Эти дислокации описаны в рамках геометрической теории дефектов.

В настоящем разделе постоянные интегрирования обозначены так же, как и в разделе 36.10.2, но имеют другой смысл.

Рассмотрим трехмерные уравнения Эйнштейна (36.190) с источником, который имеет только одну отличную от нуля компоненту

$$T_{zz} := 2\Theta\delta(r - r_*), \quad \Theta = \text{const}, \quad (36.211)$$

в цилиндрических координатах. В отличие от предыдущего случая мы заменили производную δ -функции в формуле (36.195) на саму δ -функцию. Будем искать решение, которое обладает трансляционной и вращательной инвариантностью. В дальнейшем мы увидим, что постоянная Θ играет роль угла дефицита конической особенности, соответствующей асимптотике $r \rightarrow \infty$.

Эта задача эквивалентна статической конической цилиндрической дислокации в общей теории относительности с тензором энергии-импульса

$$T_{00} = -T_{zz} = 2\Theta\delta(r - r_*). \quad (36.212)$$

Физический смысл 2Θ – это поверхностная плотность энергии тонкого цилиндра радиуса r_* . Для обычной материи $\Theta > 0$. Однако мы рассмотрим оба случая, т.к. вычисления не сильно зависят от знака Θ .

Выберем метрику в виде (36.193). Тогда вся система уравнений Эйнштейна сведется к одному линейному обыкновенному дифференциальному уравнению (вычисления те же, что и при получении уравнения (36.196))

$$\left(\frac{B'}{A}\right)' = \Theta\delta(r - r_*).$$

Оно легко интегрируется:

$$B' = \Theta A\theta(r - r_*) + c_1 A, \quad (36.213)$$

$$B = \Theta \int_{r_*}^r ds A(s) + c_1 \int_0^r ds A(s) + c_2, \quad c_{1,2} = \text{const}, \quad (36.214)$$

где $\theta(r - r_*)$ – ступенька (36.199) и $c_{1,2}$ – постоянные интегрирования. Таким образом, уравнения Эйнштейна определяют компоненту репера $e_\varphi^{\hat{\varphi}} := B$ через компоненту $e_r^{\hat{r}} := A$, которая может быть произвольной положительной функцией. Если функция $A(r)$ непрерывна, то $B(r)$ также непрерывна благодаря уравнению (36.214), но ее производная имеет скачок (36.213).

Для того, что зафиксировать решение, наложим упругую калибровку (36.43), которая сводится к уравнению (36.200). Мы ищем решение данного уравнения внутри, A_{in} , и вне, A_{ex} , поверхности склейки. В теории упругости клиновья дислокация соответствует бесконечному цилиндру конечного радиуса, $0 < r < a$, т.к. напряжения расходятся в бесконечности (см. раздел 36.7.2). Пусть $a > r_*$. Тогда наложим следующие граничные условия:

$$B_{\text{in}}|_{r=0} = 0, \quad 0 < A_{\text{in}}|_{r=0} < \infty, \quad A_{\text{ex}}|_{r=a} = 1. \quad (36.215)$$

Последнее граничное условие означает отсутствие внешних нормальных сил на границе цилиндра. Мы также накладываем условия непрерывной сшивки

$$A_{\text{in}}|_{r=r_*} = A_{\text{ex}}|_{r=r_*}, \quad B_{\text{in}}|_{r=r_*} = B_{\text{ex}}|_{r=r_*}. \quad (36.216)$$

Во внутренней области из уравнения (36.213) вытекает равенство $B' = c_1 A$. Поэтому упругая калибровка сводится к уравнению (36.203) с теми же граничными условиями, что и для цилиндрической дислокации, рассмотренной ранее. Поэтому $c_1 = 1$, $c_2 = 0$, и решение во внутренней области задается той же формулой (36.205), что и раньше.

Во внешней области $B' = \alpha A$, где $\alpha := 1 + \Theta$, т.к. $c_1 = 1$. Общее решение уравнения (36.203) имеет вид

$$B_{\text{ex}} = E_1 r^{\gamma_1} + E_2 r^{\gamma_2},$$

где $\gamma_1 > 0$ и $\gamma_2 < 0$ являются корнями квадратного уравнения

$$\gamma^2 + \frac{\Theta\sigma}{1-\sigma}\gamma - \alpha = 0, \quad \alpha := 1 + \Theta. \quad (36.217)$$

Третье граничное условие (36.215),

$$\frac{E_1 \gamma_1}{\alpha} a^{\gamma_1-1} + \frac{E_2 \gamma_2}{\alpha} a^{\gamma_2-1} = 1, \quad (36.218)$$

приводит к соотношению между E_1 и E_2 при заданном a .

После этого постоянные D_1 и E_2 определяются через Θ и r_* условиями сшивки (36.216):

$$\frac{E_1 \gamma_1}{\alpha} r_*^{\gamma_1-1} + \frac{E_2 \gamma_2}{\alpha} r_*^{\gamma_2-1} = D_1, \quad (36.219)$$

$$E_1 r_*^{\gamma_1} + E_2 r_*^{\gamma_2} = D_1 r_*. \quad (36.220)$$

На практике, мы сначала решаем уравнения (36.219) и (36.220):

$$\begin{aligned} E_1 &= D_1 \frac{\alpha - \gamma_2}{\gamma_1 - \gamma_2} r_*^{-\gamma_1+1}, \\ E_2 &= -D_1 \frac{\alpha - \gamma_1}{\gamma_1 - \gamma_2} r_*^{-\gamma_2+1}. \end{aligned} \quad (36.221)$$

Затем постоянная D_1 находится путем подстановки этих решений в равенство (36.218),

$$D_1 = \frac{\alpha}{\gamma_1^{\frac{\alpha-\gamma_2}{\gamma_1-\gamma_2}} \left(\frac{a}{r_*}\right)^{\gamma_1-1} - \gamma_2^{\frac{\alpha-\gamma_1}{\gamma_1-\gamma_2}} \left(\frac{a}{r_*}\right)^{\gamma_2-1}}. \quad (36.222)$$

Таким образом, решение для конической цилиндрической дислокации имеет вид

$$ds^2 = A^2 dr^2 + B^2 d\varphi^2 + dz^2, \quad (36.223)$$

где

$$A = \begin{cases} D_1, & 0 \leq r \leq r_*, \\ \frac{1}{\alpha}(E_1 \gamma_1 r^{\gamma_1-1} + E_2 \gamma_2 r^{\gamma_2-1}), & r_* \leq r \leq a, \end{cases} \quad (36.224)$$

$$B = \begin{cases} D_1 r, & 0 \leq r \leq r_*, \\ E_1 r^{\gamma_1} + E_2 r^{\gamma_2}, & r_* \leq r \leq a, \end{cases} \quad (36.225)$$

и постоянные D_1 и $E_{1,2}$ с помощью формул (36.221), (36.222) выражаются через постоянные Θ , r_* , и a , определяющие задачу. Все компоненты метрики являются непрерывными функциями, а первая производная функции B имеет скачок Θ на поверхности склейки $r = r_*$.

Теперь дадим физическую интерпретацию постоянной Θ . Предположим, что радиус цилиндра a является большим, $a \gg r_*$. Тогда вблизи внешней поверхности цилиндра $r \sim a$ постоянные примерно равны

$$D_1 \approx \frac{\alpha(\gamma_1 - \gamma_2)}{\gamma_1(\alpha - \gamma_2)} \left(\frac{r_*}{a}\right)^{\gamma_1-1},$$

$$E_1 \approx \frac{\alpha}{\gamma_1 a^{\gamma_1-1}},$$

и метрика (36.223) асимптотически принимает вид

$$ds^2 = \left(\frac{r}{a}\right)^{2\gamma_1-2} \left(dr^2 + \frac{\alpha^2}{\gamma_1^2} r^2 d\varphi^2\right) + dz^2. \quad (36.226)$$

Это в точности метрика конической дислокации (36.99) с углом дефицита Θ . Таким образом, постоянная Θ , стоящая перед δ -функцией в источнике (36.211), совпадает с углом дефицита среды, которая удалена, $-2\pi < \Theta < 0$, или добавлена, $\Theta > 0$.

С физической точки зрения мы нашли метрику для конической цилиндрической дислокации, которая определяет упругие напряжения вокруг дефекта. Заметим, что это решение зависит от упругих постоянных, характеризующих среду, в отличие от метрики (36.188), полученной ранее. Метрика является решением поставленной задачи даже без нахождения векторного поля смещений. Однако нахождение вектора смещений инструктивно и помогает представить процесс создания дислокации.

Во внутренней области двумерная часть метрики (36.223) имеет вид

$$dl_{\text{in}}^2 = D_1^2 dr^2 + D_1^2 r^2 d\varphi^2.$$

В терминах новой радиальной координаты $y := D_1 r$ метрика становится евклидовой,

$$dl_{\text{in}}^2 = dy^2 + y^2 d\varphi^2, \quad (36.227)$$

причем поверхность склейки $r = r_*$ соответствует следующим значениям координат $r_1 = y(r_*) = D_1 r_*$. Это означает, что цилиндр $y \leq r_1$ в евклидовом пространстве y, φ, z отображается на внутреннюю область конической цилиндрической дислокации. Векторное поле смещений (36.1) при этом имеет только одну отличную от нуля компоненту $u^{\hat{r}} := r - y = (1 - D_1)r$.

Двумерная часть метрики во внешней области имеет вид

$$dl_{\text{ex}} = \frac{1}{\alpha^2} (E_1 \gamma_1 r^{\gamma_1 - 1} + E_2 \gamma_2 r^{\gamma_2 - 1})^2 dr^2 + (E_1 r^{\gamma_1} + E_2 r^{\gamma_2})^2 d\varphi^2.$$

Преобразование к новым координатам,

$$y := \frac{1}{\alpha} (E_1 r^{\gamma_1} + E_2 r^{\gamma_2}), \quad \varphi' := \alpha \varphi, \quad (36.228)$$

приводит метрику к евклидову виду

$$dl_{\text{ex}}^2 = dy^2 + y^2 d\varphi'^2.$$

Новые координаты имеют следующую область определения

$$r_2 < y < \infty, \quad 0 < \varphi' < 2\pi\alpha,$$

где

$$r_2 := \frac{1}{\alpha} (E_1 r_*^{\gamma_1} + E_2 r_*^{\gamma_2}).$$

Это значит, что внешность цилиндра $y > r_2$ евклидова пространства с вырезанным клином $2\pi\alpha < \varphi' < 2\pi$ отображается на внешнюю часть конической цилиндрической дислокации. В этой области векторное поле смещений имеет две нетривиальные компоненты:

$$u^{\hat{r}} = r - \frac{1}{\alpha} (E_1 r^{\gamma_1} + E_2 r^{\gamma_2}), \quad u^{\hat{\varphi}} = -r\Theta\varphi,$$

что вытекает из преобразования координат (36.228). Напомним, что $u^{\hat{\varphi}} := ru^{\varphi}$.

Нетрудно проверить, что длина внутренней окружности равна длине оставшейся дуги от внешней окружности, $2\pi r_1 = 2\pi r_2 \alpha$, в соответствии с равенством (36.220). Это обеспечивает непрерывность компоненты метрики $g_{\varphi\varphi}$ на поверхности склейки.

Процесс создания конической цилиндрической дислокации показан на рис. 36.28. Для отрицательного угла дефицита Θ мы берем евклидово пространство $\mathbb{R}^3$ с цилиндрической системой координат y, φ', z , вырезаем из него толстостенную трубу $r_1 < y < r_2$ и клин с углом $2\pi\alpha < \varphi' < 2\pi$ из наружной части среды. Затем склеиваем края разрезов так, как показано на рисунке. После этого среда приходит в некоторое равновесное состояние в соответствии с упругой калибровкой, которое называется *конической цилиндрической дислокацией*.

Для положительного угла дефицита клин добавляется к внешней части среды. Поэтому внутренний цилиндр перед склейкой необходимо сжать, т.к. в этом случае $r_1 > r_2$.

Вероятно, эту задачу можно решить в рамках теории упругости, хотя решение не известно и представляется непростым из-за сложных граничных условий. В любом случае решение в рамках линейной теории упругости будет воспроизводить только линейное приближение, поскольку упругая калибровка (36.43) соответствует нелинейной теории упругости. Данный пример показывает, что некоторые задачи проще решаются в рамках геометрического подхода.

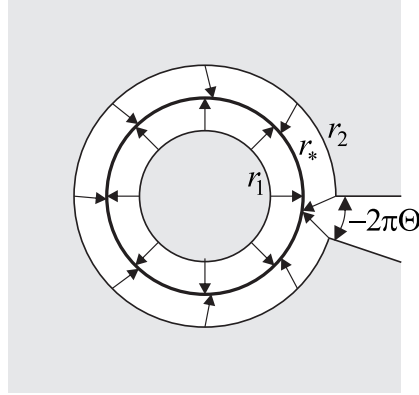


Рис. 36.28: Коническая цилиндрическая дислокация с отрицательным углом дефицита Θ .

Не просто представить себе пространство-время как топологическое произведение вещественной прямой $t \in \mathbb{R}$ с цилиндром конечного радиуса a , снабженного метрикой (36.223), потому что оно геодезически неполно (геодезическая достигает края цилиндра за конечное собственное время). Однако выражение для метрики (36.223) имеет смысл для всех значений радиуса $0 < r < \infty$. Поэтому решение для метрики можно продолжить на все пространство-время

$$ds^2 = dt^2 - A^2 dr^2 - B^2 d\varphi^2 - dz^2,$$

где $t, z \in \mathbb{R}$, $0 < r < \infty$ и $0 < \varphi < 2\pi$. Это – решение четырехмерных уравнений Эйнштейна с тензором энергии-импульса (36.212), описывающим статическую коническую цилиндрическую дислокацию. Эта метрика записана в упругой калибровке и поэтому явно зависит от коэффициента Пуассона σ вселенной. Если этот объект существует в природе, то, в принципе, коэффициент Пуассона вселенной можно измерить. Таким образом, геометрическая теория дефектов дает возможность измерить упругие постоянные вселенной.

36.10.5 Асимптотически плоская клиновая дислокация

Процесс создания дислокации из предыдущего раздела можно обратить. Возьмем бесконечный стержень среды цилиндрической формы радиуса r_1 и вырежем из него клин с углом Θ как показано на рис. 36.29. После склейки получим стержень с конической дислокацией, периметр сечения которого равен $2\pi(1 + \Theta)$. Затем вырежем из бесконечной среды цилиндр меньшего радиуса $r_2 < r_1$ таким образом, чтобы длина окружности $2\pi r_2$ была равна периметру стержня, и вставим стержень внутрь,

$$2\pi\alpha r_1 = 2\pi r_2, \quad \alpha := 1 + \Theta. \quad (36.229)$$

В результате получим асимптотически плоскую коническую дислокацию, причем равенство (36.229) обеспечивает непрерывность $g_{\varphi\varphi}$ компоненты метрики.

Если рассматривать этот дефект в гравитации, то у нас есть космическая струна, окруженная цилиндрической оболочкой материи таким образом, что пространство-время вдали от струны является асимптотически плоским.

Теперь опишем асимптотически плоскую клиновую дислокацию в рамках геометрической теории дефектов. Источник для этого дефекта отличается знаком от

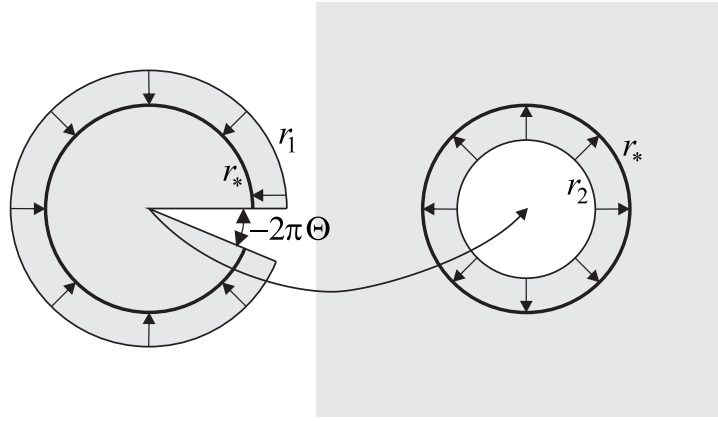


Рис. 36.29: Асимптотически плоская коническая особенность для отрицательного угла дефицита Θ .

источника для конической цилиндрической дислокации:

$$T_{zz} = -2\Theta\delta(r - r_*). \quad (36.230)$$

Кроме того, будут изменены граничные условия. В дальнейшем мы покажем, что для противоположного знака источника уравнения Эйнштейна не имеют решений, которые были бы плоскими вне поверхности склейки $r > r_*$. Поэтому решение с отрицательным углом дефицита, $\Theta < 0$, имеет физический смысл в общей теории относительности.

Интегрирование уравнений Эйнштейна аналогично случаю конической цилиндрической дислокации, рассмотренной в предыдущем разделе,

$$B = -\Theta \int_{r_*}^r ds A(s) + c_1 \int_0^r ds A(s) + c_2.$$

Наложим следующие граничные условия на линии дислокации и в бесконечности:

$$B_{\text{in}}|_{r=0} = 0, \quad \frac{r A_{\text{in}}}{B_{\text{in}}}|_{r=0} = \frac{\gamma_1}{\alpha}, \quad A_{\text{ex}}|_{r=\infty} = 1, \quad (36.231)$$

где γ_1 – положительный корень уравнения (36.217). Таким образом, правая часть второго уравнения выражена только через угол дефицита Θ и коэффициент Пуассона. Первые два граничных условия соответствуют конической особенности на оси z , где метрика должна иметь вид (36.226). Третье граничное условие соответствует асимптотически плоской метрике.

Поставленные граничные условия (36.231) отличаются от граничных условий (36.215) в нуле, которые были поставлены для конической цилиндрической дислокации в предыдущем разделе.

Условие сшивки решений на цилиндре радиуса r_* – это условие непрерывности (36.216) такое же, как и для конической цилиндрической дислокации.

Во внутренней области имеют место равенства:

$$A_{\text{in}} = \frac{B'_{\text{in}}}{c_1},$$

$$B_{\text{in}} = D_1 r^{\gamma_1} + D_2 r^{\gamma_2}, \quad D_{1,2} = \text{const},$$

и первые два граничных условия (36.231) определяют три постоянных интегрирования:

$$c_1 = \alpha, \quad c_2 = 0, \quad D_2 = 0.$$

Во внешней области

$$B'_{\text{ex}} = -\Theta A_{\text{ex}} + \alpha A_{\text{ex}} = A_{\text{ex}},$$

и мы видим необходимость выбора отрицательного знака в источнике (36.230) для асимптотической плоскости. Таким образом, внешнее решение имеет вид

$$\begin{aligned} A_{\text{ex}} &= E_1 - \frac{E_2}{r^2}, \\ B_{\text{ex}} &= E_1 r + \frac{E_2}{r}, \quad E_{1,2} = \text{const.} \end{aligned}$$

Третье граничное условие (36.231) определяет $E_1 = 1$.

Оставшиеся свободными постоянные D_1 и E_2 определяются требованием непрерывности решения (36.216),

$$\begin{aligned} D_1 r_*^{\gamma_1} &= r_* + \frac{E_2}{r_*}, \\ \frac{D_1 \gamma_1 r_*^{\gamma_1-1}}{\alpha} &= 1 - \frac{E_2}{r_*^2}. \end{aligned}$$

Эти уравнения легко решаются:

$$\begin{aligned} D_1 &= \frac{2\alpha}{\alpha + \gamma_1} r_*^{-\gamma_1+1}, \\ E_2 &= \frac{\alpha - \gamma_1}{\alpha + \gamma_1} r_*^2. \end{aligned}$$

В итоге мы получаем следующие выражения для существенной двумерной части метрики внутри и вне поверхности склейки:

$$dl_{\text{in}}^2 = \frac{4\gamma_1^2}{(\alpha + \gamma_1)^2} \left(\frac{r}{r_*} \right)^{2\gamma_1-2} \left(dr^2 + \frac{\alpha^2}{\gamma_1^2} r^2 d\varphi^2 \right), \quad (36.232)$$

$$dl_{\text{ex}}^2 = \left[1 - \frac{\alpha - \gamma_1}{\alpha + \gamma_1} \left(\frac{r_*}{r} \right)^2 \right]^2 dr^2 + \left[1 + \frac{\alpha - \gamma_1}{\alpha + \gamma_1} \left(\frac{r_*}{r} \right)^2 \right]^2 r^2 d\varphi^2. \quad (36.233)$$

Их компоненты являются непрерывными на поверхности склейки, и внешняя часть является асимптотически евклидовой.

Преобразование к декартовым координатам, где метрика имеет вид $dl^2 = dy^2 + y^2 d\varphi^2$ является различным для внутренней и внешней области. Внутри поверхности склейки оно имеет вид

$$y := \frac{2r_*}{\alpha + \gamma_1} \left(\frac{r}{r_*} \right)^{\gamma_1}, \quad \varphi' := \alpha\varphi.$$

Отсюда следует, что мы имеем коническую особенность с углом дефицита Θ и

$$r_1 = \frac{2r_*}{\alpha + \gamma_1}.$$

Вне поверхности склейки преобразование координат имеет вид

$$y := r + \frac{\alpha - \gamma_1}{\alpha + \gamma_1} \frac{r_*^2}{r}, \quad \varphi' := \varphi,$$

и

$$r_2 = \frac{2\alpha r_*}{\alpha + \gamma_1}.$$

Для отрицательного угла дефицита, $\Theta < 0$, $\alpha < 1$, справедливо неравенство $r_2 < r_1$ как это следует из элементарных геометрических построений.

Таким образом, задача с δ -образным источником имеет два типа решений. Один тип является плоским внутри поверхности склейки и коническим вне. Другой тип решений для метрики имеет коническую особенность внутри и является плоским вне поверхности склейки. Соответствующие решения в рамках теории упругости неизвестны. Скорее всего их можно найти, но не стоит ожидать, что они будут иметь простой вид из-за сложных граничных условий.

36.10.6 Непрерывное распределение цилиндрических дислокаций

Для непрерывного распределения цилиндрических дислокаций источник в уравнениях Эйнштейна (36.190) имеет вид

$$T_{zz} = 2f(r), \quad (36.234)$$

где $f(r)$ – произвольная достаточно гладкая скалярная плотность, зависящая только от радиуса. Как и раньше, эта задача инвариантна относительно трансляций вдоль оси z и вращений в плоскости r, φ . Уравнения Эйнштейна (36.190) для метрики (36.193) сводятся к обыкновенному дифференциальному уравнению

$$\left(\frac{B'}{A} \right)' = f(r),$$

которое легко интегрируется

$$B = \int_0^r ds A(s) \int_0^s dt f(t) + c_1 \int_0^r ds A(s) + c_2, \quad c_{1,2} = \text{const.} \quad (36.235)$$

Одна постоянная интегрирования фиксируется требованием, чтобы длина окружности цилиндра, окружающего ось z , стремилась к нулю при $r \rightarrow 0$. Тогда граничное условие принимает вид $B|_{r=0} = 0$, и, как следствие, $c_2 = 0$. Если метрика является конформно евклидовой на оси z , то тогда возникает второе граничное условие

$$\left. \frac{B'}{A} \right|_{r=0} = 1.$$

Оно фиксирует $c_1 = 1$. Это построение проясняет геометрический смысл постоянных интегрирования. Однако можно также рассмотреть более экзотические источники (тензоры энергии-импульса) с другими значениями постоянных $c_{1,2}$.

Упругая калибровка для решения (36.235) сводится к уравнению

$$\frac{B''}{F} - \frac{B'f}{F^2} + \frac{B'}{Fr} - \frac{B}{r^2} + \frac{\sigma}{1-2\sigma} \left(\frac{B''}{F} - \frac{B'f}{F^2} + \frac{B'}{r} - \frac{B}{r^2} \right) = 0,$$

где

$$F(r) := \int_0^r ds f(s) + c_1$$

– первообразная для источника $f(r)$. Это уравнение для заданного источника можно решить по крайней мере численно.

Таким образом, мы получили решение для произвольного распределения цилиндрических дислокаций. Эта задача показывает громадное преимущество геометрической теории дефектов перед теорией упругости. Действительно, если источник всюду отличен от нуля, $f(r) \neq 0$, тогда кривизна отлична от нуля в силу уравнений Эйнштейна, и векторное поле смещений просто не существует. В то же время в геометрической теории дефектов все хорошо определено, и можно найти метрику (т.е. упругие напряжения) как решение уравнений Эйнштейна в упругой калибровке. Векторное поле смещений может быть построено только в тех областях пространства, где источники отсутствуют, $f(r) = 0$, (напомним, что в трех измерениях полный тензор кривизны выражается через тензор Риччи, который равен нулю в силу уравнений Эйнштейна). В этих областях вектор смещений автоматически удовлетворяет уравнениям теории упругости благодаря упругой калибровке.

Таким образом, мы показали, что уравнения Эйнштейна допускают решения с δ -образным тензором энергии-импульса. В общем случае компоненты метрики для цилиндрической дислокации являются разрывными функциями и приводят к неопределенным компонентам тензора кривизны. Однако все неопределенные члены в уравнениях Эйнштейна благополучно сокращаются. Последние сводятся к линейным неоднородным обыкновенным дифференциальным уравнениям для специальных комбинаций компонент метрики.

В общей теории относительности и геометрической теории дефектов полученные в последних разделах решения описывают тонкие массивные оболочки и цилиндрические дислокации. С математической точки зрения это одна и та же модель, т.к. в обоих случаях метрика удовлетворяет уравнениям Эйнштейна (в отсутствие дисклинаций). Отличие сводится только к упругой калибровке, которая имеет физический смысл в геометрической теории дефектов и явно зависит от коэффициента Пуассона, характеризующего упругие свойства среды. Если мы предположим, что пространство-время заполнено упругим эфиром, то, поскольку метрика явно зависит от коэффициента Пуассона, его можно измерить для вселенной.

36.10.7 Двустенная трубка

Сначала опишем двустенную трубку с цилиндрической дислокацией в рамках теории упругости. Цилиндрическая дислокация в бесконечной упругой среде была описана в разделе 36.10.

Пусть в евклидовом пространстве $\mathbb{R}^3$ задана цилиндрическая система координат $\{x^\mu\} = \{r, \varphi, z\}$, $\mu = 1, 2, 3$. Рассмотрим две толстостенные трубки $r_0 \leq r \leq r_1$ и $r_2 \leq r \leq r_3$, сделанные из упругого материала, оси которых совпадают с осью z . Мы предполагаем, что выполнены неравенства $r_0 < r_1 < r_2 < r_3$ (см. рис. 36.30а, где показано сечение $z = \text{const}$). Теперь создадим двустенную трубку с цилиндрической дислокацией внутри следующим образом. Растянем симметрично внутреннюю трубку и сожмем внешнюю. Затем склеим внешнюю поверхность внутренней трубки с внутренней поверхностью внешней трубки. После этого среда придет в некоторое равновесное состояние. Благодаря вращательной и трансляционной симметрии, мы получим одну двустенную трубку $r_{\text{in}} \leq r \leq r_{\text{ex}}$, ось которой совпадает с осью z (см.

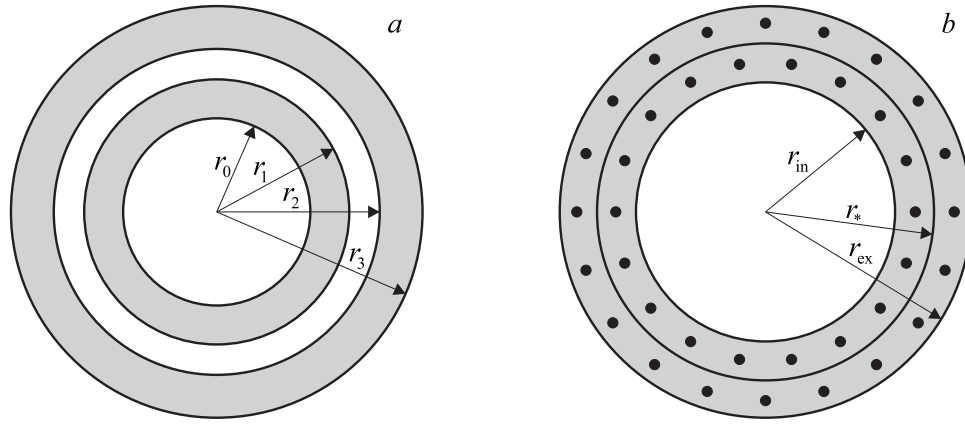


Рис. 36.30: Сечение двустенной трубки $z = \text{const}$ до (a) и после (b) создания цилиндрической дислокации.

рис. 36.30b). Радиусы поверхностных цилиндров отображаются следующим образом:

$$r_0 \mapsto r_{\text{in}}, \quad r_1, r_2 \mapsto r_*, \quad r_3 \mapsto r_{\text{ex}}.$$

Склейка проходит по цилиндру радиуса r_* , где расположена цилиндрическая дислокация, т.к. часть вещества между исходными трубками была удалена.

Полученная таким образом двустенная трубка обладает вращательной и трансляционной инвариантностью.

Построенная модель трубки можно рассматривать в качестве континуальной модели двустенной нанотрубки (см. обзоры [316, 317, 318]). Действительно, рассмотрим нанотрубку с двумя слоями атомов. Предположим, что по окружности слоев располагается, соответственно, 18 и 20 атомов, которые изображены жирными точками на рис. 36.30b. Естественной мерой длины в таком случае является межатомное расстояние. Тогда длина окружности претерпевает скачок, когда мы переходим от одного слоя к другому. В геометрической теории дефектов это означает, что $g_{\varphi\varphi}$ компонента метрики в цилиндрической системе координат также имеет скачок. Соответствующая модель построена ниже.

Для того, чтобы найти радиусы r_{in} , r_* и r_{ex} необходимо решить соответствующую граничную задачу в теории упругости.

Поле смещений определено равенством (36.1). Здесь проявляется преимущество геометрического подхода, т.к. до создания дислокации мы имели несвязное многообразие, а после создания – связное, и поле смещений удобнее рассматривать как функцию от точки среды после создания дефекта. В равновесном состоянии векторное поле смещений удовлетворяет уравнению равновесия

$$(1 - 2\sigma)\Delta u_i + \partial_i \partial_j u^j = 0, \quad (36.236)$$

где σ – коэффициент Пуассона и Δ – лапласиан. Для удобства, мы рассматриваем компоненты поля смещений относительно ортонормального базиса

$$u = u^{\hat{r}} e_{\hat{r}} + u^{\hat{\varphi}} e_{\hat{\varphi}} + u^{\hat{z}} e_{\hat{z}},$$

где

$$e_{\hat{r}} = \partial_r, \quad e_{\hat{\varphi}} = \frac{1}{r} \partial_{\varphi}, \quad e_{\hat{z}} = \partial_z.$$

Индексы компонент относительно ортонормального базиса шляпкой помечены шляпкой:

$$\{i\} = \{\hat{r}, \hat{\varphi}, \hat{z}\}, \quad \{\mu\} = \{r, \varphi, z\}.$$

Латинские индексы поднимаются и опускаются с помощью символа Кронекера: $u_i := u^j \delta_{ji}$.

Дивергенция и лапласиан имеют следующий вид в цилиндрических координатах (см. раздел 7.2):

$$\begin{aligned} \nabla_i u^i &= \frac{1}{r} \partial_r (r u^{\hat{r}}) + \frac{1}{r} \partial_{\varphi} u^{\hat{\varphi}} + \partial_z u^{\hat{z}}, \\ \Delta u_{\hat{r}} &= \frac{1}{r} \partial_r (r \partial_r u_{\hat{r}}) + \frac{1}{r^2} \partial_{\varphi\varphi}^2 u_{\hat{r}} + \partial_{zz}^2 u_{\hat{r}} - \frac{1}{r^2} u_{\hat{r}} - \frac{2}{r^2} \partial_{\varphi} u_{\hat{\varphi}}, \\ \Delta u_{\hat{\varphi}} &= \frac{1}{r} \partial_r (r \partial_r u_{\hat{\varphi}}) + \frac{1}{r^2} \partial_{\varphi\varphi}^2 u_{\hat{\varphi}} + \partial_{zz}^2 u_{\hat{\varphi}} - \frac{1}{r^2} u_{\hat{\varphi}} + \frac{2}{r^2} \partial_{\varphi} u_{\hat{r}}, \\ \Delta u_{\hat{z}} &= \frac{1}{r} \partial_r (r \partial_r u_{\hat{z}}) + \frac{1}{r^2} \partial_{\varphi\varphi}^2 u_{\hat{z}} + \partial_{zz}^2 u_{\hat{z}}. \end{aligned} \tag{36.237}$$

Из симметрии задачи вытекает, что только радиальная компонента вектора смещений отлична от нуля, и она не зависит от полярного угла φ и координаты z :

$$\{u^i\} = \{u^{\hat{r}} := u(r), u^{\hat{\varphi}} = 0, u^{\hat{z}} = 0\}.$$

Уравнение (36.236) для нулевых компонент $u_{\hat{\varphi}}$ и $u_{\hat{z}}$ выполняется автоматически. Нетрудно проверить, что радиальная производная дивергенции,

$$\partial_{\hat{r}} \partial_j u^j = \partial_r \left(\frac{1}{r} \partial_r (r u) \right) = \partial_{rr}^2 u + \frac{1}{r} \partial_r u - \frac{1}{r^2} u,$$

совпадает с лапласианом

$$\Delta u_{\hat{r}} = \frac{1}{r} \partial_r (r \partial_r u) - \frac{1}{r^2} u = \partial_{rr}^2 u + \frac{1}{r} \partial_r u - \frac{1}{r^2} u.$$

Поэтому радиальная компонента уравнения (36.236) принимает вид

$$\partial_r \left(\frac{1}{r} \partial_r (r u) \right) = 0. \tag{36.238}$$

Общее решение данного уравнения зависит от двух постоянных интегрирования:

$$u = c_1 r + \frac{c_2}{r}, \quad c_{1,2} = \text{const.}$$

Отметим, что уравнение равновесия (36.238) не зависит от коэффициента Пуассона σ . Это значит, что цилиндрическая дислокация является геометрическим дефектом, не зависящим от упругих свойств среды.

Для определения постоянных интегрирования наложим граничные условия. С этой целью введем обозначение для внутренней и внешней трубок:

$$u = \begin{cases} u_{\text{in}}, & r_{\text{in}} \leq r \leq r_*, \\ u_{\text{ex}}, & r_* \leq r \leq r_{\text{ex}}. \end{cases}$$

Теперь необходимо наложить граничные условия. Мы предполагаем, что поверхность двустенной трубки является свободной, т.е. тензор деформаций на границе равен нулю:

$$\left. \frac{du_{\text{in}}}{dr} \right|_{r=r_{\text{in}}} = 0, \quad \left. \frac{du_{\text{ex}}}{dr} \right|_{r=r_{\text{ex}}} = 0. \quad (36.239)$$

Мы также предполагаем, что среда находится в равновесии. Это значит, что на поверхности склейки упругие силы равны:

$$\left. \frac{du_{\text{in}}}{dr} \right|_{r=r_*} = \left. \frac{du_{\text{ex}}}{dr} \right|_{r=r_*}. \quad (36.240)$$

Каждое из граничных условий (36.239) определяет одну постоянную интегрирования для внутреннего и внешнего решения:

$$\begin{aligned} u_{\text{in}} &= a \left(r + \frac{r_{\text{in}}^2}{r} \right) > 0, & a &= \text{const} > 0, \\ u_{\text{ex}} &= -b \left(\frac{1}{r} + \frac{r}{r_{\text{ex}}^2} \right) < 0, & b &= \text{const} > 0. \end{aligned} \quad (36.241)$$

Знаки постоянных интегрирования a и b подобраны таким образом, что поле смещений положительно и отрицательно, соответственно, для внутренней и внешней трубок. Это согласуется с поставленной задачей.

Подстановка полученных решений (36.241) в условие склейки (36.240) определяет отношение постоянных интегрирования:

$$r_*^2 = r_{\text{ex}}^2 \frac{ar_{\text{in}}^2 + b}{ar_{\text{ex}}^2 + b} \quad \Leftrightarrow \quad b = ar_{\text{ex}}^2 \frac{r_*^2 - r_{\text{in}}^2}{r_{\text{ex}}^2 - r_*^2}. \quad (36.242)$$

Условие сплошности среды имеет вид

$$\begin{aligned} r_* &= r_1 + a \left(r_* + \frac{r_{\text{in}}^2}{r_*} \right), \\ r_* &= r_2 - b \left(\frac{1}{r_*} + \frac{r_*}{r_{\text{ex}}^2} \right). \end{aligned} \quad (36.243)$$

Из этих равенств находим расстояние между начальными трубками, которое характеризует величину дефекта:

$$l := r_2 - r_1 = 2ar_* \frac{r_{\text{ex}}^2 - r_{\text{in}}^2}{r_{\text{ex}}^2 - r_*^2}, \quad (36.244)$$

где было использовано выражение для b (36.242). После этого определяются все постоянные интегрирования:

$$a = \frac{l}{2r_*} \frac{r_{\text{ex}}^2 - r_*^2}{r_{\text{ex}}^2 - r_{\text{in}}^2}, \quad b = \frac{lr_{\text{ex}}^2}{2r_*} \frac{r_*^2 - r_{\text{in}}^2}{r_{\text{ex}}^2 - r_{\text{in}}^2}. \quad (36.245)$$

Таким образом, найдено векторное поле смещений для двустенной трубки

$$u(r) = \begin{cases} a \left(r + \frac{r_{\text{in}}^2}{r} \right) > 0, & r_{\text{in}} \leq r < r_*, \\ -b \left(\frac{1}{r} + \frac{r}{r_{\text{ex}}^2} \right) < 0, & r_* < r \leq r_{\text{ex}}, \end{cases} \quad (36.246)$$

где постоянные a и b определены равенствами (36.245). Качественное поведение векторного поля смещений показано на рис. 36.31*a*. Дифференцирование этого векторного поля в областях $r_{\text{in}} < r < r_*$, $r_* < r < r_{\text{ex}}$ и его продолжение в точку r_* по непрерывности дает функцию

$$v(r) := \frac{du}{dr} = \begin{cases} a \left(1 - \frac{r_{\text{in}}^2}{r^2}\right) > 0, & r_{\text{in}} \leq r \leq r_*, \\ b \left(\frac{1}{r^2} - \frac{1}{r_{\text{ex}}^2}\right) > 0, & r_* \leq r \leq r_{\text{ex}}, \end{cases} \quad (36.247)$$

которая показана на рис. 36.31*b*.

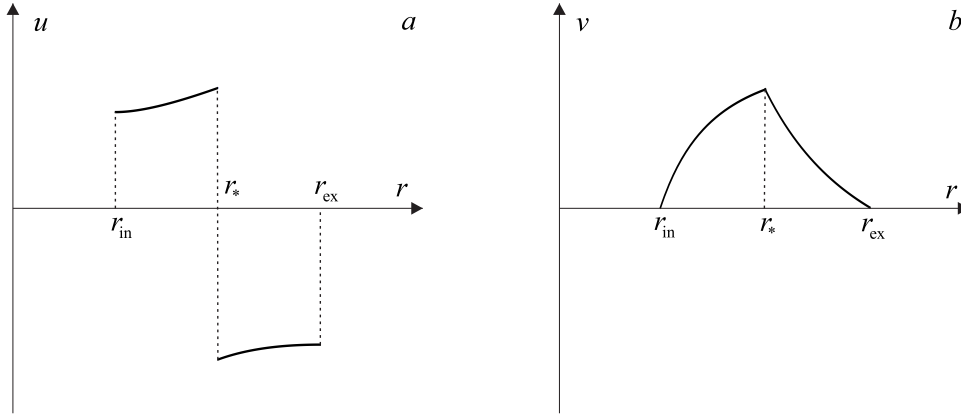


Рис. 36.31: Качественное поведение радиальной компоненты векторного поля смещений для двустенной трубки (*a*). Производная поля смещений (*b*).

Отметим, что условие сплошности среды (36.243) приводит к скачку вектора смещений в точке r_* :

$$l := r_2 - r_1 = u_{\text{in}}(r_*) - u_{\text{ex}}(r_*).$$

Поскольку векторное поле u имеет скачок в точке r_* , то формальная производная u содержит $\delta(r - r_*)$. Эта δ -функция отбрасывается в геометрической теории дефектов.

Двустенная трубка параметризуется четырьмя постоянными r_0 , r_1 , r_2 , r_3 или r_{in} , r_* , r_{ex} , l . Формулы (36.245) определяют постоянные a , b и, следовательно, векторное поле смещений через второй набор параметров. Из определения (36.1) вытекает, что между двумя наборами параметров существует взаимно однозначное соответствие.

Теперь вычислим метрику, индуцированную в двустенной трубке. По определению она имеет вид

$$g_{\mu\nu}(x) = \frac{\partial y^\rho}{\partial x^\mu} \frac{\partial y^\sigma}{\partial x^\nu} \overset{\circ}{g}_{\rho\sigma}(y), \quad (36.248)$$

где $\overset{\circ}{g}_{\rho\sigma}(y)$ – евклидова метрика в цилиндрической системе координат. Соотношение между координатами до и после создания дислокации задано равенством (36.1). Обозначим цилиндрические координаты до и после создания дефекта, соответственно, через $\{y^\mu\} = \{\tilde{r}, \varphi, z\}$ и $\{x^\mu\} = \{r, \varphi, z\}$. Отметим, что вектор смещения рассматривается в координатах r, φ, z после создания дислокации. Тогда исходная евклидова метрика имеет вид

$$ds^2 = \overset{\circ}{g}_{\mu\nu} dy^\mu dy^\nu = d\tilde{r}^2 + \tilde{r}^2 d\varphi^2 + dz^2.$$

Подстановка сюда равенства $\tilde{r} = r - u$, где вектор смещения задан уравнением (36.246), приводит к индуцированной метрике в двустенной трубке

$$ds^2 = (1 - v)^2 dr^2 + (r - u)^2 d\varphi^2 + dz^2. \quad (36.249)$$

Компонента $g_{rr} = (1 - v)^2$ этой метрики является непрерывной функцией, а ее производная имеет скачок в точке r_* . Компонента $g_{\varphi\varphi} = (r - u)^2$ в точке $r = r_*$ имеет скачок.

Элемент объема для метрики (36.249) равен

$$\sqrt{|g|} = (1 - v)(r - u).$$

Правая часть этого равенства положительна, т.к. положителен каждый из сомножителей. Второй сомножитель $r - u = y > 0$ положителен по построению. Первый сомножитель также положителен. Действительно, функция v имеет максимум в точке $r = r_*$. При этом выполнено неравенство

$$v(r_*) = a \frac{r_*^2 - r_{\text{in}}^2}{r_*^2} = \frac{r_* - r_1}{r_*} \frac{r_*^2 - r_{\text{in}}^2}{r_*^2 + r_{\text{in}}^2} < 1,$$

где использовано выражение (36.243) для a .

Длина окружности является геометрическим инвариантом. Она равна $2\pi(r - u(r))$ для метрики (36.248). Когда мы переходим из внутренней трубки к внешней, она имеет скачок $2\pi l$, где l – расстояние между трубками до создания дислокации. Это наблюдение находится в согласии с континуальной моделью двустенной трубки.

Метрика (36.249) получена полностью в рамках теории упругости. К удивлению эта метрика удовлетворяет нелинейным уравнениям Эйнштейна с источником в виде δ' -функции. Это было доказано в разделе 36.10 для $r_0 = 0$ и $r_3 = \infty$. Нетрудно проверить, что то же утверждение верно и для конечных значений r_0 и r_3 . Действительно, тензор Эйнштейна равен нулю всюду (пустое пространство) за исключением поверхности склейки, а сингулярность здесь такая же. Поэтому метрика (36.249) возникает в геометрической теории дефектов как решение уравнений Эйнштейна. Это – исключительный случай, поскольку, как правило, метрика, полученная в рамках теории упругости, отличается от метрики в геометрической теории дефектов (см., например, раздел 36.8.2, где рассмотрена клиновья дислокация).

36.11 Упругие колебания в среде с дислокациями

Среда с дефектами – дислокациями и дисклинациями – является топологически тривиальным многообразием $M \approx \mathbb{R}^3$, на котором задана геометрия Римана–Картана. Если в среде присутствуют дислокации, то евклидова метрика меняется на нетривиальную метрику

$$\delta_{ij} \mapsto g_{\mu\nu}(x) = e_\mu^i e_\nu^j \delta_{ij}, \quad (36.250)$$

которая находится из уравнений Эйнштейна. На этом фоне могут распространяться упругие волны. В разделе 36.9 упругие волны рассматривались в рамках приближения геометрической оптики. Ниже будут найдены точные решения волновых уравнений в отдельных случаях.

Упругие колебания в среде без дефектов описываются зависящим от времени векторным полем $w^i(t, x)$, которое удовлетворяет волновому уравнению (см., например, [297], глава III, §22)

$$\rho_0 \ddot{w}^i - \mu \Delta w^i - (\lambda + \mu) \partial^i \partial_j w^j = 0, \quad (36.251)$$

где точки обозначают дифференцирование по времени, и ρ_0 – плотность среды, которую будем считать постоянной. Эти уравнения получаются из уравнений равновесия

для малых деформаций (36.2), (36.3), если в закон Ньютона (36.2) добавить слагаемое вида “масса $\times$ ускорение”.

Уравнения (36.251) представляют собой систему трех линейных гиперболических однородных уравнений. Все функции, входящие в эту систему уравнений, предполагаются достаточно гладкими, и мы рассматриваем классические решения.

В случае бесконечной среды обычно рассматривается задача Коши, которая определяется некоторыми начальными условиями:

$$w^i(0, x) = w_0^i(x), \quad \dot{w}^i(0, x) = w_1^i(x),$$

где w_0 и w_1 – некоторые заданные векторные функции.

Для (частично) ограниченных упругих тел $\mathbb{M}$ можно рассмотреть задачу Коши, граничную задачу или смешанную задачу, когда одновременно задаются и начальные данные, и граничные условия. В простейших случаях, если на поверхности тела $\partial\mathbb{M}$ заданы перемещения U^i , то ставится граничное условие

$$w^i(t, x) = U^i(t, x), \quad x \in \partial\mathbb{M}.$$

Если на границе упругой среды перемещения не заданы, а определена поверхностная плотность внешних нагрузок p^i , то ставится граничное условие

$$\sigma^{ij}(t, x)n_j(t, x) = p^i(t, x), \quad x \in \partial\mathbb{M}, \quad (36.252)$$

где n^i – компоненты вектора внешней нормали к поверхности тела. При этом внешние силы на поверхности могут иметь как нормальные, так и тангенциальные компоненты. Возможны, конечно, и другие постановки граничных условий. Например, смешанная краевая задача. Мы не будем их обсуждать, т.к. нашей целью является не общность рассмотрения, а демонстрация принципиальных моментов, связанных с появлением дислокаций в упругой среде.

Если до возникновения деформации наблюдаемый инфинитезимальный элемент объема был равен $dy^1 dy^2 dy^3$, то после деформации он будет равен $dx^1 dx^2 dx^3$. Эти элементы объема связаны соотношением

$$dy^1 dy^2 dy^3 = \det \frac{\partial y}{\partial x} dx^1 dx^2 dx^3.$$

В линейном приближении по вектору смещений, который определен равенством (36.1), якобиан перехода $y \mapsto x$ имеет вид

$$\det \frac{\partial y}{\partial x} \approx 1 - \partial_i w^i$$

Дивергенция векторного поля смещений

$$\epsilon := \partial_i w^i = \operatorname{div} w. \quad (36.253)$$

называется *дилатацией*. Это безразмерное поле в линейном приближении описывает относительное сжатие среды в той или иной точке пространства. Поскольку

$$\frac{dx^1 dx^2 dx^3}{dy^1 dy^2 dy^3} \approx 1 + \epsilon,$$

то при $\epsilon < 0$ среда сжимается, а при $\epsilon > 0$ – растягивается. Если $\epsilon = 0$, то инфинитезимальные объемы среды до и после деформации сохраняются.

Для описания упругих волн мы ввели новую букву w , поскольку буква u была использована для обозначения поля смещений, описывающего дислокации. Полное поле смещений равно сумме $u + w$ (там, где u существует). Однако надо помнить, что в геометрической теории дефектов никакого разбиения поля смещений на два слагаемых, которое не является однозначным, нет, т.к. переменная u просто отсутствует. Вместо этого мы решаем уравнения равновесия и находим сразу репер, который задает нетривиальную метрику (36.250), описывающую распределение дефектов.

Если среда содержит дефекты, то метрика пространства становится нетривиальной (36.250). Предположим, что относительные смещения при упругих колебаниях малы по сравнению с напряжениями, создаваемыми дефектами:

$$\partial_\mu w^i \ll e_\mu^i. \quad (36.254)$$

Тогда в первом приближении будем считать, что упругие колебания распространяются в римановом пространстве с нетривиальной метрикой, созданной дислокациями. При этом мы пренебрегаем изменением метрики, вызванным самими упругими колебаниями. Поэтому для упругих колебаний в среде с дефектами мы постулируем следующее уравнение, которое является ковариантным обобщением (36.251) по пространственным переменным

$$\rho_0 \ddot{w}^i - \mu \tilde{\Delta} w^i - (\lambda + \mu) \tilde{\nabla}^i \tilde{\nabla}_j w^j = 0, \quad (36.255)$$

где w^i – компоненты векторного поля смещений относительно ортонормального базиса касательного пространства e_i , $\tilde{\Delta} = \tilde{\nabla}^i \tilde{\nabla}_i$ – ковариантный оператор Лапласа–Бельтрами, построенный по реперу e_μ^i , $\tilde{\nabla}_i$ – оператор ковариантного дифференцирования. Выпишем в явном виде ковариантную производную от поля смещений

$$\tilde{\nabla}_i w^j := e^\mu_i \tilde{\nabla}_\mu w^j = e^\mu_i (\partial_\mu w^j + w^k \tilde{\omega}_{\mu k}^j),$$

где $\tilde{\omega}_{\mu k}^j$ – связность, построенная при нулевом кручении и поэтому помеченная знаком тильды.

Как уже говорилось в разделе 36.2, поле смещений для дислокаций можно ввести только в тех областях среды, где дефекты отсутствуют. При этом репер определяется полем смещений для дислокации $e_\mu^i = \partial_\mu u^i$. Заметим, что малость относительных деформаций для упругих колебаний (36.254) имеет смысл даже в тех областях пространства, где смещения u^i не определены.

Разложим поле смещений w ковариантным образом на поперечную (соленоидальную) w^{Ti} и продольную (потенциальную) части w^{Li} :

$$w^i = w^{Ti} + w^{Li},$$

которые определены следующими равенствами:

$$\tilde{\nabla}_i w^{Ti} = 0, \quad (36.256)$$

$$\tilde{\nabla}_i w_j^L - \tilde{\nabla}_j w_i^L = 0. \quad (36.257)$$

Разложение векторного поля на продольную и поперечную части в трехмерном пространстве является единственным с точностью до постоянной (см. раздел 25.3.1). Напомним, что опускание латинских индексов производится с помощью символов Кронекера, $w^i = w_i$, и оно перестановочно с ковариантным дифференцированием. Последнее уравнение (36.257) можно переписать в виде

$$\tilde{\nabla}_i w_j^L - \tilde{\nabla}_j w_i^L = e^\mu_i e^\nu_j (\tilde{\nabla}_\mu w_\nu^L - \tilde{\nabla}_\nu w_\mu^L) = e^\mu_i e^\nu_j (\partial_\mu w_\nu^L - \partial_\nu w_\mu^L) = 0,$$

т.к. переход от латинских индексов к греческим перестановочен с ковариантным дифференцированием, а символы Кристоффеля симметричны по первым двум индексам. Последнее равенство означает, что 1-форма $dx^\mu w_\mu^L$ замкнута. Нетрудно проверить, что уравнение (36.255) для упругих колебаний, эквивалентно двум независимым уравнениям для поперечных и продольных колебаний:

$$\frac{1}{c_T^2} \ddot{w}^{Ti} - \tilde{\Delta} w^{Ti} = 0, \quad \frac{1}{c_L^2} \ddot{w}^{Li} - \tilde{\Delta} w^{Li} = 0, \quad (36.258)$$

где

$$c_T^2 := \frac{\mu}{\rho_0}, \quad c_L^2 := \frac{\lambda + 2\mu}{\rho_0}$$

– квадраты скоростей звука поперечных и продольных колебаний. Как видим, скорости распространения звука в упругой среде не зависят от наличия дислокаций.

Из определения скоростей звука следует равенство

$$c_L^2 = c_T^2 + \frac{\lambda + \mu}{\rho_0}.$$

Поэтому скорость распространения продольных колебаний всегда больше скорости распространения поперечных.

Разложение векторного поля на продольную и поперечную части на трехмерном римановом многообразии с точностью до постоянной является единственным при достаточно общих граничных условиях. Это разложение в пространстве Минковского для электромагнитного поля рассмотрено в разделе 25.3.1.

Для постановки граничных условий вида (36.252) для волнового уравнения (36.255) в среде с дислокациями необходимо выражение для тензора деформаций σ_{ij} . При наличии дефектов тензор деформаций является ковариантным обобщением выражения (36.4) из классической теории упругости:

$$\epsilon_{ij} := \frac{1}{2} (e^\mu_i \tilde{\nabla}_\mu w_j + e^\mu_j \tilde{\nabla}_\mu w_i). \quad (36.259)$$

В среде с дислокациями дилатация определяется ковариантным образом,

$$\epsilon := \tilde{\nabla}_i w^i, \quad (36.260)$$

и также определяет сжатие среды в первом приближении. Из определения поперечных компонент поля смещений (36.256) вытекает, что поперечные колебания среды w^T происходят без сжатия и приводят только к изменению формы. Продольные колебания, напротив, сопровождаются сжатием и растяжением среды.

Поскольку справедливо равенство $\epsilon = \tilde{\nabla}_i w^{Li}$, то волны дилатации определяются только продольными компонентами и удовлетворяют волновому уравнению.

$$\ddot{\epsilon} - c_L^2 \tilde{\Delta} \epsilon = 0,$$

которое получается после взятия дивергенции от второго уравнения (36.258).

36.11.1 Крутильные волны в волноводе с клиновой дислокацией

В качестве примера применения геометрической теории дефектов рассмотрим распространение крутильных волн в цилиндрическом волноводе радиуса a с клиновой

дислокацией, изображенном на рис. 36.9. Исходя из симметрии задачи, выберем цилиндрическую систему координат r, φ, z , где ось z совпадает с осью волновода. Если $-1 < \theta < 0$, то клин вырезается из волновода. Дислокация отсутствует, если $\theta = 0$. Для положительных углов дефицита, $\theta > 0$, клин среды добавляется к цилиндру.

Из-за наличия дислокации метрика внутри волновода становится неевклидовой (36.99)

$$ds^2 = \left(\frac{r}{a}\right)^{2\gamma-2} \left(dr^2 + \frac{\alpha^2 r^2}{\gamma^2} d\varphi^2\right) + dz^2, \quad (36.261)$$

где $\alpha := 1 + \theta$, и введена безразмерная постоянная

$$\gamma := -\theta b + \sqrt{\theta^2 b^2 + 1 + \theta}. \quad (36.262)$$

Константа $b := \sigma/(2(1 - \sigma))$ определяется безразмерным коэффициентом Пуассона σ , который характеризует упругие свойства среды. Предел

$$\theta \rightarrow 0, \quad \alpha \rightarrow 1, \quad \gamma \rightarrow 1 \quad (36.263)$$

соответствует отсутствию дислокации.

Метрика клиновой дислокации определяет репер e_μ^i , который мы выберем диагональным:

$$e_r^{\hat{r}} = \left(\frac{r}{a}\right)^{\gamma-1}, \quad e_\varphi^{\hat{\varphi}} = \left(\frac{r}{a}\right)^{\gamma-1} \frac{\alpha r}{\gamma}, \quad e_z^{\hat{z}} = 1, \quad (36.264)$$

где $\mu = r, \varphi, z$ и $i = \hat{r}, \hat{\varphi}, \hat{z}$. В отсутствие дислокации, $\gamma = 1, \alpha = 1$, он определяет обычный ортонормальный репер в касательном пространстве в цилиндрической системе координат. Обратный репер имеет вид

$$e_r^{\hat{r}} = \left(\frac{a}{r}\right)^{\gamma-1}, \quad e_\varphi^{\hat{\varphi}} = \left(\frac{a}{r}\right)^{\gamma-1} \frac{\gamma}{\alpha r}, \quad e_z^{\hat{z}} = 1. \quad (36.265)$$

Для того, чтобы найти явный вид волнового оператора, необходимо вычислить символы Кристоффеля и компоненты $\mathbb{SO}(3)$ связности. Прямые вычисления показывают, что только четыре символа Кристоффеля отличны от нуля:

$$\tilde{\Gamma}_{rr}^r = \frac{\gamma-1}{r}, \quad \tilde{\Gamma}_{r\varphi}^\varphi = \tilde{\Gamma}_{\varphi r}^\varphi = \frac{\gamma}{r}, \quad \tilde{\Gamma}_{\varphi\varphi}^r = -\frac{\alpha^2 r}{\gamma}. \quad (36.266)$$

Репер (36.264) определяет также компоненты $\mathbb{SO}(3)$ связности

$$\omega_{\mu i}^j = -\partial_\mu e_\nu^j e^\nu_i + e^\nu_i \tilde{\Gamma}_{\mu\nu}^\rho e_\rho^j.$$

Здесь и в дальнейшем, для краткости, мы опустим знак тильды у компонент $\mathbb{SO}(3)$ связности. Легко проверить, что только две компоненты $\mathbb{SO}(3)$ связности отличны от нуля:

$$\omega_{\varphi\hat{r}}^{\hat{\varphi}} = -\omega_{\varphi\hat{\varphi}}^{\hat{r}} = \alpha. \quad (36.267)$$

Теперь можно вычислить лапласиан

$$\begin{aligned} \tilde{\Delta} w_i &= g^{\mu\nu} \partial_{\mu\nu}^2 w_i - g^{\mu\nu} \partial_\mu \omega_{\nu i}^j w_j - 2g^{\mu\nu} \omega_{\mu i}^j \partial_\nu w_j - \\ &\quad - g^{\mu\nu} \tilde{\Gamma}_{\mu\nu}^\rho (\partial_\rho w_i - \omega_{\rho i}^j w_j) + g^{\mu\nu} \omega_{\mu i}^k \omega_{\nu k}^j w_j. \end{aligned} \quad (36.268)$$

Подстановка явных выражений для символов Кристоффеля (36.266) и $\mathbb{SO}(3)$ связности (36.267) приводит к следующим выражениям:

$$\begin{aligned}\tilde{\Delta}w_{\hat{r}} &= \left(\frac{a}{r}\right)^{2\gamma-2} \frac{1}{r} \partial_r (r \partial_r w_{\hat{r}}) + \left(\frac{a}{r}\right)^{2\gamma-2} \frac{\gamma^2}{\alpha^2 r^2} \partial_{\varphi\varphi}^2 w_{\hat{r}} + \partial_{zz}^2 w_{\hat{r}} - \\ &\quad - \left(\frac{a}{r}\right)^{2\gamma-2} \frac{\gamma^2}{r^2} w_{\hat{r}} - 2 \left(\frac{a}{r}\right)^{2\gamma-2} \frac{\gamma^2}{\alpha r^2} \partial_{\varphi} w_{\hat{\varphi}}, \\ \tilde{\Delta}w_{\hat{\varphi}} &= \left(\frac{a}{r}\right)^{2\gamma-2} \frac{1}{r} \partial_r (r \partial_r w_{\hat{\varphi}}) + \left(\frac{a}{r}\right)^{2\gamma-2} \frac{\gamma^2}{\alpha^2 r^2} \partial_{\varphi\varphi}^2 w_{\hat{\varphi}} + \partial_{zz}^2 w_{\hat{\varphi}} - \\ &\quad - \left(\frac{a}{r}\right)^{2\gamma-2} \frac{\gamma^2}{r^2} w_{\hat{\varphi}} + 2 \left(\frac{a}{r}\right)^{2\gamma-2} \frac{\gamma^2}{\alpha r^2} \partial_{\varphi} w_{\hat{r}}, \\ \tilde{\Delta}w_{\hat{z}} &= \left(\frac{a}{r}\right)^{2\gamma-2} \frac{1}{r} \partial_r (r \partial_r w_{\hat{z}}) + \left(\frac{a}{r}\right)^{2\gamma-2} \frac{\gamma^2}{\alpha^2 r^2} \partial_{\varphi\varphi}^2 w_{\hat{z}} + \partial_{zz}^2 w_{\hat{z}}.\end{aligned}\tag{36.269}$$

Соответствующее волновое уравнение (36.255) имеет много решений. Простейшее решение описывает распространение крутильных волн, для которых только угловая компонента поля смещений $w_{\hat{\varphi}}$ отлична от нуля и не зависит от угла φ :

$$w_{\hat{r}} = 0, \quad w_{\hat{\varphi}} = w_{\hat{\varphi}}(t, r, z), \quad w_{\hat{z}} = 0.$$

Легко видеть, что крутильные колебания происходят без сжатия среды, т.к. дилатация равна нулю

$$\epsilon := \tilde{\nabla}_i w^i = 0.$$

Тем самым, крутильные волны описывают только поперечные колебания.

Для крутильных колебаний волновые уравнения $\tilde{\square}w_{\hat{r}} = 0$ и $\tilde{\square}w_{\hat{z}} = 0$ тождественно удовлетворяются, и остается одно волновое уравнение для поперечных колебаний

$$\frac{1}{c_T^2} \ddot{w}_{\hat{\varphi}} - \left(\frac{a}{r}\right)^{2\gamma-2} \frac{1}{r} \partial_r (r \partial_r w_{\hat{\varphi}}) - \partial_{zz}^2 w_{\hat{\varphi}} + \left(\frac{a}{r}\right)^{2\gamma-2} \frac{\gamma^2}{r^2} w_{\hat{\varphi}} = 0.\tag{36.270}$$

Будем искать решение этого уравнения в виде плоской волны

$$w_{\hat{\varphi}} = \text{re} \left(W e^{i(kz - \omega t)} \right),$$

где $W(r) \in \mathbb{R}$ – амплитуда волны, $k \in \mathbb{R}$ – волновой вектор и $\omega \in \mathbb{R}$ – частота волны. Тогда уравнение (36.270) принимает вид

$$r \partial_r (r \partial_r W) + \kappa^2 a^2 \left(\frac{r}{a}\right)^{2\gamma} W - \gamma^2 W = 0,\tag{36.271}$$

где

$$\kappa^2 := \frac{\omega^2}{c_T^2} - k^2.\tag{36.272}$$

Теперь введем новую радиальную координату

$$r = a r'^{\frac{1}{\gamma}}, \quad 0 < r' < 1.$$

Тогда уравнение (36.272) сведется к уравнению Бесселя

$$r'^2 \frac{d^2 W}{dr'^2} + r' \frac{dW}{dr'} + \lambda^2 r'^2 W - W = 0,\tag{36.273}$$

где

$$\lambda^2 := \frac{\kappa^2 a^2}{\gamma^2}.$$

Общее решение уравнения Бесселя содержит две постоянных интегрирования. Мы требуем, чтобы решение было конечным в нуле $r = 0$. Тогда амплитуда волны примет вид

$$W = C J_1 \left(\lambda \left(\frac{r}{a} \right)^\gamma \right), \quad (36.274)$$

где $C \in \mathbb{R}$ – постоянная интегрирования и J_1 – функция Бесселя первого рода и первого порядка (см., например, [319]).

Если дислокация отсутствует, $\gamma = 1$, то решение принимает хорошо известный вид (см., например, [320], глава E, раздел 73)

$$W \xrightarrow{\gamma \rightarrow 1} C J_1(\kappa r)$$

Для анализа граничных условий необходимо знать тензор деформаций (36.259). Прямые вычисления дают следующий результат

$$\begin{aligned} \epsilon_{\hat{r}\hat{r}} &= \left(\frac{a}{r} \right)^{\gamma-1} \partial_r w_{\hat{r}}, & \epsilon_{\hat{r}\hat{\varphi}} &= \frac{1}{2} \left(\frac{a}{r} \right)^{\gamma-1} \left[\partial_r w_{\hat{\varphi}} + \frac{\gamma}{\alpha r} \partial_\varphi w_{\hat{r}} - \frac{\gamma}{r} w_{\hat{\varphi}} \right], \\ \epsilon_{\hat{\varphi}\hat{\varphi}} &= \left(\frac{a}{r} \right)^{\gamma-1} \frac{\gamma}{\alpha r} (\partial_\varphi w_{\hat{\varphi}} + \alpha w_{\hat{r}}), & \epsilon_{\hat{r}\hat{z}} &= \frac{1}{2} \left[\left(\frac{a}{r} \right)^{\gamma-1} \partial_r w_{\hat{z}} + \partial_z w_{\hat{r}} \right], \\ \epsilon_{\hat{z}\hat{z}} &= \partial_z w_{\hat{z}}, & \epsilon_{\hat{\varphi}\hat{z}} &= \frac{1}{2} \left[\left(\frac{a}{r} \right)^{\gamma-1} \frac{\gamma}{\alpha r} \partial_\varphi w_{\hat{z}} + \partial_z w_{\hat{\varphi}} \right]. \end{aligned} \quad (36.275)$$

Для крутильных волн только две компоненты отличны от нуля:

$$\epsilon_{\hat{r}\hat{\varphi}} = \epsilon_{\hat{\varphi}\hat{r}} = \frac{1}{2} \left(\frac{a}{r} \right)^{\gamma-1} \left[\partial_r w_{\hat{\varphi}} - \frac{\gamma}{r} w_{\hat{\varphi}} \right].$$

Мы требуем, чтобы поверхность волновода была свободной, т.е. упругие силы на поверхности цилиндра должны обращаться в нуль. Это дает граничное условие

$$\left[\partial_r w_{\hat{\varphi}} - \frac{\gamma}{r} w_{\hat{\varphi}} \right]_{r=a}. \quad (36.276)$$

Для решения (36.274) это граничное условие принимает вид

$$\lambda J'_1(\lambda) - J_1(\lambda) = 0,$$

где штрих обозначает производную функции Бесселя по аргументу. Учтем равенство

$$J'_1(\lambda) = J_0(\lambda) - \frac{1}{\lambda} J_1(\lambda).$$

Тогда граничное условие эквивалентно равенству, определяющему дисперсионное соотношение

$$\frac{\kappa a}{\gamma} = \xi \quad \Leftrightarrow \quad \omega = c_T \sqrt{k^2 + \frac{\gamma^2 \xi^2}{a^2}}, \quad (36.277)$$

где ξ – корень уравнения

$$\xi J_0(\xi) = 2 J_1(\xi). \quad (36.278)$$

Функции Бесселя J_ν имеет следующую асимптотику при больших аргументах, $\xi \gg 1$, $\xi \gg \nu$,

$$\begin{aligned} J_0(\lambda) &\approx \sqrt{\frac{2}{\pi\lambda}} \cos\left(\lambda - \frac{\pi}{4}\right), \\ J_1(\lambda) &\approx \sqrt{\frac{2}{\pi\lambda}} \sin\left(\lambda - \frac{\pi}{4}\right). \end{aligned}$$

Тогда уравнение (36.277) принимает вид

$$\xi \cos\left(\xi - \frac{\pi}{4}\right) = 2 \sin\left(\xi - \frac{\pi}{4}\right).$$

Это уравнение имеет счетное число решений, каждое из которых определяет дисперсионное соотношение.

Нетрудно найти фазовую скорость крутильных колебаний $v := \omega/k$

$$v = c_T \sqrt{1 + \frac{\gamma^2 \xi^2}{a^2 k^2}}. \quad (36.279)$$

При этом групповая скорость равна

$$v_g := \frac{d\omega}{dk} = \frac{c_T^2}{v}. \quad (36.280)$$

Если дислокация отсутствует, то уравнение (36.277) для ξ остается прежним. Соотношение (36.280) между фазовой и групповой скоростью также остается прежним. Поэтому наличие клиновой дислокации меняет только дисперсионное соотношение (36.277).

Для малых углов дефицита из уравнения (36.262) следует равенство

$$\gamma \approx 1 + \theta \frac{1 - 2\sigma}{2(1 - \sigma)}$$

и первая поправка к дисперсионному соотношению имеет вид

$$\omega \approx c_T \sqrt{k^2 + \frac{\xi^2}{a^2}} \left[1 + \theta \frac{1 - 2\sigma}{2(1 - \sigma)} \frac{1}{1 + \frac{k^2 a^2}{\xi^2}} \right]. \quad (36.281)$$

Она линейна по углу дефицита.

Таким образом, наличие клиновой дислокации в волноводе приводит к изменению дисперсионного соотношения для крутильных волн. Для положительных углов дефицита $\gamma > 1$ фазовая скорость возрастает, а групповая – падает как следствие уравнений (36.279) и (36.280). Для отрицательных углов дефицита ситуация противоположна.

36.11.2 Крутильные волны в двустенной трубке

По определению векторное поле смещений w , описывающее упругие колебания, удовлетворяет волновому уравнению

$$\rho_0 \ddot{w}_i - \mu \Delta w_i - (\lambda + \mu) \nabla_i \nabla_j w^j = 0, \quad (36.282)$$

где ρ_0 – плотность среды, λ, μ – коэффициенты Ламе, а ковариантная производная $\nabla_i := e^\mu_j \nabla_\mu$ и оператор Лапласа–Бельтрами $\Delta := g^{\mu\nu} \nabla_\mu \nabla_\nu$ определены метрикой (36.249) двустенной трубки. Здесь e^μ_i – обратный репер, $e^\mu_i e^\mu_j = \delta^j_i$.

Волновое уравнение (36.282) ковариантно относительно общих преобразований координат, и может быть решено в цилиндрической системе координат r, φ, z с метрикой (36.249) после создания дефекта. Однако проще поступить по другому. Мы решим волновое уравнение в исходной цилиндрической системе координат $y := \tilde{r}, \varphi, z$, где координата y обозначает старый радиус до создания дислокации

$$y := r - u,$$

а после этого наложим необходимые граничные условия. Это проще, т.к. метрика в исходной системе координат является евклидовой.

Рассмотрим крутильные колебания. В этом случае отлична от нуля только угловая компонента вектора смещений:

$$\{w^i\} = \{w^{\hat{r}} = 0, w^{\hat{\varphi}} = w^{\hat{\varphi}}(t, r, z), w^{\hat{z}} = 0\}.$$

Из соображений симметрии угловая компонента $w^{\hat{\varphi}}$ не зависит от угла φ . Для такого векторного поля $\hat{r}$ и $\hat{z}$ компоненты уравнения (36.282) автоматически удовлетворяются. Дилатация для крутильных волн обращается в нуль

$$\epsilon := \partial_i w^i = 0,$$

т.е. крутильные колебания происходят без сжатия среды.

Ищем решение уравнения (36.282) в виде плоской волны

$$w_{\hat{\varphi}} = \text{re} [W(y) e^{i(kz - \omega t)}], \quad (36.283)$$

где $W(y)$ – амплитуда, $k \in \mathbb{R}$ – волновой вектор и $\omega \in \mathbb{R}$ – частота волны. Тогда уравнение (36.282) в цилиндрических координатах принимает вид

$$y^2 \frac{d^2 U}{dy^2} + y \frac{dU}{dy} + (\kappa^2 y^2 - 1)U = 0, \quad (36.284)$$

где

$$\kappa^2 := \frac{\omega^2}{c_T^2} - k^2, \quad c_T^2 := \frac{\mu}{\rho_0}. \quad (36.285)$$

Общее решение данного уравнения содержит две постоянных интегрирования. Поэтому решение для внутренней и внешней трубок можно записать следующим образом

$$W = \begin{cases} W_{\text{in}} = C_1 J_1(\kappa y) + C_2 N_1(\kappa y), & r_0 \leq y \leq r_1, \\ W_{\text{ex}} = C_3 J_1(\kappa y) + C_4 N_1(\kappa y), & r_2 \leq y \leq r_3, \end{cases} \quad (36.286)$$

где J_1 – функция Бесселя первого рода и первого порядка, N_1 – функция Неймана первого порядка (см., например, [319]) и $C_{1,2,3,4}$ – постоянные интегрирования.

Для того, чтобы найти постоянные интегрирования, наложим граничные условия. Предположим, что внутренняя и внешняя поверхности двустенной трубки являются свободными, т.е. тензор деформаций равен нулю:

$$\left. \frac{dW_{\text{in}}}{dr} \right|_{r=r_{\text{in}}} = 0, \quad \left. \frac{dW_{\text{ex}}}{dr} \right|_{r=r_{\text{ex}}} = 0.$$

Поскольку

$$\frac{dW}{dr} = \frac{dy}{dr} \frac{dW}{dy} = (1-v) \frac{dW}{dy},$$

где $v(r_{\text{in}}) = v(r_{\text{ex}}) = 0$, эти уравнения в исходной системе координат принимают вид

$$\begin{aligned} C_1 J_1'(z_0) + C_2 N_1'(z_0) &= 0 \\ C_3 J_1'(z_3) + C_4 N_1'(z_3) &= 0, \end{aligned}$$

где

$$z := \kappa y,$$

и штрих обозначает дифференцирование по аргументу z . Полученные условия определяют две постоянные интегрирования:

$$\begin{aligned} C_2 &= -k_0 C_1, & k_0 &:= \frac{J_1'(z_0)}{N_1'(z_0)}, \\ C_4 &= -k_3 C_1, & k_3 &:= \frac{J_1'(z_3)}{N_1'(z_3)}. \end{aligned} \quad (36.287)$$

На поверхности склейки наложим два граничных условия: сплошность среды и равенство упругих напряжений,

$$W_{\text{in}}(r_*) = W_{\text{ex}}(r_*), \quad \left. \frac{dW_{\text{in}}}{dr} \right|_{r=r_*} = \left. \frac{dW_{\text{ex}}}{dr} \right|_{r=r_*}. \quad (36.288)$$

В результате возникают два уравнения

$$\begin{aligned} C_1 [J_1(z_1) - k_0 N_1(z_1)] - C_3 [J_1(z_2) - k_3 N_1(z_2)] &= 0, \\ C_1 [J_1'(z_1) - k_0 N_1'(z_1)] - C_3 [J_1'(z_2) - k_3 N_1'(z_2)] &= 0. \end{aligned} \quad (36.289)$$

Необходимым и достаточным условием разрешимости этой системы уравнений является равенство нулю определителя:

$$\begin{aligned} [J_1(z_1) - k_0 N_1(z_1)] [J_1'(z_2) - k_3 N_1'(z_2)] - \\ - [J_1(z_2) - k_3 N_1(z_2)] [J_1'(z_1) - k_0 N_1'(z_1)] = 0. \end{aligned} \quad (36.290)$$

Для заданных параметров двустенной трубки r_0 , r_1 , r_2 и r_3 , это равенство представляет собой уравнение на постоянную κ . Пусть κ является корнем уравнения (36.290), тогда равенство

$$\omega = c_T \sqrt{k^2 + \kappa^2}. \quad (36.291)$$

задает уравнение дисперсии для крутильных волн.

Фазовая скорость крутильных волн $v := \omega/k$ легко находится из дисперсионного соотношения (36.291):

$$v = c_T \sqrt{1 + \frac{\kappa^2}{k^2}}. \quad (36.292)$$

Легко также вычисляется групповая скорость

$$v_g := \frac{d\omega}{dk} = \frac{c_T^2}{v}. \quad (36.293)$$

Мы видим, что фазовая скорость всегда больше поперечной скорости, а групповая – меньше. Дисперсионное соотношение (36.291) зависит от параметров двустенной трубки через соотношение (36.290).

Помимо чисто академического интереса, двустенная трубка может являться континуальной моделью двухслойной нанотрубки.

36.12 Примесь в поле клиновой дислокации

Рассмотрим упругую среду с клиновой дислокацией, которая содержит один атом примеси или вакансию. Если считать влияние примеси на распределение упругих напряжений малым по сравнению с упругими напряжениями, вызванными самой дислокацией, то можно полагать, что движение примеси происходит в трехмерном пространстве с нетривиальной метрикой (36.99). В геометрическом подходе мы считаем, что потенциальная энергия взаимодействия примеси с дислокацией равна нулю, и все взаимодействие происходит за счет изменения кинетической энергии, куда явным образом входит нетривиальная метрика.

Решим соответствующую квантово-механическую задачу. Рассмотрим связанные состояния примеси при наличии клиновой дислокации, двигающейся в цилиндре радиуса a . При этом мы предполагаем, что ось цилиндра совпадает с ядром дислокации. Стационарное уравнение Шредингера имеет вид

$$-\frac{\hbar^2}{2M}\Delta\Psi = E\Psi, \quad (36.294)$$

где $\hbar$ – постоянная Планка, M , Ψ и E – масса, волновая функция и энергия примеси. Нетривиальность взаимодействия примеси с дислокацией сводится к нетривиальному оператору Лапласа–Бельтрами

$$\Delta\Psi = \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}g^{\mu\nu}\partial_\nu\Psi),$$

где метрика была найдена ранее (36.99) и $g := \det g_{\mu\nu}$.

Исходя из симметрии задачи, будем решать уравнение Шредингера (36.294) в цилиндрических координатах методом разделения переменных. Пусть

$$\Psi(r, \varphi, z) = Z(z) \sum_{m=-\infty}^{\infty} \psi_m(r) e^{im\varphi},$$

где для нормированной функции $Z(z)$ имеются две следующие возможности. Если вдоль оси z примесь движется свободно с импульсом $\hbar k$, то

$$Z(z) = \frac{1}{\sqrt{2\pi}} e^{ikz}.$$

Если же ее движение ограничено плоскостями $z = 0$ и $z = z_0$, то

$$Z(z) = \sqrt{\frac{2}{z_0}} \sin(k_l z), \quad k_l = \frac{\pi l}{z_0}.$$

В дальнейшем целочисленный индекс l , возникающий при ограниченном движении, мы для краткости опустим, имея в виду обе возможности.

Условие целочисленности постоянной m (собственного значения проекции момента импульса на ось z) возникает из условия периодичности

$$\Psi(r, \varphi, z) = \Psi(r, \varphi + 2\pi, z).$$

Тогда для радиальной волновой функции $\psi_m(r)$ получаем уравнение

$$\frac{a^{2\gamma-2}}{r^{2\gamma-1}}\partial_r(r\partial_r\psi_m) + \left(\frac{2ME}{\hbar^2} - \frac{\gamma^2}{\alpha^2} \frac{a^{2\gamma-2}}{r^{2\gamma}} m^2 - k^2\right) \psi_m = 0. \quad (36.295)$$

Введем новую радиальную координату

$$\rho := \frac{r^\gamma}{\gamma a^{\gamma-1}}.$$

С точностью до постоянной, это – преобразование (36.96), (36.98). Тогда радиальное уравнение примет вид

$$\frac{1}{\rho} \partial_\rho (\rho \partial_\rho \psi_m) + \left(\beta^2 - \frac{\nu^2}{\rho^2} \right) \psi_m = 0, \quad (36.296)$$

где

$$\beta^2 := \frac{2ME}{\hbar^2} - k^2, \quad \nu := \frac{|m|}{\alpha} > 0.$$

Это есть уравнение Бесселя. Будем решать его при граничном условии

$$\psi_m|_{\rho=a/\gamma} = 0, \quad (36.297)$$

что соответствует движению примеси внутри цилиндра с непроницаемой границей. Общее решение уравнения Бесселя (36.296) содержит две постоянные интегрирования

$$\psi_m = c_m J_\nu(\beta \rho) + d_m N_\nu(\beta \rho), \quad c_m, d_m = \text{const},$$

где J_ν и N_ν – функции Бесселя и Неймана порядка ν [319]. Из условия ограниченности волновой функции на оси цилиндра следует, что $d_m = 0$. Постоянные интегрирования c_m находятся из условия нормировки

$$\int_0^a dr r |\psi_m|^2 = 1.$$

Из граничного условия (36.297) следует уравнение на β

$$J_\nu(\beta R/\gamma) = 0, \quad (36.298)$$

которое определяет уровни энергии связанных состояний. Хорошо известно, что при вещественных $\nu > -1$ и a/γ , это уравнение имеет только вещественные нули. Положительные нули образуют бесконечное счетное множество, и все они простые [319]. Отсюда следует неравенство

$$\beta^2 = \frac{2ME}{\hbar^2} - k^2 \geq 0.$$

Пронумеруем положительные нули уравнения (36.298) индексом $n = 1, 2, \dots$ (главное квантовое число): $\beta \rightarrow \beta_n(m, \alpha, \gamma, R)$. Тогда спектр собственных состояний имеет вид

$$E_n = \frac{\hbar^2}{2M} (k^2 + \beta_n^2). \quad (36.299)$$

При больших радиусах ($\beta \rho \gg 1$ и $\beta \rho \gg \nu$) справедлива асимптотика

$$J_\nu(\beta \rho) \approx \sqrt{\frac{2}{\pi \beta \rho}} \cos \left(\beta \rho - \frac{\nu \pi}{2} - \frac{\pi}{4} \right).$$

Отсюда следует явное выражение для спектра

$$\beta_n = \frac{\gamma \pi}{a} \left(n + \frac{|m|}{2\alpha} - \frac{1}{4} \right). \quad (36.300)$$

В отсутствие дефекта $\alpha = 1$, $\gamma = 1$, $\rho = r$, и радиальные функции ψ_m выражаются через функции Бесселя целого порядка $\nu = |m|$. В этом случае спектр энергии собственных состояний зависит только от размеров цилиндра. При наличии клиновой дислокации функции Бесселя в общем случае будут иметь нецелый порядок. При этом спектр уровней энергии собственных состояний приобретает зависимость от угла дефицита клиновой дислокации θ и коэффициента Пуассона σ , характеризующего упругие свойства среды.

Если масса примеси и вакансии определена интегралом (36.16), то для примеси $M > 0$ и собственные значения энергии положительны. Для вакансии $M < 0$ и собственные значения энергии отрицательны. В этом случае спектр энергии не ограничен снизу, что вызывает серьезные проблемы для физической интерпретации.

Рассмотренный пример показывает, как в первом приближении учитывается влияние дефектов в геометрической подходе. Если расчеты в некоторой задаче проведены в упругой среде без дефектов, то для учета влияния дислокаций необходимо заменить плоскую евклидову метрику на нетривиальную метрику, описывающую данное распределение дефектов. Математически эта задача может оказаться сложной, т.к. для определения метрики необходимо решить трехмерные уравнения Эйнштейна. Однако принципиальных трудностей здесь нет: влияние дислокаций сводится к изменению метрики.

36.13 О динамической теории

Проведенное геометрическое построение в теории дефектов можно обратить, и рассматривать гравитационное взаимодействие масс во Вселенной, как взаимодействие дефектов упругого эфира. При этом точечные массы будут соответствовать точечным дефектам (вакансиям и примесям), а космические струны [314, 321] – клиновым дислокациям. При таком взгляде на гравитацию возникает вопрос об упругой калибровке, которой в геометрической теории дефектов придается прямой физический смысл. Если принять точку зрения из теории дефектов, то упругие свойства эфира соответствуют некоторому значению коэффициента Пуассона, который можно измерить экспериментально.

Представляется интересным и важным для приложений включить в рассматриваемый статический подход время для описания движения дефектов в среде. Такая модель в настоящее время отсутствует. С геометрической точки зрения обобщение провести несложно, по крайней мере, в принципе. Достаточно заменить евклидово пространство $\mathbb{R}^3$ на пространство Минковского $\mathbb{R}^{1,3}$ и написать подходящий лагранжиан, квадратичный по тензору кривизны и кручения, что соответствует настоящей модели гравитации с динамическим кручением. Одна из возникающих трудностей заключается в физической интерпретации дополнительных компонент репера и лоренцевой связности, которые содержат временной индекс. Физический смысл временной компоненты репера $e_0^i \rightarrow \partial_0 u^i = v^i$ прост – это скорость точки среды. Такая интерпретация естественна с физической точки зрения, поскольку движение непрерывно распределенных дислокаций означает течение среды. Действительно, жидкость можно представить, как упругую среду с непрерывным распределением подвижных дислокаций. Это значит, что динамическая теория дефектов, основанная на геометрии Римана–Картана, должна включать в себя гидродинамику. В настоящее время неясно, как это происходит. Остается также неясной физическая интерпретация других компонент репера и лоренцевой связности с временным индексом.

Глава 37

Дополнение

37.1 Матрицы

Мы рассматриваем вещественные или комплексные матрицы. Бóльшую часть утверждений и формул настоящего дополнения можно найти, например, в монографиях [17, 23].

Для любого натурального $n \in \mathbb{N}$ справедливо разложение

$$\begin{aligned}(a+b)^n &= \sum_{k=0}^n C_n^k a^{n-k} b^k = \\ &= a^n + na^{n-1}b + \frac{n(n-1)}{2}a^{n-2}b^2 + \dots + nab^{n-1} + b^n,\end{aligned}\quad (37.1)$$

которое называется *биномом Ньютона*, где

$$C_n^k := \frac{n!}{k!(n-k)!} \quad (37.2)$$

– *биномиальные коэффициенты*. В частности, при $a = b = 1$ и $a = -b = 1$ справедливы равенства

$$2^n = \sum_{k=0}^n C_n^k, \quad 0 = \sum_{k=0}^n (-1)^k C_n^k. \quad (37.3)$$

Пусть дана 2×2 матрица

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Если ее определитель отличен от нуля, $\det A := ad - bc \neq 0$, то у нее существует обратная матрица, которая имеет вид

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Для двух 2×2 матриц A и B справедлива формула

$$\det(A+B) = \det A + \operatorname{tr} A \operatorname{tr} B - \operatorname{tr}(AB) + \det B. \quad (37.4)$$

Пусть A – произвольная 3×3 матрица. Тогда справедлива формула

$$A^3 - A^2 \operatorname{tr} A - \frac{1}{2}A (\operatorname{tr} A^2 - (\operatorname{tr} A)^2) - \det A = 0. \quad (37.5)$$

Приведем также формулы, выражающие определитель $n \times n$ матрицы через следы степеней данной матрицы в низших размерностях:

$$n = 2 : \quad \det A = \frac{1}{2!} [(\operatorname{tr} A)^2 - \operatorname{tr} A^2], \quad (37.6)$$

$$n = 3 : \quad \det A = \frac{1}{3!} [(\operatorname{tr} A)^3 - 3 \operatorname{tr} A \operatorname{tr} A^2 + 2 \operatorname{tr} A^3], \quad (37.7)$$

$$n = 4 : \quad \det A = \frac{1}{4!} [(\operatorname{tr} A)^4 - 6(\operatorname{tr} A)^2 \operatorname{tr} A^2 + 3(\operatorname{tr} A^2)^2 + 8 \operatorname{tr} A \operatorname{tr} A^3 - 6 \operatorname{tr} A^4]. \quad (37.8)$$

Эти формулы получаются из представления определителя матрицы A с помощью полностью антисимметричного тензора (37.59) и использования его свойств (37.57). Формулу (37.7) можно также получить, взяв след от равенства (37.5).

Пусть $A = (A_{ab})$ – произвольная квадратная невырожденная $n \times n$ матрица, $\det A \neq 0$. Тогда элементы обратной матрица A^{-1ab} равны

$$A^{-1ab} = \frac{A^{ba}}{\det A},$$

где A^{ba} – алгебраическое дополнение элемента A_{ba} . Если матрица A симметрична, то обратная матрица A^{-1} также симметрична. Для произведения двух невырожденных квадратных матриц справедлива формула $(AB)^{-1} = B^{-1}A^{-1}$. Кроме того $\det A^{-1} \det A = 1$.

Определитель квадратной $n \times n$ матрицы A является полиномом от n^2 элементов матрицы. Поэтому его можно дифференцировать по каждому элементу. Если матрица A невырождена, то справедлива формула

$$A^{-1ba} = \frac{1}{\det A} \frac{\partial \det A}{\partial A_{ab}}. \quad (37.9)$$

Если элементы квадратной матрицы $A_{ab}(x)$ являются дифференцируемыми функциями, то производная от определителя равна

$$\partial_\alpha \det A = \det A A^{-1ab} \partial_\alpha A_{ba} = - \det A \partial_\alpha A^{-1ab} A_{ba}. \quad (37.10)$$

Пусть A и B – две квадратные *обратимые* матрицы размеров $m \times m$ и $n \times n$, соответственно, и пусть C – произвольная матрица размера $n \times m$. Тогда для квадратных $(m+n) \times (m+n)$ матриц справедливо равенство

$$\begin{aligned} \det \begin{pmatrix} A & -C^T \\ C & B \end{pmatrix} &= \det A \det (B + CA^{-1}C^T) = \\ &= \det B \det (A + C^T B^{-1}C). \end{aligned} \quad (37.11)$$

Заметим, что если матрицы A и B антисимметричны, то вся блочная матрица также антисимметрична.

Определитель нижнетреугольной матрицы равен произведению определителей диагональных блоков:

$$\det \begin{pmatrix} A & 0 \\ C & B \end{pmatrix} = \det A \det B. \quad (37.12)$$

Рассмотрим квадратную $m \times m$ матрицу C , которая равна произведению двух прямоугольных матриц

$$C = AB.$$

Размеры матриц A и B равны, соответственно, $m \times n$ и $n \times m$. Тогда справедлива формула Бине–Коши [17], глава I, §2, раздел 5:

$$\det C = \sum_{1 \leq \beta_1 < \beta_2 < \dots < \beta_m \leq n} \det \begin{pmatrix} A_{1\beta_1} & \dots & A_{1\beta_m} \\ \vdots & & \vdots \\ A_{m\beta_1} & \dots & A_{m\beta_m} \end{pmatrix} \det \begin{pmatrix} B_{\beta_1 1} & \dots & B_{\beta_m 1} \\ \vdots & & \vdots \\ B_{\beta_1 m} & \dots & B_{\beta_m m} \end{pmatrix}, \quad (37.13)$$

где сумма берется по всем упорядоченным выборкам m индексов $\beta_1, \dots, \beta_m$ из $1, \dots, n$. Если $m > n$, то сумма считается равной нулю. Согласно формуле Бине–Коши определитель матрицы C равен сумме произведений всех возможных миноров максимального m -того порядка матриц A и B . При $m > n$ $\det C = 0$.

Если матрица A невырождена, то справедливо тождество

$$\begin{aligned} \det(A + B) &= \det A \exp(\operatorname{tr} \ln(1 + A^{-1}B)) = \\ &= \det A \exp \left[\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \operatorname{tr}(A^{-1}B)^n \right]. \end{aligned} \quad (37.14)$$

В частном случае, если $A = 1$, то

$$\det(1 + B) = 1 + \operatorname{tr} B + \frac{1}{2}(\operatorname{tr}^2 B - \operatorname{tr} B^2) + \dots + \det B, \quad (37.15)$$

где точки обозначают сумму определителей матриц, получающихся из B всеми возможными вычеркиваниями одинаковых строк и столбцов.

Если норма матрицы A меньше единицы, то матрица $1 - A$ обратима, и справедлива формула

$$\sum_{k=1}^{\infty} A^k = (1 - A)^{-1}, \quad (37.16)$$

которая является аналогом формулы для суммы членов геометрической прогрессии.

Для произвольных квадратных некоммутирующих матриц A и B справедливо равенство

$$e^{-B} A e^B = A + [A, B] + \frac{1}{2}[[A, B], B] + \dots = \sum_{k=0}^{\infty} \frac{1}{k!} [A, B]_{(k)}, \quad (37.17)$$

где члены ряда определены рекуррентными соотношениями

$$[A, B]_{(0)} = A, \quad [A, B]_{(k+1)} = [[A, B]_{(k)}, B].$$

В частности, если $[B, A] = \lambda A$, то равенство (37.17) принимает вид

$$e^B A e^{-B} = e^{\lambda} A. \quad (37.18)$$

Если коммутатор двух матриц пропорционален единичной матрице,

$$[A, B] = c\mathbb{I},$$

то ряд (37.17) обрывается и справедливы следующие формулы:

$$e^A e^B = e^{[A, B]} e^B e^A = e^{\frac{1}{2}[A, B]} e^{A+B}, \quad (37.19)$$

$$[A, e^B] = [A, B] e^B. \quad (37.20)$$

Последнее равенство можно переписать в эквивалентном виде

$$e^{-B} A e^B = A + [A, B]. \quad (37.21)$$

Равенство (37.19) известно, как *формула Хаусдорфа*.

Теорема 37.1.1 (Троттер). Для двух произвольных вещественных или комплексных квадратных матриц A и B справедлива формула

$$e^{t(A+B)} = \lim_{n \rightarrow \infty} \left(e^{\frac{t}{n}A} e^{\frac{t}{n}B} \right)^n. \quad (37.22)$$

Доказательство. См. [322]. □

Пусть заданы две произвольные симметричные матрицы A и B . Тогда, если равенство

$$A_{ab}B_{cd} - A_{ad}B_{cb} - A_{cd}B_{ab} + A_{cb}B_{ad} = 0 \quad (37.23)$$

выполнено для всех значений индексов, то матрицы A и B пропорциональны.

Приведем без доказательства несколько теорем о свойствах матриц с вещественными элементами, которые часто используются в приложениях. Напомним, что *собственными значениями* λ_i , $i = 1, \dots, n$ квадратной $n \times n$ матрицы A называются корни характеристического (векового) уравнения

$$\det(A - \lambda \mathbb{1}) = 0.$$

Справедливы формулы

$$\det A = \lambda_1 \dots \lambda_n, \quad \operatorname{tr} A^k = \sum_{i=1}^n \lambda_i^k, \quad k = 1, 2, \dots \quad (37.24)$$

У симметричной матрицы все собственные значения вещественны. При этом симметричная матрица положительно определена тогда и только тогда, когда все собственные значения вещественны и положительны $\lambda_i > 0$.

Если симметричная 2×2 матрица, $A = A^T$, имеет положительные собственные значения, то из нее можно извлечь квадратный корень, т.е. представить в виде $A = B^2$. Этот корень не является единственным. Однако, с точностью до знака существует единственный симметричный корень $B = B^T$. Поскольку собственные значения матрицы A положительны, то на диагонали матрицы должны стоять положительные числа. Пусть

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}, \quad \det A = ad - b^2 > 0, \quad a, d > 0.$$

тогда “положительный” квадратный корень имеет вид

$$B = \begin{pmatrix} \sqrt{a} \cos \theta & \sqrt{a} \sin \theta \\ \sqrt{a} \sin \theta & \sqrt{d} \cos \psi \end{pmatrix},$$

где

$$\sin^2 \theta := \frac{b^2}{a(a + d + 2\sqrt{ad - b^2})},$$

$$\sin^2 \psi := \frac{b^2}{d(a + d + 2\sqrt{ad - b^2})},$$

и для углов выбираются положительные корни: $0 \leq \theta, \psi \leq \pi/2$. Отметим, что справедливо тождество

$$\sqrt{a} \sin \theta = \sqrt{d} \sin \psi.$$

Предложение 37.1.1. Пусть A – произвольная вещественная квадратная невырожденная матрица. Тогда матрица $A^T A$ симметрична и положительно определена.

Пусть $\lambda_i(A^T A) \geq 0$ – собственные значения симметричной матрицы $A^T A$. Тогда сингулярными числами $\nu_i(A)$ матрицы A называется набор чисел

$$\nu_i(A) := \sqrt{\lambda_i(A^T A)}.$$

Теорема 37.1.2. О сингулярном разложении матриц. Пусть A произвольная вещественная квадратная матрица и $\nu_i(A)$ – ее сингулярные числа, занумерованные произвольным образом, тогда найдутся ортогональные матрицы R и S , такие что

$$A = R^T \operatorname{diag} \{\nu_1(A), \dots, \nu_n(A)\} S,$$

где $\operatorname{diag} \{\nu_1(A), \dots, \nu_n(A)\}$ – диагональная матрица, у которой на диагонали стоят сингулярные числа.

Доказательство. См., например, [23], теорема 7.3.5. $\square$

Теорема 37.1.3. Пусть A и B произвольные квадратные матрицы с сингулярными числами $\nu_i(A)$ и $\nu_i(B)$, занумерованными в порядке возрастания, тогда справедлива оценка

$$|\operatorname{tr}(AB)| \leq \sum_{i=1}^n \nu_i(A) \nu_i(B).$$

Теорема 37.1.4. Пусть B – произвольная вещественная симметричная положительно определенная матрица. Тогда существует единственная вещественная симметричная положительно определенная матрица A , такая что $A^2 = B$. При этом мы пишем $A = \sqrt{B}$.

Теорема 37.1.5. Полярное разложение вещественных матриц. Любая вещественная матрица A может быть представлена в виде произведения

$$A = RS, \tag{37.25}$$

где R – неотрицательно определенная симметричная матрица и S – ортогональная матрица. Матрица R всегда однозначно определена соотношением

$$R := \sqrt{AA^T}.$$

Если матрица A невырождена, то симметричная матрица R является положительно определенной, а ортогональная матрица S также определена однозначно, $S = R^{-1}A$.

Доказательство. См., например, [23], следствие 7.3.3. $\square$

Пример 37.1.1. Вещественное число $a \in \mathbb{R}$ является 1×1 матрицей и его можно представить в виде $a = |a| \operatorname{sign} a$, где

$$\operatorname{sign} a := \begin{cases} 1 & a > 0, \\ -1, & a < 0, \end{cases}$$

– ортогональная 1×1 матрица. Если $a \neq 0$, то представление существует и единственно. При $a = 0$ модуль числа также равен нулю, $|a| = 0$, а матрица $S = \operatorname{sign} a$ неопределена. $\square$

Полярное разложение вещественных матриц имеет простую геометрическую интерпретацию. Рассмотрим $n \times n$ матрицу A , как линейный оператор, действующий в евклидовом пространстве $\mathbb{R}^n$. Тогда ортогональной матрице соответствует некоторое вращение евклидова пространства вокруг начала координат. Симметричная положительно определенная матрица осуществляет дилатацию евклидова пространства вдоль n взаимно перпендикулярных направлений с различными в общем случае коэффициентами растяжения. Следовательно, полярное разложение матриц означает последовательное выполнение некоторого вращения и некоторой дилатации.

Теорема 37.1.6. Полярное разложение комплексных матриц. *Любая комплексная матрица A может быть представлена в виде произведения*

$$A = RU, \quad (37.26)$$

где R – неотрицательно определенная эрмитова матрица и U – унитарная матрица. Матрица R всегда однозначно определена соотношением

$$R = \sqrt{AA^\dagger}.$$

Если матрица A невырождена, то эрмитова матрица R является положительно определенной, а унитарная матрица U также определена однозначно, $U = R^{-1}A$.

Доказательство. См., например, [23], следствие 7.3.3. □

Пример 37.1.2. Комплексные числа $z = x + iy$ являются комплексными 1×1 матрицами. Для них полярное разложение имеет хорошо известный вид $z = r e^{i\varphi}$. Если $z \neq 0$, то модуль r и аргумент φ комплексного числа определены однозначно. При $z = 0$ модуль комплексного числа также равен нулю, $r = 0$, а аргумент неопределен. □

Замечание. Полярные разложения матриц справедливы также для другого порядка матриц в формулах (37.25) и (37.26): $A = S'R'$ и $A = U'R'$. Вообще говоря, новые матрицы будут отличаться от старых. Сомножители R, S и R, U в полярном разложении матриц перестановочны между собой тогда и только тогда, когда матрицы A, A^T и $A, A^\dagger$ коммутируют, соответственно, для вещественного и комплексного случая. □

Определение. Две квадратные $n \times n$ матрицы A и B называются *подобными*, если существует такая невырожденная матрица S , что

$$B = SAS^{-1}. \quad (37.27)$$

Отображение $A \mapsto SAS^{-1}$ называется *преобразованием подобия*. □

Легко проверить, что преобразование подобия является отношением эквивалентности в множестве квадратных $n \times n$ матриц. Если матрицы A и B подобны, то матрица S , с помощью которой проводится преобразование, определена неоднозначно. Ее можно умножить слева на произвольную невырожденную матрицу, коммутирующую с A .

С помощью преобразования подобия любую матрицу можно преобразовать к некоторому каноническому виду. Одним из таких видов является жорданова форма матрицы.

Определение. Жордановым блоком или жордановой клеткой $J_k(\lambda)$ называется верхнетреугольная $k \times k$ матрица вида

$$J_k(\lambda) := \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \\ 0 & & & & \lambda \end{pmatrix}, \quad (37.28)$$

у которой на главной диагонали стоят одинаковые числа λ , а над ними – единицы. Все остальные элементы матрицы равны нулю. По определению $J_1(\lambda) = \lambda$. Жордановой (вещественной или комплексной) $n \times n$ матрицей называется любая прямая сумма жордановых клеток:

$$J := \begin{pmatrix} J_{n_1}(\lambda_1) & & 0 \\ & J_{n_2}(\lambda_2) & \\ & & \ddots \\ 0 & & & J_{n_k}(\lambda_k) \end{pmatrix}, \quad n_1 + n_2 + \dots + n_k = n, \quad (37.29)$$

где порядки n_i каких то клеток могут совпадать и числа λ_i не обязательно различны. □

Все собственные числа жордановой клетки одинаковы и равны λ , т.е. λ является собственным числом жордановой клетки (37.28) кратности k . λ_i , $i = 1, \dots, n$ – собственные числа матрицы A . Если собственное число $\lambda_i = \lambda_j$ и отличается от всех остальных собственных чисел, то кратность собственного числа λ_i равна $n_i + n_j$.

Теорема 37.1.7. Пусть задана комплексная $n \times n$ матрица A . Тогда существует невырожденная $n \times n$ матрица S такая, что

$$A = SJS^{-1}, \quad (37.30)$$

где J – жорданова матрица (37.29). Матрица A определяет подобную ей жорданову матрицу однозначно с точностью до перестановки жордановых клеток на ее главной диагонали. Собственные значения λ_i не обязательно различны. Если матрица A вещественна и обладает только вещественными собственными значениями, то подобие может быть реализовано с помощью вещественной матрицы S .

Доказательство. См., например, [23], теорема 3.1.11. □

Число k жордановых клеток матрицы A (с учетом возможных повторов) равно максимальному числу ее линейно независимых векторов. При этом для каждой жордановой клетки существует только один независимый собственный вектор. Матрица A диагонализирована тогда и только тогда, когда $k = n$. В частности, если все собственные числа различны, то матрицу A можно привести к диагональному виду с помощью преобразования подобия.

Если порядок всех жордановых клеток равен единице, $n_i = 1$, $i = 1, \dots, k$, то жорданова матрица является диагональной. Если один из жордановых блоков имеет больший порядок, $n_i > 1$, то жорданова матрица не может быть приведена к диагональному виду преобразованием подобия.

Предложение 37.1.2. Любая комплексная матрица A подобна своей транспонированной A^T .

Любая жорданова клетка записывается в виде суммы единичной и нильпотентной матриц:

$$J_k(\lambda) = \lambda \mathbb{1} + N_k, \quad (N_k)^k = 0. \quad (37.31)$$

Любая жорданова матрица записывается в виде суммы диагональной и нильпотентной матриц:

$$J = D + N, \quad N^k = 0, \quad (37.32)$$

где k – порядок наибольшей жордановой клетки. Если задана произвольная матрица и ее жорданова форма, то

$$A = SJS^{-1} = SDS^{-1} + SNS^{-1} := A_D + A_N,$$

где A_D диагонализируемая и A_N нильпотентная матрицы. При этом $[A_D, A_N] = 0$.

Приведем некоторые результаты для эрмитовых и симметричных матриц.

Определение. Комплексная $n \times n$ матрица A называется *эрмитовой*, если выполнено равенство $A^\dagger = A$, где символ $\dagger$ обозначает эрмитово сопряжение: $A^\dagger := (\bar{A})^T$. Если выполнено равенство $A^\dagger = -A$, то матрица A называется *антиэрмитовой*. $\square$

Если матрица A вещественна, то эрмитовы и антиэрмитовы матрицы соответствуют симметричным и антисимметричным матрицам.

Теорема 37.1.8. Все собственные значения эрмитовой $n \times n$ матрицы A вещественны. Матрица A эрмитова тогда и только тогда, когда существует унитарная матрица U такая, что выполнено равенство

$$A = U\Lambda U^{-1}, \quad U \in \mathbb{U}(n), \quad (37.33)$$

где Λ – вещественная диагональная матрица. Матрица Λ составлена из собственных значений матрицы A . Унитарную матрицу U можно выбрать из односвязной подгруппы $\mathbb{SU}(n) \subset \mathbb{U}(n)$.

Все собственные значения вещественной симметричной $n \times n$ матрицы B вещественны. Матрица B симметрична тогда и только тогда, когда существует ортогональная матрица S такая, что выполнено равенство

$$B = S\Lambda S^{-1}, \quad S \in \mathbb{O}(n), \quad (37.34)$$

где Λ – вещественная диагональная матрица. Матрица Λ составлена из собственных значений матрицы B . Ортогональную матрицу S можно выбрать из связной компоненты единицы $\mathbb{SO}(n)$.

Доказательство. См., например, [23], теорема 2.5.6. $\square$

С помощью одной унитарной матрицы иногда удается диагонализировать сразу несколько эрмитовых матриц.

Теорема 37.1.9. Пусть $\mathcal{F}$ – семейство эрмитовых матриц. Унитарная матрица U такая, что матрица UAU^{-1} – диагональна для всех $A \in \mathcal{F}$, существует тогда и только тогда, когда $[A, B] = 0$ для всех $A, B \in \mathcal{F}$.

Доказательство. См., например, [23], теорема 4.1.6. $\square$

В заключение рассмотрим некоторые свойства суперматриц, элементами которых являются коммутирующие и антикоммутирующие объекты. Пусть задана квадратная суперматрица

$$M = \begin{pmatrix} A_m^n & B_m^\nu \\ C_\mu^n & D_\mu^\nu \end{pmatrix}, \quad (37.35)$$

где A и D – квадратные невырожденные матрицы с коммутирующими элементами, а B и C – матрицы с антикоммутирующими элементами, которые могут быть прямоугольными. Суперслед и суперопределитель матрицы M определены соотношениями:

$$\text{str } M := \text{tr } A - \text{tr } D, \quad (37.36)$$

$$\text{sdet } M := \frac{\det A}{\det (D - CA^{-1}B)} = \frac{\det (A - BD^{-1}C)}{\det D}. \quad (37.37)$$

Из определения вытекает, что ряд свойств обычного следа и определителя наследуется суперматрицами:

$$\begin{aligned} \text{str } (M + N) &= \text{str } M + \text{str } N, \\ \text{str } (MN) &= \text{str } (NM), \\ \text{sdet } M^{-1} &= \frac{1}{\text{sdet } M}, \\ \text{sdet } (MN) &= \text{sdet } M \text{ sdet } N, \\ \text{sdet } M &= \exp \text{str } \ln M. \end{aligned} \quad (37.38)$$

37.2 Матрицы Паули

Определение. Три эрмитовых 2×2 матрицы $\sigma = (\sigma_i)$, $i = 1, 2, 3$:

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (37.39)$$

называются *матрицами Паули*. □

Они обладают следующими свойствами:

$$\begin{aligned} \sigma_1 \sigma_2 &= -\sigma_2 \sigma_1 = i\sigma_3, \\ \sigma_2 \sigma_3 &= -\sigma_3 \sigma_2 = i\sigma_1, \\ \sigma_3 \sigma_1 &= -\sigma_1 \sigma_3 = i\sigma_2, \\ \sigma_i^2 &= 1 \quad (\text{суммирования по } i \text{ нет}), \\ \sigma_1 \sigma_2 \sigma_3 &= i. \end{aligned} \quad (37.40)$$

Свойства (37.40) можно записать в виде коммутатора и антикоммутатора матриц Паули:

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ij}^k \sigma_k, \quad \{\sigma_i, \sigma_j\} = 2\delta_{ij} \quad (37.41)$$

или

$$\sigma_i \sigma_j = \delta_{ij} + i\varepsilon_{ij}^k \sigma_k, \quad (37.42)$$

где ε_{ijk} — полностью антисимметричный тензор третьего ранга и подъем индексов осуществляется с помощью евклидовой метрики δ_{ij} . Отметим правило комплексного сопряжения матриц Паули:

$$\sigma_1^* = \sigma_1, \quad \sigma_2^* = -\sigma_2, \quad \sigma_3^* = \sigma_3.$$

Матрицы Паули встречаются в различных контекстах. Во-первых, они являются генераторами группы двумерных унитарных матриц $\mathrm{SU}(2)$ (первое соотношение в (37.41)). Во-вторых, второе из соотношений (37.41) означает, что матрицы Паули являются образующими алгебры Клиффорда $\mathfrak{cl}(3, 0, \mathbb{F})$, где $\mathbb{F} = \mathbb{R}$ или $\mathbb{C}$. В-третьих, матрицы Паули, умноженные на $-i$, можно рассматривать, как представление образующих алгебры кватернионов (37.52). Кроме того, матрицы Паули вместе с единичной 2×2 матрицей $\mathbb{1}$ образуют базис линейного пространства всех эрмитовых 2×2 матриц.

Для следов произведения матриц Паули справедливы тождества:

$$\mathrm{tr} \sigma_i = 0, \quad \mathrm{tr} (\sigma_i \sigma_j) = 2\delta_{ij}, \quad \mathrm{tr} (\sigma_i \sigma_j \sigma_k) = 2i\varepsilon_{ijk}. \quad (37.43)$$

Справедливы также равенства

$$\det \sigma_i = -1, \quad i = 1, 2, 3.$$

Рассмотрим вектор $\mathbf{x} = \{x^i\} \in \mathbb{R}^3$ в трехмерном евклидовом пространстве с обычным скалярным произведением. Тогда справедлива формула

$$e^{i(\mathbf{x}, \boldsymbol{\sigma})} = \cos |\mathbf{x}| + i \frac{(\mathbf{x}, \boldsymbol{\sigma})}{|\mathbf{x}|} \sin |\mathbf{x}|, \quad (37.44)$$

где

$$(\mathbf{x}, \boldsymbol{\sigma}) := x^1 \sigma_1 + x^2 \sigma_2 + x^3 \sigma_3, \quad |\mathbf{x}| := \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}.$$

Свертка равенства (37.42) с произвольными векторами $\mathbf{x}$ и $\mathbf{y}$ приводит к тождеству

$$(\mathbf{x}, \boldsymbol{\sigma})(\mathbf{y}, \boldsymbol{\sigma}) = (\mathbf{x}, \mathbf{y}) + i([\mathbf{x}, \mathbf{y}], \boldsymbol{\sigma}), \quad (37.45)$$

где $[\mathbf{x}, \mathbf{y}]^i := x^j y^k \varepsilon_{jk}^i$ — векторное произведение векторов. В общем случае из свойства (37.42) следует, что всякая функция от матриц Паули сводится к линейной функции.

При проведении вычислений с двумерными спинорами удобно использовать условие полноты матриц Паули:

$$\sigma_{\alpha}^i \sigma_{i\gamma}^{\beta} = 2\delta_{\alpha}^{\beta} \delta_{\gamma}^{\delta} - \delta_{\alpha}^{\beta} \delta_{\gamma}^{\delta}, \quad (37.46)$$

где $\sigma^i := \delta^{ij} \sigma_j$, и греческие индексы являются матричными. Если матрицы Паули дополнить единичной матрицей, то условие полноты запишется в следующем виде

$$\mathbb{1}_{\alpha}^{\beta} \mathbb{1}_{\gamma}^{\delta} + \sigma_{\alpha}^i \sigma_{i\gamma}^{\beta} = 2\delta_{\alpha}^{\beta} \delta_{\gamma}^{\delta}. \quad (37.47)$$

Свертка формулы (37.46) по индексам β и γ приводит к равенству

$$\sigma^i \sigma_i = 3,$$

где матричные индексы опущены. Это же равенство возникает после свертки формулы (37.42) по индексам i и j с евклидовой метрикой δ^{ij} .

С групповой точки зрения матрицы Паули являются генераторами группы $\mathrm{SU}(2)$ или $\mathrm{SO}(3)$. “Скалярное произведение” $(\mathbf{x}, \boldsymbol{\sigma})$ и экспонента $e^{i(\mathbf{x}, \boldsymbol{\sigma})}$ являются, соответственно, элементами алгебры Ли $\mathfrak{su}(2)$ и группы Ли $\mathrm{SU}(2)$. Отображение (37.44) представляет собой экспоненциальное отображение алгебры Ли в группу Ли.

37.3 Кватернионы

Кватернионы играют большую роль в математической физике. Особенно в теории алгебр Клиффорда (см. главу 26).

Определение. Рассмотрим четырехмерное векторное пространство $\mathbb{H}$ над полем вещественных чисел, которое содержит поле вещественных чисел как подпространство. Обозначим базис этого пространства символами $1, i, j, k$, где 1 – единица из поля вещественных чисел. Тогда произвольный элемент $q \in \mathbb{H}$ представим в виде

$$q = a + bi + cj + dk, \quad a, b, c, d \in \mathbb{R}. \quad (37.48)$$

Введем билинейное умножение в $\mathbb{H}$, положив следующее правило умножения базисных векторов:

$$\begin{aligned} ij &= -ji = k, \\ jk &= -kj = i, \\ ki &= -ik = j, \\ i^2 &= j^2 = k^2 = -1. \end{aligned} \quad (37.49)$$

Элементы вида (37.48) с операциями сложения, умножения на действительные числа и умножения (37.49) называются *кватернионами*. $\square$

С помощью прямых вычислений можно проверить, что кватернионы с так определенным умножением образуют ассоциативную, но некоммутативную алгебру с единицей $\mathbb{H}$ над полем вещественных чисел. Напомним, что ассоциативная алгебра с единицей называется унитарной.

Алгебру кватернионов можно представить в виде комплексных матриц

$$A(q) = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} \in \mathbf{mat}(2, \mathbb{C}), \quad (37.50)$$

где i – мнимая единица. Прямые вычисления показывают, что отображение $q \mapsto A(q)$ является гомоморфизмом, т.е.

$$A(q_1 q_2) = A(q_1) A(q_2).$$

Для доказательства этого утверждения достаточно проверить его для $q = i, j, k$. В частности, при $q = i, j, k$ имеем равенства:

$$A(i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad A(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad A(k) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \quad (37.51)$$

Эти матрицы с точностью до мнимой единицы $-i \in \mathbb{C}$ совпадают с матрицами Паули (37.39):

$$\sigma_1 = -iA(k), \quad \sigma_2 = -iA(j), \quad \sigma_3 = -iA(i). \quad (37.52)$$

Введем операцию сопряжения в $\mathbb{H}$, т.е. поставим каждому кватерниону вида (37.48) его сопряженный элемент

$$\bar{q} := a - bi - cj - dk. \quad (37.53)$$

Это отображение $q \mapsto \bar{q}$ является антиавтоморфизмом алгебры кватернионов, т.е.

$$\overline{q_1 + q_2} = \bar{q}_1 + \bar{q}_2, \quad \overline{q_1 q_2} = \bar{q}_2 \bar{q}_1.$$

Число a называется вещественной частью $\operatorname{re} q := a$, а остаток $\operatorname{im} q := bi + cj + dk$ – мнимой частью кватерниона. Легко проверяется свойство

$$\operatorname{re}(q_1 q_2) = \operatorname{re}(q_2 q_1).$$

Поскольку выполнены также равенства

$$\bar{\bar{q}} = q \quad \text{и} \quad a\bar{q} = \bar{q}a, \quad a \in \mathbb{R},$$

то сопряжение в вещественной алгебре кватернионов является инволюцией. Кроме того справедливы формулы:

$$\operatorname{re} q = \frac{1}{2}(q + \bar{q}), \quad \operatorname{im} q = \frac{1}{2}(q - \bar{q}),$$

как и в поле комплексных чисел.

В алгебре кватернионов можно ввести симметричное скалярное произведение

$$(q_1, q_2) := \operatorname{re} q_1 \bar{q}_2 = a_1 a_2 + b_1 b_2 + c_1 c_2 + d_1 d_2, \quad (37.54)$$

которое совпадает с обычным евклидовым произведением в $\mathbb{R}^4$. Поскольку каждый кватернион представим в виде $q = \operatorname{re} q + \operatorname{im} q$, то алгебра кватернионов распадается в прямую сумму ортогональных подпространств $\mathbb{H} = \mathbb{R} \oplus \mathbb{R}^3$, где $\mathbb{R}$ – прямая, состоящая из вещественных кватернионов, а $\mathbb{R}^3$ – ортогональное трехмерное евклидово пространство мнимых кватернионов.

Предложение 37.3.1. *Кватернион q является вещественным тогда и только тогда, когда он коммутирует со всеми кватернионами из $\mathbb{H}$. То есть первое слагаемое $\mathbb{R}$ в прямой сумме $\mathbb{R} \oplus \mathbb{R}^3$ является центром алгебры кватернионов.*

Доказательство. Прямая проверка. □

Определим норму кватерниона

$$\|q\|^2 := q\bar{q} = \bar{q}q = a^2 + b^2 + c^2 + d^2. \quad (37.55)$$

Нетрудно проверить, что

$$\|q\|^2 = \det A(q).$$

Отсюда следует, что $\|q_1 q_2\|^2 = \|q_1\|^2 \|q_2\|^2$.

Каждый отличный от нуля кватернион обладает единственным обратным элементом

$$qq^{-1} = q^{-1}q = 1,$$

где

$$q^{-1} = \frac{\bar{q}}{\|q\|^2}.$$

Это означает, что множество кватернионов относительно операции сложения и умножения (37.49) (без умножения на действительные числа) образует тело, при этом умножение некоммутативно.

Предложение 37.3.2. Пусть $I \in \mathbb{H}$ – фиксированный чисто мнимый кватернион с единичной нормой, $\|I\| = 1$. Тогда множество кватернионов вида $a1 + bI \in \mathbb{H}$ для всех $a, b \in \mathbb{R}$ образует двумерную подалгебру в $\mathbb{H}$, которая изоморфна алгебре комплексных чисел $\mathbb{C}$. В частности, эта подалгебра коммутативна.

Доказательство. Поскольку кватернион I чисто мнимый и обладает единичной нормой, то справедливы равенства:

$$\bar{I} = -I, \quad I^2 = -I\bar{I} = -1.$$

Поэтому кватернион $a + bI \in \mathbb{H}$ можно отождествить с комплексным числом $a + ib \in \mathbb{C}$. Согласованность умножения очевидна, т.к. $1 \cdot I = I$. $\square$

Алгебра кватернионов $\mathbb{H}$ допускает естественное отождествление с двумерным комплексным пространством $\mathbb{C}^2$. Используя определение умножения в алгебре кватернионов $k = -ji$ (37.49), запишем кватернион в виде

$$q = (a + ib) + j(c - id) = z_1 + j\bar{z}_2,$$

где линейные комбинации $z_1 := a + ib$ и $z_2 := c + id$ можно отождествить с парой комплексных чисел. При этом 1 и j рассматриваются, как базис в $\mathbb{C}^2$.

Рассмотрим мнимый кватернион, не содержащий вещественной части,

$$r = bi + cj + dk, \quad \|r\| = b^2 + c^2 + d^2.$$

Его можно представить, как вектор в трехмерном евклидовом пространстве $r \in \mathbb{R}^3$. Пусть $q \in \mathbb{H}$ произвольный обратимый (в смысле кватернионного умножения) кватернион. Тогда кватернион qrq^{-1} снова является мнимым кватернионом с той же нормой:

$$qrq^{-1} \in \mathbb{R}^3, \quad \|qrq^{-1}\| = \|r\|.$$

Это значит, что каждому обратимому кватерниону, т.е. отличному от нуля, ставится в соответствие некоторое вращение. При этом двум кватернионам q и $-q$, отличающимся знаком, соответствует одно и то же вращение. Можно проверить, что любое вращение можно реализовать с помощью кватерниона единичной нормы. Поскольку кватернионы единичной нормы параметризуют трехмерную сферу единичного радиуса, то существует гомеоморфизм

$$\mathrm{SO}(3) \approx \frac{\mathbb{S}^3}{\mathbb{Z}_2}.$$

Рассмотрим отображение кватернионов $q \in \mathbb{H}$, задаваемое двумя кватернионами $a, b \in \mathbb{H}$ единичной нормы:

$$q \mapsto aqb^{-1}, \quad \|a\| = \|b\| = 1.$$

Это отображение задает вращение четырехмерного евклидова пространства $\mathbb{H} = \mathbb{R}^4$, поскольку сохраняет норму q . Можно проверить, что любое вращение определяет пару a и b с точностью до знака, т.к. кватернионы a, b и $-a, -b$ определяют одно и то же вращение. Это значит, что существует гомеоморфизм

$$\mathrm{SO}(4) \approx \frac{\mathbb{S}^3 \times \mathbb{S}^3}{\mathbb{Z}_2},$$

где $\mathbb{Z}_2 := \{(1, 1), (-1, -1)\}$.

Множество обратимых кватернионов образует группу по умножению, которую мы обозначим $\mathbb{H}^\times$.

Предложение 37.3.3. Множество кватернионов с единичной нормой, $\|q\| = 1$ образует подгруппу $\mathbb{SU}(2) \subset \mathbb{H}^\times$, которая, как многообразие, совпадает с трехмерной сферой $\mathbb{S}^3 \hookrightarrow \mathbb{R}^4$.

Доказательство. Групповые аксиомы легко проверяются. Изоморфизм следует из точного представления (37.50), поскольку для кватернионов единичной нормы справедливы равенства:

$$A^\dagger A = AA^\dagger = \begin{pmatrix} a^2 + b^2 + c^2 + d^2 & 0 \\ 0 & a^2 + b^2 + c^2 + d^2 \end{pmatrix} = \mathbb{1}.$$

Вложение является прямым следствием определения нормы (37.55). $\square$

Предложение 37.3.4. Множество кватернионов, удовлетворяющих условию $q^2 = -1$, образует двумерную сферу $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$.

Доказательство. Поскольку

$$q^2 = (\operatorname{re} q)^2 + 2 \operatorname{re} q \operatorname{im} q + (\operatorname{im} q)^2 = -1,$$

то $\operatorname{re} q = 0$, т.е. кватернион является чисто мнимым. Тогда из условия $q^2 = -1$ следует, что $\|q\| = 1$, что влечет за собой вложение $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. $\square$

37.4 Полностью антисимметричные тензоры

Полностью антисимметричный тензор ранга n в n -мерном евклидовом пространстве единственен с точностью до умножения на произвольную ненулевую постоянную и имеет всего одну независимую компоненту. Его всегда можно представить в виде константы, умноженной на тензор, все ненулевые компоненты которого равны по модулю единице. Этот тензор часто используется в приложениях и имеет специальное обозначение

$$\begin{aligned} \varepsilon^{a_1 \dots a_n} &= \varepsilon^{[a_1 \dots a_n]}, & \varepsilon^{12 \dots n} &= 1, \\ \varepsilon_{a_1 \dots a_n} &= \varepsilon^{a_1 \dots a_n}, & \varepsilon_{12 \dots n} &= 1. \end{aligned} \quad (37.56)$$

Здесь подъем индексов произведен с помощью евклидовой метрики. Произведения и свертки двух таких тензоров выражаются через символы Кронекера:

$$\begin{aligned} \varepsilon^{a_1 \dots a_n} \varepsilon_{b_1 \dots b_n} &= n! \delta_{b_1 \dots b_n}^{[a_1 \dots a_n]} = \det \begin{pmatrix} \delta_{b_1}^{a_1} & \delta_{b_1}^{a_2} & \dots & \delta_{b_1}^{a_n} \\ \delta_{b_2}^{a_1} & \delta_{b_2}^{a_2} & \dots & \delta_{b_2}^{a_n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{b_n}^{a_1} & \delta_{b_n}^{a_2} & \dots & \delta_{b_n}^{a_n} \end{pmatrix} \\ \varepsilon^{b_1 a_2 \dots a_n} \varepsilon_{b_1 \dots b_n} &= (n-1)! \delta_{b_2 \dots b_n}^{[a_2 \dots a_n]} = \frac{n!}{C_n^1} \delta_{b_2 \dots b_n}^{[a_2 \dots a_n]} \\ \varepsilon^{b_1 b_2 a_3 \dots a_n} \varepsilon_{b_1 \dots b_n} &= 2(n-2)! \delta_{b_3 \dots b_n}^{[a_3 \dots a_n]} = \frac{n!}{C_n^2} \delta_{b_3 \dots b_n}^{[a_3 \dots a_n]} \\ \varepsilon^{b_1 \dots b_k a_{k+1} \dots a_n} \varepsilon_{b_1 \dots b_n} &= k!(n-k)! \delta_{b_{k+1} \dots b_n}^{[a_{k+1} \dots a_n]} = \frac{n!}{C_n^k} \delta_{b_{k+1} \dots b_n}^{[a_{k+1} \dots a_n]} \\ \varepsilon^{b_1 \dots b_{n-1} a_n} \varepsilon_{b_1 \dots b_n} &= (n-1)! \delta_{b_n}^{a_n} = \frac{n!}{C_n^{n-1}} \delta_{b_n}^{a_n} \\ \varepsilon^{b_1 \dots b_n} \varepsilon_{b_1 \dots b_n} &= n! \end{aligned} \quad (37.57)$$

Полностью антисимметричный тензор максимального ранга в пространстве с метрикой произвольной сигнатуры определяется соотношениями

$$\varepsilon_{12\dots n} = 1, \quad \varepsilon^{12\dots n} = \text{sgn}, \quad (37.58)$$

где множитель $\text{sgn} := \text{sign}(\det g_{\alpha\beta})$ равен знаку определителя метрики. Этот множитель возникает при подъеме индексов. Для таких пространств правые части (37.57) необходимо умножить на множитель sgn .

Определитель произвольной квадратной матрицы A в евклидовом пространстве можно записать в виде

$$\begin{aligned} \det(A_a^b) &= \varepsilon_{b_1 b_2 \dots b_n} A_1^{b_1} A_2^{b_2} \dots A_n^{b_n} = \varepsilon^{a_1 a_2 \dots a_n} A_{a_1}^1 A_{a_2}^2 \dots A_{a_n}^n \\ &= \frac{1}{n!} \varepsilon^{a_1 a_2 \dots a_n} \varepsilon_{b_1 b_2 \dots b_n} A_{a_1}^{b_1} A_{a_2}^{b_2} \dots A_{a_n}^{b_n}. \end{aligned} \quad (37.59)$$

Полностью антисимметричный тензор в n -мерном евклидовом пространстве инвариантен относительно $\mathbb{SO}(n)$ вращений с матрицей $S_a^{a'} \in \mathbb{SO}(n)$:

$$\varepsilon^{a'_1 \dots a'_n} = \varepsilon^{a_1 \dots a_n} S_{a_1}^{a'_1} \dots S_{a_n}^{a'_n} = \det(S_a^{a'}) \varepsilon^{a'_1 \dots a'_n}. \quad (37.60)$$

Полностью антисимметричный тензор в голономном базисе на многообразии имеет следующие компоненты:

$$\varepsilon^{\alpha_1 \dots \alpha_n} = e^{\alpha_1}_{a_1} \dots e^{\alpha_n}_{a_n} \varepsilon^{a_1 \dots a_n}, \quad (37.61)$$

и в геометрии Римана–Картана является ковариантно постоянным:

$$\nabla_\beta \varepsilon^{\alpha_1 \dots \alpha_n} = 0. \quad (37.62)$$

По определению, полностью антисимметричный тензор в пространстве Минковского четного числа измерений меняет знак при пространственном отражении $(x^0, x^i) \mapsto (x^0, -x^i)$, т.е. является псевдотензором. Приставку “псевдо” мы, как правило, опускаем.

Полностью антисимметричная тензорная плотность $\hat{\varepsilon}^{\alpha_1 \dots \alpha_n}$ веса -1 определяется соотношениями

$$\begin{aligned} \hat{\varepsilon}^{\alpha_1 \dots \alpha_n} &:= \sqrt{|g|} \varepsilon^{\alpha_1 \dots \alpha_n}, \\ \hat{\varepsilon}_{\alpha_1 \dots \alpha_n} &:= \frac{1}{\sqrt{|g|}} \varepsilon_{\alpha_1 \dots \alpha_n}. \end{aligned} \quad (37.63)$$

Ее ненулевые компоненты равны по модулю единице в любой системе координат,

$$\hat{\varepsilon}^{\alpha_1 \dots \alpha_n} = \begin{cases} +1 & \text{— четная перестановка индексов,} \\ -1 & \text{— нечетная перестановка индексов,} \\ 0 & \text{— совпадение любой пары индексов.} \end{cases} \quad (37.64)$$

Отсюда следует, что все частные производные антисимметричной тензорной плотности равны нулю:

$$\partial_\beta \hat{\varepsilon}^{\alpha_1 \dots \alpha_n} = 0. \quad (37.65)$$

Приведем явные формулы для трехмерного евклидова пространства $\mathbb{R}^3$, которые часто используются в физических приложениях,

$$\varepsilon_{123} = 1, \quad \varepsilon^{123} = 1. \quad (37.66)$$

Свертки:

$$\begin{aligned}\varepsilon^{lmn}\varepsilon_{ijk} &= \delta_{ijk}^{lmn} + \delta_{ijk}^{mnl} + \delta_{ijk}^{nlm} - \delta_{ijk}^{lnm} - \delta_{ijk}^{mln} - \delta_{ijk}^{nml} = \det \begin{pmatrix} \delta_i^l & \delta_i^m & \delta_i^n \\ \delta_j^l & \delta_j^m & \delta_j^n \\ \delta_k^l & \delta_k^m & \delta_k^n \end{pmatrix} \\ \varepsilon^{klm}\varepsilon_{ijm} &= \delta_{ij}^{kl} - \delta_{ij}^{lk}, \\ \varepsilon^{jkl}\varepsilon_{ikl} &= 2\delta_i^j, \\ \varepsilon^{ijk}\varepsilon_{ijk} &= 6.\end{aligned}\tag{37.67}$$

Формулы (37.67) справедливы и при лоренцевой сигнатуре метрики $(+ - -)$, т.к. $\varepsilon_{012} = 1$ и $\varepsilon^{012} = 1$.

Полностью антисимметричный тензор второго ранга в евклидовом пространстве имеет две ненулевые компоненты

$$\varepsilon_{12} = 1, \quad \varepsilon^{12} = 1\tag{37.68}$$

или

$$\varepsilon_{ab} = \varepsilon^{ab} = \varepsilon^a{}_b = \varepsilon_a{}^b = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.\tag{37.69}$$

Свертки:

$$\begin{aligned}\varepsilon^{ab}\varepsilon_{cd} &= \delta_{cd}^{ab} - \delta_{cd}^{ba} = \det \begin{pmatrix} \delta_c^a & \delta_c^b \\ \delta_d^a & \delta_d^b \end{pmatrix} \\ \varepsilon^{ab}\varepsilon_{ad} &= \delta_d^b, \\ \varepsilon^{ab}\varepsilon_{ab} &= 2.\end{aligned}\tag{37.70}$$

В пространстве Минковского нечетного числа измерений формулы для сверток полностью антисимметричных тензоров те же, что и в евклидовом пространстве. В пространстве Минковского четного числа измерений полностью антисимметричные тензоры с верхними и нижними индексами отличаются знаком. Определим антисимметричный тензор четвертого ранга в пространстве Минковского следующим образом

$$\varepsilon_{0123} = 1, \quad \varepsilon^{0123} = -1.\tag{37.71}$$

Свертки:

$$\begin{aligned}\varepsilon^{abcd}\varepsilon_{efgh} &= -\delta_{efgh}^{abcd} - \delta_{efgh}^{acdb} - \delta_{efgh}^{adbc} + \delta_{efgh}^{acbd} + \delta_{efgh}^{abdc} + \delta_{efgh}^{adcb} \\ &\quad + \delta_{efgh}^{bcd a} + \delta_{efgh}^{bdac} + \delta_{efgh}^{bacd} - \delta_{efgh}^{bdca} - \delta_{efgh}^{bcad} - \delta_{efgh}^{badc} \\ &\quad - \delta_{efgh}^{cdab} - \delta_{efgh}^{cabd} - \delta_{efgh}^{cbda} + \delta_{efgh}^{cadb} + \delta_{efgh}^{cdba} + \delta_{efgh}^{cbad} \\ &\quad + \delta_{efgh}^{dabc} + \delta_{efgh}^{dbca} + \delta_{efgh}^{dcab} - \delta_{efgh}^{dbac} - \delta_{efgh}^{dacb} - \delta_{efgh}^{dcba} \\ \varepsilon^{abcd}\varepsilon_{afgh} &= -\delta_{fgh}^{bcd} - \delta_{fgh}^{cdb} - \delta_{fgh}^{dbc} + \delta_{fgh}^{cbd} + \delta_{fgh}^{bdc} + \delta_{fgh}^{dcb} \\ \varepsilon^{abcd}\varepsilon_{abgh} &= -2(\delta_{gh}^{cd} - \delta_{gh}^{dc}), \\ \varepsilon^{abcd}\varepsilon_{abch} &= -6\delta_h^d, \\ \varepsilon^{abcd}\varepsilon_{abcd} &= -24.\end{aligned}\tag{37.72}$$

Полностью антисимметричные тензоры в трехмерном евклидовом пространстве и пространстве Минковского связаны следующими соотношениями

$$\varepsilon_{ijk} = \varepsilon_{0ijk}, \quad \varepsilon^{ijk} = -\varepsilon^{0ijk}.\tag{37.73}$$

Полностью антисимметричный тензор второго ранга в пространстве Минковского $\mathbb{R}^{1,1}$ имеет две ненулевые компоненты:

$$\varepsilon_{01} = 1, \quad \varepsilon^{01} = -1 \quad (37.74)$$

или

$$\varepsilon_{ab} = -\varepsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \varepsilon^a_b = -\varepsilon_a^b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (37.75)$$

Свертки:

$$\begin{aligned} \varepsilon^{ab}\varepsilon_{cd} &= -\delta_{cd}^{ab} + \delta_{cd}^{ba}, \\ \varepsilon^{ab}\varepsilon_{ad} &= -\delta_d^b, \\ \varepsilon^{ab}\varepsilon_{ab} &= -2. \end{aligned} \quad (37.76)$$

В двумерном пространстве Минковского $\mathbb{R}^{1,1}$ можно ввести проекционные операторы

$$\Pi_a^{\pm b} := \frac{1}{2}(\delta_a^b \pm \varepsilon_a^b).$$

Отметим, что аналогичная конструкция в евклидовом пространстве не приводит к проекционным операторам.

Пусть в римановом пространстве $(\mathbb{M}, g)$, $\dim \mathbb{M} = n$, задано тензорное поле с компонентами $K_{\alpha\beta}$ типа $(0, 2)$ такое, что $\det K_{\alpha\beta} \neq 0$. Тогда существует обратное тензорное поле $K^{\alpha\beta}$ типа $(2, 0)$,

$$K^{\alpha\beta}K_{\beta\gamma} = \delta_\gamma^\alpha,$$

и его можно представить в виде

$$K^{\alpha\beta} = \frac{1}{|\det K|(n-1)!} \hat{\varepsilon}^{\alpha\gamma_2 \dots \gamma_n} \hat{\varepsilon}^{\beta\delta_2 \dots \delta_n} K_{\delta_2\gamma_2} \dots K_{\delta_n\gamma_n}, \quad (37.77)$$

Если метрика имеет лоренцеву сигнатуру, то в пространстве четного числа измерений в правой части появляется знак минус.

37.5 δ -функция

В простейшем случае в качестве основных (пробных) функций рассматривается пространство гладких финитных (с компактным носителем) функций $\varphi(x) \in \mathcal{D}(\mathbb{R}^n)$ в евклидовом пространстве $x \in \mathbb{R}^n$, $n = 1, 2, \dots$. Множество обобщенных функций – это множество линейных непрерывных функционалов на основных функциях из $\mathcal{D}(\mathbb{R}^n)$. Пространство обобщенных функций является сопряженным к пространству основных функций и обозначается $\mathcal{D}'(\mathbb{R}^n)$ (см., например, [3]).

Мы используем следующие обозначения для одномерной δ -функции и ее производных:

$$\begin{aligned} (\delta, \varphi) &:= \int_{-\infty}^{\infty} dx \delta(x) \varphi(x) = \varphi(0), \\ (\delta', \varphi) &= \int_{-\infty}^{\infty} dx \delta'(x) \varphi(x) = -\varphi'(0), \\ (\delta'', \varphi) &= \int_{-\infty}^{\infty} dx \delta''(x) \varphi(x) = \varphi''(0), \\ (\delta^{(m)}, \varphi) &= \int_{-\infty}^{\infty} dx \delta^{(m)}(x) \varphi(x) = (-1)^m \varphi^{(m)}(0). \end{aligned} \quad (37.78)$$

где производные по $x \in \mathbb{R}$ обозначены штрихами. δ -функция в общем случае зависит от двух аргументов: $\delta = \delta(x, x')$, при этом $\delta(x) := \delta(x, 0)$. Производные δ -функции можно представить в следующем виде

$$\delta'(x', x) = -\delta'(x, x') = \frac{\partial}{\partial x'} \delta(x', x) = -\frac{\partial}{\partial x} \delta(x', x), \quad (37.79)$$

$$\delta''(x', x) = \delta''(x, x') = \frac{\partial^2}{\partial x'^2} \delta(x', x) = \frac{\partial^2}{\partial x^2} \delta(x', x) = -\frac{\partial^2}{\partial x' \partial x} \delta(x', x). \quad (37.80)$$

Здесь и далее равенства выполняются в слабом смысле после интегрирования по x и x' с основными функциями. δ -функция равна производной

$$\delta(x) = \theta'(x) \quad (37.81)$$

от функции Хевисайда

$$\theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0, \end{cases} \quad (37.82)$$

которую также называют *ступенькой*. Равенство (37.81) можно переписать в виде

$$\delta(x) = \frac{1}{2} \left[\frac{\partial}{\partial x} \theta(x) - \frac{\partial}{\partial x} \theta(-x) \right]. \quad (37.83)$$

Для функции Хевисайда справедливо интегральное представление

$$\theta(\pm x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \frac{e^{ikx}}{\epsilon \pm ik}, \quad \epsilon \in \mathbb{R}, \quad \epsilon \rightarrow 0. \quad (37.84)$$

В пространстве обобщенных функций $f, g \in \mathcal{D}'(\mathbb{R}^n)$ умножение отсутствует. Тем не менее некоторые обобщенные функции умножать можно. Например, можно умножать обобщенные функции, которые задаются гладкими функциями в $\mathbb{R}^n$. Кроме этого, δ -функцию и ее производные можно умножать на достаточно гладкие функции.

Пусть $f, g \in \mathcal{C}^1(\mathbb{R})$, тогда нетрудно проверить следующие соотношения:

$$\begin{aligned} f'\delta &= f\delta, \\ (f - f')\delta' &= \partial f \delta, \\ fg'\delta' &= fg\delta' - f\partial g \delta, \\ f'g\delta' &= f'g'\delta' + f\partial g \delta, \\ (fg' + f'g)\delta' &= (fg + f'g')\delta', \\ (fg' - f'g)\delta' &= (\partial fg - f\partial g)\delta, \\ (f - f')\delta'' &= (\partial f + \partial' f')\delta', \\ fg'\delta'' &= fg\delta'' - 2f\partial g \delta' + f\partial^2 g \delta, \\ f'g\delta'' &= f'g'\delta'' + 2f'\partial' g'\delta' + f\partial^2 g \delta, \\ (fg' + f'g)\delta'' &= (fg + f'g')\delta'' - 2\partial f \partial g \delta, \\ (fg' - f'g)\delta'' &= (\partial fg + \partial' f'g' - f\partial g - f'\partial' g')\delta'. \end{aligned} \quad (37.85)$$

В приведенных формулах δ -функция и ее производные δ' и δ'' зависят от аргументов (x', x) . Для краткости, штрих у функций f и g означает, что они рассматриваются в точке x' , а частные производные обозначены через $\partial := \partial/\partial x$ и $\partial' := \partial/\partial x'$.

Обобщенная функция $\mathcal{P}\frac{1}{x}$ определяется соотношением

$$\left(\mathcal{P}\frac{1}{x}, \varphi\right) := \text{vp} \int dx \frac{\varphi(x)}{x}, \quad (37.86)$$

где vp обозначает главное значение интеграла. Справедливы формулы:

$$\mathcal{P}\frac{1}{x} = \frac{d}{dx} \ln|x|, \quad (37.87)$$

$$x\mathcal{P}\frac{1}{x} = 1. \quad (37.88)$$

Формулы Соболевского:

$$\frac{1}{x \pm i0} = \mp i\pi\delta(x) + \mathcal{P}\frac{1}{x}. \quad (37.89)$$

В квантовой механике часто используются следующие обобщенные функции

$$\delta^+ := \frac{\delta}{2} + \frac{1}{2\pi i} \mathcal{P}\frac{1}{x}, \quad (37.90)$$

$$\delta^- := \frac{\delta}{2} - \frac{1}{2\pi i} \mathcal{P}\frac{1}{x}, \quad (37.91)$$

$$\delta = \delta^+ + \delta^-. \quad (37.92)$$

В n -мерном евклидовом пространстве $\mathbb{R}^n$ под δ -функцией понимается произведение δ -функций по каждой из декартовых координат:

$$\delta(x) := \delta(x^1)\delta(x^2)\dots\delta(x^n) = \prod_{\alpha} \delta(x^{\alpha}). \quad (37.93)$$

Тогда

$$\int dx \delta(x)\varphi(x) = \varphi(0).$$

Рассмотрим преобразование координат $x^{\alpha} \rightarrow y^{\alpha}(x)$

$$\int dx \delta(x)\varphi(x) = \int dy J_0^{-1} \delta(x(y))\varphi(x(y)), \quad (37.94)$$

где обратный якобиан преобразования координат достаточно рассмотреть в одной точке,

$$J_0^{-1} := \det \left(\frac{\partial x^{\alpha}}{\partial y^{\beta}} \right) \Big|_{x=0}.$$

Пусть $y(0) = y_0$. Тогда интеграл (37.94) можно переписать в виде

$$\int dy \tilde{\delta}(y, y_0) f(y) = f(y_0), \quad \tilde{\delta}(y, y_0) = J_0^{-1} \delta(x).$$

То есть δ -функция является скалярной плотностью веса -1 , как и элемент объема $\sqrt{|g|}$. В частном случае формула невырожденной линейной замены переменных интегрирования $x := Ay + b$ в $\mathbb{R}^n$, где $A \in \text{GL}(n, \mathbb{R})$ и $b \in \mathbb{R}^n$, имеет вид

$$\int dy \delta(Ay + b)\varphi(y) = \frac{1}{|\det A|} \int dx \delta(x)\varphi(A^{-1}(x - b)) = \frac{1}{|\det A|} f(-A^{-1}b). \quad (37.95)$$

Производные δ -функции в n -мерном евклидовом пространстве $x = \{x^\alpha\} \in \mathbb{R}^n$ мы, для краткости, обозначаем индексами. Они имеют представления

$$\begin{aligned}\delta_\alpha &:= \delta_\alpha(x', x) = \frac{\partial}{\partial x'^\alpha} \delta(x', x) = -\frac{\partial}{\partial x^\alpha} \delta(x' - x), \\ \delta_{\alpha\beta} &:= \delta_{\alpha\beta}(x', x) = \frac{\partial^2}{\partial x'^\alpha \partial x'^\beta} \delta(x', x) = \frac{\partial^2}{\partial x^\alpha \partial x^\beta} \delta(x', x) = \\ &= -\frac{\partial^2}{\partial x^\alpha \partial x'^\beta} \delta(x', x) = -\frac{\partial^2}{\partial x'^\alpha \partial x^\beta} \delta(x', x).\end{aligned}\quad (37.96)$$

Формулы (37.85) без труда обобщаются на случай δ -функций в евклидовом пространстве $\mathbb{R}^n$ произвольной размерности:

$$\begin{aligned}(f - f')\delta_\alpha &= \partial_\alpha f \delta, \\ fg'\delta_\alpha &= fg\delta_\alpha - f\partial_\alpha g \delta, \\ f'g\delta_\alpha &= f'g'\delta_\alpha + f\partial_\alpha g \delta, \\ (fg' + f'g)\delta_\alpha &= (fg + f'g')\delta_\alpha, \\ (fg' - f'g)\delta_\alpha &= (\partial_\alpha fg - f\partial_\alpha g)\delta, \\ (f - f')\delta_{\alpha\beta} &= \partial_\alpha f \delta_\beta + \partial'_\beta f' \delta_\alpha = \partial_\beta f \delta_\alpha + \partial'_\alpha f' \delta_\beta, \\ fg'\delta_{\alpha\beta} &= fg\delta_{\alpha\beta} - f\partial_\alpha g \delta_\beta - f\partial_\beta g \delta_\alpha + f\partial_{\alpha\beta}^2 g \delta, \\ f'g\delta_{\alpha\beta} &= f'g'\delta_{\alpha\beta} + f'\partial'_\alpha g' \delta_\beta + f'\partial'_\beta g' \delta_\alpha + f\partial_{\alpha\beta}^2 g \delta, \\ (fg' + f'g)\delta_{\alpha\beta} &= (fg + f'g')\delta_{\alpha\beta} - (\partial_\alpha f \partial_\beta g + \partial_\beta f \partial_\alpha g)\delta, \\ (fg' - f'g)\delta_{\alpha\beta} &= (\partial_\alpha fg - f'\partial'_\alpha g')\delta_\beta - (f\partial_\beta g - \partial'_\beta f' g')\delta_\alpha.\end{aligned}\quad (37.97)$$

37.6 Функции Грина

Приведем явные выражения для функций Грина, наиболее часто встречающихся в приложениях (подробности можно найти, например, в книге [3]). Все функции Грина определены с точностью до добавления произвольного решения соответствующего однородного уравнения, которое обычно фиксируется некоторыми граничными условиями.

Пусть L – линейный дифференциальный оператор порядка m с переменными коэффициентами

$$L = \frac{d^m}{dt^m} + a_1(t) \frac{d^{m-1}}{dt^{m-1}} + \dots + a_m(t),$$

где $a_1(t), a_2(t), \dots$ – достаточно гладкие функции $t \in \mathbb{R}$. Тогда функция Грина $G(t)$ уравнения

$$LG = \delta(t) \quad (37.98)$$

имеет вид

$$G(t) = \theta(t)f(t), \quad (37.99)$$

где f – решение однородного дифференциального уравнения

$$Lf = 0$$

с начальными условиями

$$f(0) = f'(0) = \dots = f^{(m-2)}(0) = 0, \quad f^{(m-1)}(0) = 1.$$

В частности, функции

$$G(t) = \theta(t) e^{-at}, \quad (37.100)$$

$$G(t) = \theta(t) \frac{\sin at}{a} \quad (37.101)$$

являются, соответственно, функциями Грина для операторов

$$\frac{d}{dt} + a, \quad \frac{d^2}{dt^2} + a^2.$$

Уравнение Лапласа для функции Грина $G(x)$ в евклидовом пространстве $x \in \mathbb{R}^n$ имеет вид

$$\Delta G := (\partial_1^2 + \partial_2^2 + \dots \partial_n^2)G = \delta(x). \quad (37.102)$$

Функция Грина уравнения Лапласа на евклидовой плоскости $\mathbb{R}^2$ имеет вид

$$G(x) = \frac{1}{2\pi} \ln|x|. \quad (37.103)$$

Если размерность евклидова пространства больше или равна трем, то решение уравнения (37.102) имеет вид

$$G(x) = -\frac{1}{(n-2)\sigma_n} \frac{1}{|x|^{n-2}}, \quad n \geq 3, \quad (37.104)$$

где σ_n – площадь поверхности единичной сферы $\mathbb{S}^{n-1}$, вложенной в $\mathbb{R}^n$:

$$\sigma_n = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}, \quad (37.105)$$

где $\Gamma(z)$ – гамма функция Эйлера (см. пример ??). В частности,

$$\sigma_1 = 2, \quad \sigma_2 = 2\pi, \quad \sigma_3 = 4\pi, \quad \sigma_4 = 2\pi^2, \quad \sigma_5 = \frac{3}{8}\pi^2, \quad \sigma_6 = \pi^3.$$

В частности, в трехмерном евклидовом пространстве функция Грина для уравнения Лапласа совпадает с *ньютоновым потенциалом*

$$G(x) = -\frac{1}{4\pi|x|}. \quad (37.106)$$

Уравнение Гельмгольца для функции Грина в $\mathbb{R}^n$ имеет вид

$$\Delta G + k^2 G = \delta(x), \quad k = \text{const.} \quad (37.107)$$

Приведем явные выражения для функций Грина в низших размерностях:

$$G(x) = \frac{1}{2ik} e^{ik|x|}, \quad n = 1, \quad (37.108)$$

$$G(x) = -\frac{i}{4} H_0^{(1)}(k|x|), \quad n = 2, \quad (37.109)$$

$$G(x) = -\frac{e^{ik|x|}}{4\pi|x|}, \quad n = 3, \quad (37.110)$$

где $H_0^{(1)}$ – функция Ханкеля [323]. Комплексно сопряженные функции $\bar{G}(x)$ также являются решением уравнения Гельмгольца (37.107).

Уравнение теплопроводности для функции Грина в $\mathbb{R}^{n+1}$ имеет вид

$$\frac{dG}{dt} - a^2 \Delta G = \delta(t, x), \quad a = \text{const}, \quad (37.111)$$

где Δ – оператор Лапласа в $\mathbb{R}^n$. Функция Грина при любом n имеет вид

$$G(t, x) = \frac{\theta(t)}{(2a\sqrt{\pi t})^n} e^{-\frac{x^2}{4a^2 t}}. \quad (37.112)$$

Уравнение Даламбера для функции Грина в пространстве Минковского $\mathbb{R}^{1,n-1}$ с декартовыми координатами $t, x^i, i = 1, \dots, n-1$ имеет вид

$$\left(\frac{\partial^2}{\partial t^2} - c^2 \Delta \right) G = \delta(t, x), \quad (37.113)$$

где для общности введена скорость света $c = \text{const}$. В низших размерностях функции Грина записываются в виде:

$$G(t, x) = \frac{1}{2c} \theta(ct - |x|), \quad n = 2, \quad (37.114)$$

$$G(t, x) = \frac{\theta(ct - |x|)}{2\pi c \sqrt{c^2 t^2 - x^2}}, \quad n = 3, \quad (37.115)$$

$$G(t, x) = \frac{\theta(t)}{2\pi c} \delta(c^2 t^2 - x^2), \quad n = 4. \quad (37.116)$$

37.7 Ряды и интегралы Фурье

Для функции $f(t)$, заданной и абсолютно интегрируемой на отрезке $t \in [-\pi, \pi]$, определен ряд Фурье,

$$f(t) = \frac{1}{2} a_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt) = \sum_{k=-\infty}^{\infty} c_k e^{ikt}, \quad (37.117)$$

коэффициенты которого равны

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} dt f(t) \cos kt, \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} dt f(t) \sin kt, \\ c_k &= \bar{c}_{-k} = \frac{1}{2} (a_k - ib_k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} dt f(t) e^{-ikt}. \end{aligned}$$

Для непрерывных функций ряд (37.117) сходится равномерно к $f(t)$ на любом отрезке $[a, b] \subset (-\pi, \pi)$. Если функция не является непрерывной, то поточечной сходимости нет, но равенство (37.117) справедливо в слабом смысле (после интегрирования с пробной функцией).

δ -функция представляется в виде ряда Фурье

$$\delta(t', t) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{ik(t'-t)} = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{\infty} \cos[k(t' - t)]. \quad (37.118)$$

Евклидову метрику можно представить в виде интегралов

$$\delta_{ab} = \int_{-\pi}^{\pi} dt e^{i(a-b)t} = \int_{-\pi}^{\pi} dt e^{-i(a-b)t}. \quad (37.119)$$

Для функции $f(t)$, абсолютно интегрируемой на вещественной прямой $t \in \mathbb{R}$, определено преобразование Фурье,

$$F[f](\omega) = \int_{-\infty}^{\infty} dt f(t) e^{i\omega t}. \quad (37.120)$$

Обратное преобразование Фурье имеет вид

$$f(t) = F^{-1}[F[f]] = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega F(\omega) e^{-i\omega t}. \quad (37.121)$$

Отсюда следует, что преобразование Фурье $\delta(t', t)$ по аргументу t' равно

$$F[\delta(t', t)] = e^{i\omega t}.$$

В частности $F[\delta(t')] = 1$. В результате получаем интегральное представление для δ -функции

$$\delta(t', t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{i\omega(t'-t)} = \frac{1}{\pi} \int_0^{\infty} d\omega \cos[\omega(t' - t)]. \quad (37.122)$$

В приложениях часто встречаются преобразования Фурье от следующих функций:

$$F \left[e^{-\frac{a^2 t^2}{2}} \right] = \frac{\sqrt{2\pi}}{a} e^{-\frac{\omega^2}{2a^2}}, \quad (37.123)$$

$$F \left[e^{i\frac{t^2}{2}} \right] = \sqrt{2\pi} e^{-i\frac{(\omega^2 - \pi)}{2}}, \quad (37.124)$$

$$F[\theta(t)] = \pi\delta(\omega) + i\mathcal{P}\frac{1}{\omega}, \quad (37.125)$$

Пусть квадратичная форма

$$(xAx) := \sum_{i,j=1}^n x_i A_{ij} x_j$$

вещественна и положительно определена. Тогда

$$F \left[e^{-\frac{1}{2}(xAx)} \right] (\omega) = \frac{(2\pi)^{n/2}}{\sqrt{\det A}} e^{-\frac{1}{2}(\omega A^{-1}\omega)}, \quad (37.126)$$

где $\omega = (\omega_1, \dots, \omega_n)$.

Справедливы следующие формулы для сумм:

$$\sum_{k=-\infty}^{\infty} \delta(x - 2\pi k) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{ikx}, \quad (37.127)$$

$$\sum_{k=-\infty}^{\infty} \delta(x - k) = \sum_{k=-\infty}^{\infty} e^{i2\pi kx}, \quad (37.128)$$

$$\sum_{k=-\infty}^{\infty} e^{-tk^2} = \sqrt{\frac{\pi}{t}} \sum_{k=-\infty}^{\infty} e^{-\frac{\pi^2}{t}k^2}. \quad (37.129)$$

37.8 Интегралы

Приведем формулы для простейших гауссовых интегралов:

$$\int_{-\infty}^{\infty} dx e^{-ax^2} = \sqrt{\frac{\pi}{a}}, \quad (37.130)$$

$$\int_{-\infty}^{\infty} dx x^2 e^{-ax^2} = \sqrt{\frac{\pi}{a}} \left(\frac{1}{2a} \right), \quad (37.131)$$

$$\int_{-\infty}^{\infty} dx x^4 e^{-ax^2} = \sqrt{\frac{\pi}{a}} \left(\frac{3}{4a^2} \right), \quad (37.132)$$

$$\int_{-\infty}^{\infty} dx x^6 e^{-ax^2} = \sqrt{\frac{\pi}{a}} \left(\frac{15}{8a^3} \right), \quad (37.133)$$

$$\int_{-\infty}^{\infty} V dx e^{-ax^2+bx} = \sqrt{\frac{\pi}{a}} e^{b^2/4a}, \quad (37.134)$$

$$\int_{-\infty}^{\infty} dx x e^{-ax^2+bx} = \sqrt{\frac{\pi}{a}} \left(\frac{b}{2a} \right) e^{b^2/4a}, \quad (37.135)$$

$$\int_{-\infty}^{\infty} dx x^2 e^{-ax^2+bx} = \sqrt{\frac{\pi}{a}} \left(\frac{1}{2a} \right) \left(1 + \frac{b^2}{2a} \right) e^{b^2/4a}, \quad (37.136)$$

Пусть $x = \{x^i\}$ и $J = \{J_i\}$, $i = 1, \dots, n$ – векторы евклидова пространства, и A_i^j – положительно определенная квадратная матрица, тогда

$$\int dx \exp \left[-\frac{1}{2}(xAx) + (Jx) \right] = \frac{(2\pi)^{n/2}}{\sqrt{\det A}} \exp \left[\frac{1}{2}(JA^{-1}J) \right], \quad (37.137)$$

$$\int d^n x \exp \left[\frac{i}{2}(xAx) + i(Jx) \right] = \frac{(2\pi i)^{n/2}}{\sqrt{\det A}} \exp \left[\frac{-i}{2}(JA^{-1}J) \right], \quad (37.138)$$

где $(xAx) := x^i A_i^j x_j$ и $(Jx) := J^i x_i$.

Приведем интегральное представление произведения двух чисел,

$$\frac{1}{ab} = \int_0^1 \frac{dx}{[ax + b(1-x)]^2}. \quad (37.139)$$

$$\frac{1}{ab^2} = 2 \int_0^1 dx \frac{1-x}{[ax + b(1-x)]^3}. \quad (37.140)$$

В общем случае справедливы две формулы Фейнмана, позволяющие преобразовать произведение n множителей в интеграл от n -той степени некоторого выражения (см., например, [324], упражнение I-8):

$$\begin{aligned} \frac{1}{a_1 a_2 \dots a_n} &= (n-1)! \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \dots \int_0^{1-x_1-x_2-\dots-x_{n-1}} dx_{n-1} \times \\ &\times \frac{1}{[(a_{n-1}-a_n)x_{n-1} + (a_{n-2}-a_n)x_{n-2} + \dots + (a_1-a_n)x_1 + a_n]^n}, \end{aligned} \quad (37.141)$$

$$\begin{aligned} \frac{1}{a_1 a_2 \dots a_n} &= (n-1)! \int_0^1 dx_1 \int_0^1 dx_2 \dots \int_0^1 dx_{n-1} \times \\ &\times \frac{x_1^{n-2} x_2^{n-3} \dots x_{n-2}}{[(a_n - a_{n-1})x_{n-1} \dots x_1 + (a_{n-1} - a_{n-2})x_{n-2} \dots x_1 + \dots + (a_2 - a_1)x_1 + a_1]^n}. \end{aligned} \quad (37.142)$$

В этих формулах все сомножители считаются положительными, $a_i > 0$, $i = 1, \dots, n$. Заметим, что первое представление имеет переменные пределы интегрирования, в то время как второе – постоянные.

Пусть θ^i , $i = 1, \dots, n$ – набор антикоммутирующих переменных:

$$\{\theta^i, \theta^j\} = 0.$$

По определению,

$$\int d\theta^i = 0, \quad \int d\theta^i \theta^j = \delta^{ij}. \quad (37.143)$$

Тогда интеграл от произвольной функции $F(\theta)$ определяется следующим образом. Разложим функцию $F(\theta)$ в ряд Тейлора по θ^i . Он имеет конечное число слагаемых и может зависеть не более, чем линейно от каждой переменной θ^i . После этого соотношения (37.143) определяют интеграл

$$\int d\theta F(\theta)$$

единственным образом.

Справедлива формула

$$\det A = \int d\theta d\bar{\theta} e^{i\theta A \bar{\theta}}, \quad (37.144)$$

где θ и $\bar{\theta}$ – независимые наборы антикоммутирующих переменных.

Рассмотрим область в евклидовом пространстве $\mathbb{U} \subset \mathbb{R}^n$. Пусть на $\mathbb{U}$ задан набор функций $f_i(x)$, $i = 1, \dots, \kappa$. Если все интегралы, выписанные ниже, определены, то справедливо *неравенство Гёльдера*

$$\left| \int_{\mathbb{U}} dx f_1 \dots f_{\kappa} \right| \leq \prod_{i=1}^{\kappa} \left(\int_{\mathbb{U}} dx |f_i|^{p_i} \right)^{1/p_i}, \quad (37.145)$$

где числа p_i удовлетворяют равенству

$$\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_{\kappa}} = 1, \quad p_i > 1, \quad 1 \leq i \leq \kappa.$$

В частности, для двух функций f, g имеет место неравенство

$$\left| \int_{\mathbb{U}} dx fg \right| \leq \left(\int_{\mathbb{U}} dx |f|^p \right)^{1/p} \left(\int_{\mathbb{U}} dx |g|^q \right)^{1/q}, \quad (37.146)$$

где

$$\frac{1}{p} + \frac{1}{q} = 1, \quad p > 1, \quad q > 1.$$

При $p = 2$ и $q = 2$ получаем *неравенство Коши–Буняковского*

$$\left| \int_{\mathbb{U}} dx fg \right| \leq \left(\int_{\mathbb{U}} dx f^2 \right)^{1/2} \left(\int_{\mathbb{U}} dx g^2 \right)^{1/2}. \quad (37.147)$$

Справедливо также *неравенство Минковского*

$$\left(\int_{\mathbb{U}} dx |f + g|^p \right)^{1/p} \leq \left(\int_{\mathbb{U}} dx |f|^p \right)^{1/p} + \left(\int_{\mathbb{U}} dx |g|^p \right)^{1/p}, \quad p \geq 1. \quad (37.148)$$

Литература

- [1] В. И. Арнольд. *Обыкновенные дифференциальные уравнения*. Наука, Москва, 1975. Второе изд. 240 с.
- [2] В. И. Арнольд, В. В. Козлов, and А. И. Нейштадт. *Математические аспекты классической и небесной механики*. Эдиториал УРСС, Москва, 2002. Второе изд. 414 с.
- [3] В. С. Владимиров. *Уравнения математической физики*. Наука, Москва, издание пятое, 1988.
- [4] И. М. Гельфанд. *Лекции по линейной алгебре*. МЦНМО, Москва, 1998. Пятое изд.
- [5] И. М. Гельфанд, Р. А. Минлос, and З. Я. Шапиро. *Представления группы вращений и группы Лоренца, их применения*. Физ.-мат. лит., Москва, 1958.
- [6] Ю. Н. Дрождинов and Б. Н. Завьялов. *Введение в теорию обобщенных функций*. МИАН, Москва, 2006.
- [7] Б. А. Дубровин, С. П. Новиков, and А. Т. Фоменко. *Современная геометрия. Методы и приложения*. Наука, Москва, издание четвертое, 1998.
- [8] В. В. Жаринов. *Алгебро-геометрические основы математической физики*. Изд-во МИАН им. В. А. Стеклова, Москва, 2008.
- [9] J. L. Kelley. *General Topology*. D. Van Nostrand Company, Inc., Toronto – London, 1957. Перевод: Келли Дж. Л. *Общая топология*. М.: Наука, 1968.
- [10] А. А. Кириллов. *Элементы теории представлений*. Наука, Москва, Второе edition, 1978.
- [11] А. И. Кострикин. *Введение в алгебру. Часть I. Основы алгебры. Часть II. Линейная алгебра. Часть III. Основные структуры алгебры*. Наука, Москва, 2000.
- [12] А. И. Кострикин and Ю. И. Манин. *Линейная алгебра и геометрия*. Наука, Москва, 1986.
- [13] С. П. Новиков and И. А. Тайманов. *Современные геометрические структуры и поля*. МЦНМО, Москва, 2005.
- [14] М. М. Постников. *Аналитическая геометрия*. Наука, Москва, 1979.
- [15] П. К. Рашевский. *Риманова геометрия и тензорный анализ*. Наука, Москва, 1967. Третье изд.

- [16] В. А. Рохлин and Д. Б. Фукс. *Начальный курс топологии. Геометрические главы*. Наука, Москва, 1977.
- [17] Ф. Р. Гантмахер. *Теория матриц. 4-е изд.* Наука, Москва, 1988.
- [18] М. А. Евграфов. *Аналитические функции*. Наука, Москва, Третье edition, 1991.
- [19] Czes Kosniowski. *A First Course In Algebraic Topology*. Cambridge University Press, Cambridge – London – New York, 1980. Перевод: Коснёвски Ч. *Начальный курс алгебраической топологии*. М.: Мир, 1983.
- [20] R. Engelking. *General Topology*. Państwowe Wydawnictwo Naukowe, Warszawa, second edition, 1977. Перевод: Энгелькинг, Р. *Общая топология*. М.: Мир, 1986. 751 с.
- [21] A. N. Tychonoff. Ein Fixpunktsatz. *Math. Ann.*, 11:767–776, 1935.
- [22] L. Schwartz. *Analyse Mathématique. Vols. I, II*. Hermann, Paris, 1967. Перевод: Л. Шварц. *Анализ*. Том 1,2. М.: Мир, 1972.
- [23] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge U.P., Cambridge, 1986. Перевод: Хорн Р., Джонсон Ч. *Матричный анализ* М.: Мир. 1989.
- [24] Л. Д. Кудрявцев. *Курс математического анализа. Том I, II, III*. Высшая школа, Москва, 1981.
- [25] М. А. Наймарк. *Линейные представления группы Лоренца*. Физматгиз, Москва, 1958.
- [26] A. Einstein. Zur Elektrodynamik der bewegter Körper. *Ann. Phys.*, 17(2):891–921, 1905. Перевод: Альберт Эйнштейн. Собрание научных трудов. Т. 1. М.: Наука, 1965. С. 7–35.
- [27] Н. Х. Ибрагимов. *Группы преобразований в математической физике*. Наука, Москва, 1983.
- [28] E. Cunningham. The principle of relativity in electrodynamics and an extension thereof. *Proc. London Math. Soc.*, 8, Series 2:77–98, 1909.
- [29] H. Bateman. The transformation of the electrodynamical equations. *Proc. London Math. Soc.*, 8, Series 2:223–264, 1909.
- [30] P. A. M. Dirac. Wave equations in conformal space. *Ann. Math.*, 37(2):429–442, 1936. Перевод в сб. П. А. М. Дирак “Собрание научных трудов. Том 2.” М.: Наука, 2003, сс. 485–497.
- [31] E. S. Fradkin and M. Ya. Palchik. *Conformal Quantum Field Theory in D-dimensions*. Kluwer Academic Publishers, Dordrecht, 1996.
- [32] A. O. Barut and R. Rączka. *Theory of Group Representations and Applications*. PWN – Polish Scientific Publishers, Warszawa, 1977. Перевод: А. Барут, Р. Рончка. *Теория представлений групп и ее приложения*. Т.1,2. М.: Мир, 1980.

- [33] A. A. Michelson and E. W. Morley. On the relative motion of the earth and the luminiferous ether. *Am. J. Sci*, 34:333–345, 1887. Воспроизведена в кн.: *Relativity Theory: Its Origins and Impact on Modern Thought*. Ed. L. Pearce Williams, John Wiley and Sons, 1968.
- [34] T. S. Jaseja, A. Javan, J. Murray, and C.H. Townes. Test of special relativity or of the isotropy of space by use of infrared masers. *Phys. Rev.*, 133:A1221–A1225, 1964.
- [35] G. Herglotz. Über die Mechanik des deformierbaren Körpers vom Standpunkte der Relativitätstheorie. *Ann. der Physik*, 36(3):493–533, 1911.
- [36] H. Whitney. Differentiable manifolds. *Ann. of Math. (2)*, 37(5):645–680, 1936.
- [37] J. Munkres. Obstructions to the smoothing of piecewise differentiable homeomorphisms. *Ann. of Math. (2)*, 72:521–554, 1960.
- [38] J. Milnor. On manifolds homeomorphic to the 7-sphere. *Ann. Math.*, 64:394–405, 1956. Перевод: Дж. Милнор. Сб. переводов "Математика". 1990, т. 1, 1. 3, сс. 394–405.
- [39] S. K. Donaldson and P. B. Kronheimer. *The Geometry of Four-Manifolds*. Clarendon Press, Oxford, 1991.
- [40] M. Kervaire. A manifold which does not admit any differentiable structure. *Comment. Math. Helv.*, 34:257–270, 1960.
- [41] T. Aubin. *A Course in Differential Geometry*. AMS, Providence, Rhode Island, 2001.
- [42] F. W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Springer-Verlag, Berlin – Heidelberg, 1983. Перевод: Ф. Уорнер. *Основы теории гладких многообразий и групп Ли*. М.: Мир, 1987.
- [43] von R. Sulanke and P. Wintgen. *Differentialgeometrie und Faserbündel*. Veb Deutscher Verlag Der Wissenschaften, Berlin, 1972. Перевод: Р. Зуланке, П. Винтген. *Дифференциальная геометрия и расслоения*. М.: ИЛ, 1975.
- [44] C. Godbillon. *Géométrie Différentielle et Mécanique Analytique*. Hermann, Paris, 1969. Перевод: Годбийон К. *Дифференциальная геометрия и аналитическая механика*. М.: Мир, 1973. 188 с.
- [45] S. S. Chern. Curves and surfaces in euclidean space. In S. S. Chern, editor, *"Global Geometry and Analysis"*. *MAA studies in Mathematics*. Vol. 4., pages 16–56, Inglewood Cliffs, 1967. The Mathematical Association of America, Prentice Hall, Inc.
- [46] И. Г. Петровский. *Лекции по теории обыкновенных дифференциальных уравнений*. Наука, Москва, 1970.
- [47] C. J. Isham. *Modern Differential Geometry*. World Scientific, Singapore, 1999.

- [48] S. Kobayashi and K. Nomizu. *Foundations of differential geometry*, volume 1, 2. Interscience publishers, New York – London, 1963. Перевод: Кобаяси Ш., Номидзу К. *Основы дифференциальной геометрии*. Том 1, 2. – М.: Наука, 1981.
- [49] S. S. Chern, W. H. Chen, and K. S. Lam. *Lectures on Differential Geometry*. World Scientific, Singapore, 2000.
- [50] C. Chevalley. *Theory of Lie Groups*. Princeton University Press, Princeton, 1946. К. Шевалле. *Теория групп Ли. Т. 1. М.: ИЛ, 1948*.
- [51] H. Grassmann. *Die lineale Ausdehnungslehre – Ein neuer Zweig der Mathematik*. Leipzig, 1844.
- [52] G. de Rham. Sur la théorie des formes différentielles harmoniques. *Ann. Univ. Grenoble*, 22:132–152, 1946.
- [53] G. de Rham. *Variétés Différentiables*. Hermann, Paris, 1955. Перевод: де Рам Ж. *Дифференцируемые многообразия*. М.: URSS, 2006. 2-е изд.
- [54] J. A. Wolf. *Spaces of constant curvature*. University of California, Berkley, California, 1972. Перевод: Вольф Дж. *Пространства постоянной кривизны*. М.: Наука, 1982. 480 с.
- [55] N. E. Steenrod. *The Topology of Fiber Bundles*. Princeton University Press, Princeton, New Jersey, 1951.
- [56] H. Weyl. Gravitation und Elektrizität. *Sitz. Preuss. Akad. Wiss.*, page S. 465, 1918. Перевод в сб. “Альберт Эйнштейн и теория гравитации”. М.: Мир, 1979, с. 513.
- [57] T. Levi-Civita. Nozione di parallelismo una varietà qualunque e conseguente specificazione geometrica della curvatura Riemanniana. *Rendicinti di Palermo*, 42:173–205, 1917.
- [58] М. О. Катанаев. Геометрическая теория дефектов. *УФН*, 175(7):705–733, 2005.
- [59] A. M. Gleason. Groups without small subgroups. *Ann. Math.*, 56(2):193–212, 1952.
- [60] D. Montgomery and L. Zippin. Small subgroups of finite-dimensional groups. *Ann. Math.*, 56(2):213–241, 1952.
- [61] Понтрягин Л. С. *Непрерывные группы*. Наука, Москва, 1984. Четвертое изд.
- [62] L. Bianchi. Sugli spazii a tre dimensioni che ammettono un gruppo continuo di movimenti. *Soc. Ital. Sci. Mem. di Mat.*, 11:267, 1897.
- [63] M. Goto and Grosshans F. *Semisimple Lie Algebras*. Marcel Dekker, New York, 1978.
- [64] S. Helgason. *Differential Geometry and Symmetric Spaces*. Academic Press, New York, 1962. Перевод: Хелгасон С. *Дифференциальная геометрия и симметрические пространства*. М.: Мир, 1983.
- [65] E. Hewitt and K.A. Ross. *Abstract Harmonic Analysis, I*. Springer–Verlag, Heidelberg, 1963. Перевод: Хьюитт, Э., Росс, К. *Абстрактный гармонический анализ, I*. М.: Мир, 1974.

- [66] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. American Mathematical Society, Rhode Island, 2001. Перевод: Хелгасон С. *Дифференциальная геометрия, группы Ли и симметрические пространства*. М.: Факториал Пресс, 2005.
- [67] М. М. Постников. *Группы и алгебры Ли*. Наука, Москва, 1982.
- [68] Е. Б. Дынкин. Структура полупростых алгебр Ли. *УМН*, 2(4(20)):59–127, 1947.
- [69] И. Д. Адо. Представление алгебр Ли матрицами. *Усп. Матем. наук*, 2(6):159, 1947.
- [70] E. Cartan. Groupes simples clos et ouverts et g'eometrie riemannienne. *J. Math. pure appl.*, 8:1–33, 1929.
- [71] M. Hausner and J. Schwartz. *Lie Groups; Lie algebras*. Gordon & Breach, New York–London–Paris, 1968.
- [72] J. F. Adams. Vector fields on spheres. *Ann. Math.*, 75:603–632, 1962.
- [73] R. Narasimhan. *Analysis on Real and Complex Manifolds*. Masson and North–Holland, Paris and Amsterdam, 1971. Перевод: Р. Нарасимхан. *Анализ на действительных и комплексных многообразиях*. М.: Мир, 1971.
- [74] E. H. Spanier. *Algebraic Topology*. McGraw–Hill Book Company, New York – London, 1966. Перевод: Э. Спеньер. *Алгебраическая топология*. М.: Мир, 1971.
- [75] А. А. Болибрух. *Фуксовы дифференциальные уравнения и голоморфные расслоения*. МЦНМО, Москва, 2000.
- [76] K. Iwasawa. On some types of topological groups. *Ann. Math.*, 50:507–558, 1949.
- [77] K. Nomizu and H. Ozeki. On the degree of differentiability of curves used in the definition of the holonomy groups. *Bull. Amer. Math. Soc.*, 68:74–75, 1962.
- [78] W. Ambrose and I. M. Singer. A theorem on holonomy. *Trans. Amer. Math. Soc.*, 75:428–443, 1953.
- [79] J. Hano and H. Ozeki. On the holonomy groups of linear connexions. *Nagoya Math. J.*, 10:97–100, 1956.
- [80] K. Nomizu. Un théorème sur les groupes d'holonomie. *Nagoya Math. J.*, 10:101–103, 1956.
- [81] H. Ozeki. Infinitesimal holonomy groups of bundle connections. *Nagoya Math. J.*, 10:105–123, 1956.
- [82] H. C. Wang. On invariant connections over a principal fibre bundle. *Nagoya Math. J.*, 13:1–19, 1958.
- [83] M. V. Berry. Quantal phase factors accompanying adiabatic changes. *Proc. Roy. Soc. London*, A392(1802):45–57, 1984.
- [84] Y. Aharonov and D. Bohm. Significance of electromagnetic potentials in the quantum theory. *Phys. Rev.*, 115(3):485–491, 1959.

- [85] F. Wilczek and A. Zee. Appearance of gauge structure in simple dynamical systems. *Phys. Rev. Lett.*, 52:2111, 1984.
- [86] M. Born and V. Fock. Beweis des Adiabatenatzes. *Z. Phys.*, 51:165–180, 1928. English translation in “V.A. Fock – Selected Works: Quantum Mechanics and Quantum Field Theory” ed. by L.D. Faddeev, L.A. Khalin, I.V. Komarov. Chapman & Hall/CRC, Boca Raton, 2004.
- [87] E. Schrödinger. Quantizierung als Eigenwertproblem (Erste Mitteilung). *Ann. Phys. Leipzig*, 79(4):361–376, 1926.
- [88] E. Schrödinger. Quantizierung als Eigenwertproblem (Zweite Mitteilung). *Ann. Phys. Leipzig*, 79(6):489–527, 1926.
- [89] В. С. Владимиров and И. В. Волович. Локальные и нелокальные токи для нелинейных уравнений. *ТМФ*, 61(3):3–29, 1984.
- [90] В. С. Владимиров and И. В. Волович. Законы сохранения для нелинейных уравнений. *УМН*, 40(4(244)):17–26, 1985.
- [91] A. Messiah. *Quantum Mechanics*, volume 2. North Holland, Amsterdam, 1962. Перевод: А. Мессиа. *Квантовая механика*. Т.2. М.: Наука, 1979.
- [92] В. А. Фок. *Начала квантовой механики*. Наука, Москва, 2-е изд. edition, 1976.
- [93] T. Bitter and D. Dubbers. Manifestation of berry’s topological phase in neutron spin rotation. *Phys. Rev. Lett.*, 59:251–254, 1987.
- [94] F. G. Werner and D. R. Brill. Significance of electromagnetic potentials on the quantum theory in the interpretation of electron interferometer fringe observations. *Phys. Rev. Lett.*, 4(7):344–347, 1960.
- [95] R. G. Chambers. Shift of an electron interference pattern by enclosed magnetic flux. *Phys. Rev. Lett.*, 5(1):3–5, 1960.
- [96] H. Boersch, H. Hamisch, D. Wohlleben, and K. Grohmann. Weissche bereiche als bi-prisme für elektroneninterferenzen. *Z. Phys.*, 159:397–404, 1960.
- [97] W. Killing. Über die Grunlagen der Geometrie. *J. Reine Angew. Math.*, 109:121–186, 1892.
- [98] L. Bianchi. *Lezioni sulla teoria dei gruppi continui finiti di trasformazioni*. Sporerri, Pisa, 1918.
- [99] W. de Sitter. On the relativity of inertia: Remark concerning einstein’s latest hypothesis. *Proc. Sect. Sci. K. ned. Akad. Wet.*, 19:1217–1225, 1916–1917.
- [100] W. de Sitter. On the curvature of space. *Proc. Sect. Sci. K. ned. Akad. Wet.*, 20:229–243, 1917.
- [101] В. А. Фок. *Теория пространства, времени и тяготения*. Издание второе. Физматгиз, Москва, 1961.
- [102] P. Stäckel. *Über die Integration der Hamilton–Jacobischen Differentialgleichung mittelst Separation der Variablen*. Habilitationsschrift, Halle, 1891.

- [103] P. Stäckel. Über die Bewegung eines Punktes in einer n -fachen Mannigfaltigkeit. *Math. Ann.*, 42:537–563, 1893.
- [104] P. Stäckel. Sur des problem de dynamique se reduisent a des quadratures. *Comptes rendus hebd, S. Acad. Sci. (Paris)*, 116:1284–1286, 1893.
- [105] P. Stäckel. Sur une classe de problemes de dynamique. *Comptes rendus hebd, S. Acad. Sci. (Paris)*, 116:485–487, 1893.
- [106] P. Stäckel. Sur l'integration de l'équation différentielle de Hamilton. *Comptes rendus hebd, S. Acad. Sci. (Paris)*, 121:489–492, 1895.
- [107] R. P. Kerr. Gravitational field of a spinning mass as an example of algebraically special metrics. *Phys. Rev. Lett.*, 11(5):237–238, 1963.
- [108] В. В. Обухов. *Штеккелевы пространства в теории гравитации*. Изд. Томского государственного педагогического университета, Томск, 2006.
- [109] В. В. Обухов and К. Е. Осетрин. *Классификационные проблемы в теории гравитации*. Изд. Томского государственного педагогического университета, Томск, 2007.
- [110] Th. De Donder. *La Gravifique Einsteinienne*. Gauthier-Villars & cie, Paris, 1921.
- [111] K. Lanczos. Ein vereinfachendes Koordinatensystem für die Einsteinschen Gravitationsgleichungen. *Phys. Zs.*, 23:537–539, 1922.
- [112] В. А. Фок. О движении конечных масс в общей теории относительности. *ЖЭТФ*, 9(4):375–410, 1939.
- [113] E. Fermi. Sopra i fenomeni che avvengono in vicinanza di una linea oraria. *Atti Acad. Naz. Lincei Rend. Cl. Sci. Fiz. Mat. Nat.*, 31:Three parts, pp. 21–23, 51–52, 101–103, 1922.
- [114] J. H. C. Whitehead. Convex regions in the geometry of paths. *Quart. J. Math. Oxford Ser.*, 3:33–42, 1932.
- [115] S. Myers and N. Steenrod. The group of isometries of a Riemannian manifold. *Ann. Math.*, 40:400–416, 1939.
- [116] H. Hopf and W. Rinow. Über den Begriff des vollständigen differentialgeometrischen Fläche. *Comment. Math. Helv.*, 3(2):209–225, 1931.
- [117] M. Hamermesh. *Group Theory and its Application to Physical Problems*. Addison–Wesley Publishing Company, Reading, Massachusetts, 1964.
- [118] É. Cotton. Sur les variétés a trois dimensions. *Annales de la Faculté des Sciences de Toulouse. Thèse, Paris.*, I(4):385–438, 1899.
- [119] H. Weyl. Reine Ifinitesimalgeometrie. *Mathematische Zeitschrift*, 2:384–411, 1918.
- [120] J. A. Schouten. über die Konforme Abbildung n -dimensionaler Mannigfaltigkeiten mit quadratischer Massbestimmung auf eine Mannigfaltigkeit mit euklidischer Massbestimmung. *Mathematische Zeitschrift*, 11:58–88, 1918.

- [121] А. З. Петров. Классификация пространств, определяемых полями тяготения. *Уч. зап. Каз. ун-та.*, 114(8):55, 1954. English translation, with historical comments: A. Z. Petrov “The classification of Spaces Defining Gravitational Fields” *Gen. Rel. Grav.* 32, No.8 (2000) 1665–1685.
- [122] А. З. Петров. *Новые методы в общей теории относительности*. Наука, Москва, 1966.
- [123] C. H. Brans. Some restrictions on algebraically general vacuum metrics. *J. Math. Phys.*, 16(4):1008, 1975.
- [124] В. Д. Захаров. Физические характеристики пространств Эйнштейна вырожденного типа ii по классификации Петрова. *ДАН*, 161:563, 1965.
- [125] В. Д. Захаров. Алгебраические и теоретико-групповые методы в общей теории относительности: Инвариантная характеристика типа Петрова некоторых пространств Эйнштейна. *Пробл. теории грав. и элем. частиц*. М: Атомиздат, 3:128, 1971.
- [126] A. Einstein and B. Kaufman. A new form of the general relativistic field equations. *Ann. Math. Princeton*, 62:128–138, 1955. Перевод в кн.: Эйнштейн А. Собрание научных трудов. Т. 2. М.: Наука, 1966. С. 835–848.
- [127] A. Einstein. *The Meaning of Relativity. 5-th edn.* Princeton University Press, Princeton, 1955. Перевод в кн.: А. Эйнштейн. *Собрание научных трудов*. Т. 2. М.: Наука, 1966. С. 849–873.
- [128] Yu. N. Obukhov. Conformal invariance and space-time torsion. *Phys. Lett.*, 90A(1,2):13–16, 1982.
- [129] T. Dereli and R. W. Tucker. A note on a generalization of Weyl’s theory of gravitation. *J. Phys. A*, 15:L7–L11, 1982.
- [130] S. Kobayashi and K. Nomizu. *Foundations of differential geometry*, volume 1, 2. Interscience publishers, New York – London, 1963, 1969.
- [131] А. Т. Фоменко. *Симплектическая геометрия*. Издательство МГУ, Москва, 1988.
- [132] Л. Д. Ландау and Е. М. Лифшиц. *Механика*. Наука, Москва, четвертое edition, 1988.
- [133] P. J. Olver. *Application of Lie Groups to Differential Equations*. Springer–Verlag, Berlin – Heidelberg, 1986. Перевод: Олвер П. *Приложения групп Ли к дифференциальным уравнениям*. М.: Мир, 1989. 637 с.
- [134] Л. Э. Эльсгольц. *Дифференциальные уравнения и вариационное исчисление*. УРСС, Москва, 1998.
- [135] E. Noether. Invariante variationsprobleme. *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys. Klasse*, pages 235–257, 1918. Перевод в кн. *Вариационные принципы в механике*. М.: Физматгиз, 1959, с. 611.

- [136] B. DeWitt. *Dynamical Theory of Groups and Fields*. Gordon and Breach, New York – London, 1965. Перевод: ДеВитт Б. С. *Динамическая теория групп и полей*. М.: Наука, 1987. 287 с.
- [137] S. Coleman. Classical lumps and their quantum descendants. lectures at the 1975 international school of subnuclear physics "ettore majorana"(erice). In A. Zichichi, editor, *New Phenomena in Subnuclear Physics*, pages 297–421, New York, 1977. Plenum Press.
- [138] Л. Д. Фаддеев. В поисках многомерных солитонов. In Д2-9788, editor, *Нелокальные, нелинейные и неренормируемые теории поля*, pages 207–223, Дубна, 1977. ОИЯИ.
- [139] R. Palais. The principle of symmetric criticality. *Comm. Math. Phys.*, 69(1):19–30, 1979.
- [140] О. А. Ладыженская and Л. В. Капитанский. О принципе Коулмена нахождения стационарных точек инвариантных функционалов. *Зап. науч. сем. ЛОМИ*, 127(15):84–102, 1983.
- [141] R. Schmid and L. Simoni. On infinite-dimensional variational principles with constraints. *J. Math. Phys.*, 30(5):1171–1176, 1989.
- [142] L. Michel and L. Radicati. On the dynamical breaking of $su(3)$. In B. Krsunoglu A. Pearlmutter, C. A. Hurst, editor, *Proc. 5th Coral Gables Conf. Symmetry Principle at High Energy*, New York, 1968. Benjamin.
- [143] В. И. Арнольд. *Математические методы классической механики*. Наука, Москва, 1989. Третье изд. 472 с.
- [144] Lee Hwa-Chung. Invariants of hamilton systems and applications to to the theory of canonical transformations. *Proc. Roy. Soc. Edinburg*, A62:237–247, 1947.
- [145] Ф. Р. Гантмахер. *Лекции по аналитической механике*. Третье изд. Физматлит, Москва, 2001.
- [146] P. A. M. Dirac. Generalized Hamiltonian dynamics. *Proc. Roy. Soc. London*, A246:326–332, 1958. Перевод в сб. "Новейшие проблемы гравитации" под редакцией Д. Иваненко. М.: ИЛ, 1961.
- [147] P. A. M. Dirac. The theory of gravitation in hamiltonian form. *Proc. Roy. Soc. London*, A246(1246):333–343, 1958. Перевод в сб. "Новейшие проблемы гравитации" под редакцией Д. Иваненко. М.: ИЛ, 1961.
- [148] P. A. M. Dirac. *Lectures on Quantum Mechanics*. Belfer, Yeshiva University, New York, 1964. Перевод в кн.: Дирак П. А. М. *Принципы квантовой механики*. М.: Наука, 1979. С. 408–475.
- [149] D. M. Gitman and I. V. Tyutin. *Quantization of Fields with Constraints*. Springer-Verlag, Berlin – Heidelberg, 1990.
- [150] M. Henneaux and C. Teitelboim. *Quantization of Gauge Systems*. Princeton University Press, Princeton, 1992.

- [151] Л. Д. Фаддеев. Интеграл Фейнмана для сингулярных лагранжианов. *ТМФ*, 1(1):3–18, 1969.
- [152] E. S. Fradkin and G. A. Vilkovisky. Quantization of relativistic systems with constraints – equivalence of canonical and covariant formalism in quantum theory of gravitational field. *CERN Preprint TH 2332*, 1977.
- [153] Н. Н. Боголюбов and Д. В. Ширков. *Введение в теорию квантованных полей*. Наука, Москва, издание четвертое, 1984.
- [154] A. Einstein. Die Feldgleichungen der Gravitation. *Sitzungsber. preuss. Akad. Wiss.*, 48(2):844–847, 1915. Перевод: Альберт Эйнштейн. Собрание научных трудов. Т. 1. М.: Наука, 1965. С. 448–451.
- [155] D. Hilbert. Die Grundlagen der Physik. *Nachrichten K. Gesellschaft Wiss. Göttingen, Math.-phys.*, Heft 3:395, Klasse 1915. Перевод в сб.: Альберт Эйнштейн и теория гравитации. М.: Мир, 1979. С. 133–145.
- [156] A. Einstein. Die-Grundlage der allgemeinen Relativitätstheorie. *Ann. d. Phys.*, 49:769–822, 1916. Перевод в сб.: Альберт Эйнштейн и теория гравитации. М.: Мир, 1979. С. 146–196.
- [157] von M. Fierz. über die physikalische deutung der erweiterten gravitationstheorie P. Jordans. *Helv. Phys. Acta*, 29:128–134, 1956.
- [158] P. Jordan. Zum gegenwärtigen stand der diracschen kosmologischen hypothesen. *Z. Phys.*, 157:112–121, 1959.
- [159] C. Brans and R. H. Dicke. Mach’s principle and a relativistic theory of gravitation. *Phys. Rev.*, 124(3):925–935, 1961.
- [160] P. A. M. Dirac. A new basis for cosmology. *Proc. Roy. Soc. London*, A165(921):199–208, 1938.
- [161] A. Peres. Polynomial expansion of gravitational lagrangian. *Nuovo Cimento*, 28(4):865–867, 1963.
- [162] M. O. Katanaev. Polynomial form of the Hilbert–Einstein action. *Gen. Rel. Grav.*, 38:1233–1240, 2006. gr-qc/0507026.
- [163] R. Penrose. A remarkable property of plane waves in general relativity. *Rev. Mod. Phys.*, 37(1):215–220, 1965.
- [164] N. A. Chernikov and E. A. Tagirov. Quantum theory of scalar field in de Sitter space-time. *Ann. Inst. Henri Poincaré*, A9(2):109–141, 1968.
- [165] Н. Х. Ибрагимов. К групповой классификации дифференциальных уравнений второго порядка. *ДАН*, 183(2):274–277, 1968.
- [166] C. G. Callan, S. Coleman, and R. Jackiw. A new improved energy–momentum tensor. *Ann. Phys.*, 59:42–73, 1970.
- [167] L. Infeld and J. Plebański. *Motion and Relativity*. Pergamon Press, Oxford – London, second edition, 1960. Перевод: Л. Инфельд, Е. Плебаньский. *Движение и релятивизм*. М.: ИЛ. 1962.

- [168] M. Fierz and W. Pauli. Relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc. Roy. Soc. London*, A173:211–232, 1939.
- [169] Л. Д. Ландау and Е. М. Лифшиц. *Теория поля*. Наука, Москва, седьмое edition, 1988.
- [170] W. G. Unruh. Experimental black-hole evaporation? *Phys. Rev. Lett.*, 46:1351–1353, 1981.
- [171] W. G. Unruh. Sonic analogue of black holes and the effect of high frequencies on black holes evaporation. *Phys. Rev.*, D51(6):2827–2838, 1995.
- [172] M. Visser. Acoustic black holes: Horizons, ergospheres, and Hawking radiation. *Class. Quantum Grav.*, 15:1767–1791, 1998.
- [173] Л. Д. Ландау and Е. М. Лифшиц. *Гидродинамика*. Наука, Москва, третье edition, 1986.
- [174] R. Arnowitt, S. Deser, and S. W. Misner. The dynamics of general relativity. In L. Witten, editor, *Gravitation: an introduction to current research*, New York – London, 1962. John Wiley & Sons, Inc. gr-qc/0405109.
- [175] B. S. DeWitt. Quantum theory of gravity. I. The canonical theory. *Phys. Rev.*, 160(5):1113–1148, 1967.
- [176] Д. М. Гитман and И. В. Тютин. *Каноническое квантование полей со связями*. Наука, Москва, 1986.
- [177] H. Weyl. *Raum – Zeit – Materie*. Springer, Berlin, 1918. Перевод: Г. Вейль. *Пространство, время, материя*. М.: Янус, 1996.
- [178] P. A. M. Dirac. *General Theory of Relativity*. John Wiley & Sons, Inc., New York – London, 1975. Перевод: *Общая теория относительности*. М.: Атомиздат, 1978.
- [179] W. Pauli. *Theory of Relativity*. Pergamon Press, New York, 1958. Перевод: Паули В. *Теория относительности*. М.: Наука, 1991.
- [180] E. Schrödinger. *Space-time Structure*. Cambridge U.P., Cambridge, 1950. Перевод: Шредингер Э. *Пространственно-временная структура Вселенной*. М.: Наука, 1986.
- [181] W. Gordon. Der Comptoneffekt nach der Schrödingerschen Theorie. *Zs. f. Phys.*, 40(1,2):117–133, 1926.
- [182] V. A. Fock. Zur Schrödingerschen Wellenmechanik. *Zs. f. Phys.*, 38(3):242–250, 1926.
- [183] V. A. Fock. über die invariante Form der Wellen- und der Bewegungs- gleichungen für einen geladenen Massenpunkt. *Zs. f. Phys.*, 39(2,3):226–232, 1926.
- [184] O. Klein. Elektrodynamik und Wellenmechanik vom Standpunkt des Korrespondenzprinzips. *Zs. f. Phys.*, 41(10):407–442, 1927.
- [185] J. Goldstone. Field theories with “superconductor” solutions. *Nuovo Cim.*, 19(1):154–164, 1961.

- [186] E. P. Wigner. *Group Theory and its application to the quantum mechanics of atomic spectra*. Academic Press, New York – London, 1959. Перевод: Вейль Е. *Теория групп и ее приложения к квантовомеханической теории атомных спектров*. М.: Иностранная литература, 1961.
- [187] J. van Bladel. Lorenz or Lorentz ? *IEEE Antennas and Propagation Magazin*, 33(2):69, 1991.
- [188] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13(9):321–323, 1964.
- [189] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12(2):132–133, 1964.
- [190] A. Proca. Sur la théorie ondulatoire des électrons positifs et négatifs. *J. Physique*, 8:347–353, 1936.
- [191] A. Proca. Sur la théorie du positon. *C. R. Acad. Sci. Paris*, 202:1366, 1936.
- [192] А. Н. Тихонов and А. А. Самарский. *Уравнения математической физики*. 5 изд. Наука, Москва, 1977.
- [193] H. B. Jr. Lawson and M-L. Michelsohn. *Spin Geometry*. Princeton University Press, Princeton, 1989.
- [194] I. M. Benn and R. W. Tucker. *An Introduction to Spinors and Geometry with Applications in Physics*. Adam Hilger, Bristol and Philadelphia, 1987.
- [195] P. Lounesto, editor. *Clifford Algebras and Spinors*. Cambridge University Press, Cambridge, 1997.
- [196] Г. И. Марчук and Д. С. Широков. *Введение в теорию алгебр Клиффорда*. Фазис, Москва, 2012.
- [197] E. Cartan. Nombres complexes. In J. Molk, editor, *"Encyclopédie des sciences mathématiques". Tome I, vol. 1, Fasc. 4, art. I5*, pages 329–468, 1908. Reprinted in E. Cartan, *Euvres Complètes*, Partie II. Gauthier–Villars, Paris, 1953, pp. 107–246.
- [198] S. Lang. *Algebra*. Addison Wesley, Reading, Mass., 1965.
- [199] E. Artin. *Geometric Algebra*. Interscience, New York, 1957. Перевод: Э. Артин. *Геометрическая алгебра*. М.: Наука, 1969.
- [200] Н. Н. Боголюбов, А. А. Логунов, А. И. Оксак, and И. Т. Тодоров. *Общие принципы квантовой теории поля*. Наука, Москва, 1987.
- [201] W. Pauli. Contributions mathematiques a la theorie des matrices de dirac. *Ann. Inst. H. Poincare*, 6:109, 1936.
- [202] Jr. R. H. Good. Properties of the dirac matrices. *Rev. Mod. Phys.*, 27:187, 1955.
- [203] M. Fierz. Zur fermischen theorie des β -zerfalls. *Z. Phys.*, 104:553–556, 1937.

- [204] R. F. Streater and A. S. Wightman. *PCT, Spin and Statistics and All That*. W. A. Benjamin, Inc., New York – Amsterdam, 1964. Перевод: Р. Стритер, А. С. Вайтман. *PCT, спин, статистика и все такое*. М.: Наука, 1966.
- [205] P. A. M. Dirac. The quantum theory of the electron. *Proc. Roy. Soc. London*, A117:610–624, 1928. Перевод в сб. П. А. М. Дирак “К созданию квантовой теории поля” Ред. Б. В. Медведев. М.: Наука, 1990, сс. 113–128.
- [206] P. A. M. Dirac. The quantum theory of the electron. part ii. *Proc. Roy. Soc. London*, A118:351–361, 1928. Перевод в сб. П. А. М. Дирак “К созданию квантовой теории поля” Ред. Б. В. Медведев. М.: Наука, 1990, сс. 129–141.
- [207] E. Majorana. Teoria simmetrica dell’elettrone e del positrone. *Nuovo Cim.*, 14:171–195, 1937.
- [208] V. A. Fock. Geometrizierung der Diracschen Theorie des Electrons. *Zs. f. Phys.*, 57(3,4):261–277, 1929. Перевод в сб.: Альберт Эйнштейн и теория гравитации. М.: Мир, 1979. С. 415–432.
- [209] H. Weyl. Gravitation and the electron. *Proc. Nat. Acad. Sci. USA*, 15:323–334, 1929.
- [210] В. С. Владимиров and В. В. Жаринов. *Уравнения математической физики*. Физ-мат литература, Лаборатория базовых знаний, Москва, 2000.
- [211] J. A. Thorpe. *Elementary Topics in Differential Geometry*. Springer-Verlag, Berlin – Heidelberg, 1979. Перевод: Дж. Торп. *Начальные главы дифференциальной геометрии*. М.: Мир, 1982.
- [212] T. K. Milnor. Efimov’s theorem about complete immersed surfaces of negative curvature. *Advances in Math.*, 8:474–543, 1972.
- [213] N. H. Kuiper. On C^1 isometric embeddings. II. *Nederl. Akad. Wetensch. Proc. Ser. A*, 58:683–689, 1955.
- [214] M. O. Katanaev. All universal coverings of two-dimensional gravity with torsion. *J. Math. Phys.*, 34(2):700–736, 1993.
- [215] L. Liouville. Sur l’équation aux différences partielles $\frac{\partial^2 \ln \lambda}{\partial u \partial v} \pm \lambda a^2 = 0$. *J. Math. Pures Appl.*, 18:71–72, 1853.
- [216] M. D. Kruskal. Maximal extension of Schwarzschild metric. *Phys. Rev.*, 119(5):1743–1745, 1960.
- [217] G. Szekeres. On the singularities of a riemannian manifold. *Publ. Mat. Debrecen*, 7(1–4):285–301, 1960.
- [218] K. Schwarzschild. Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie. *Sitzungsber. Akad. Wiss. Berlin*, pages 189–196, 1916. Перевод в сб.: Альберт Эйнштейн и теория гравитации. М.: Мир, 1979. С. 199–207.
- [219] J. Droste. Over het veld van een enkel centrum in einstein’s theorie der zwaartekracht. *Proc. Kon. Ned. Akad. Wetensch. Amsterdam*, 23:968–981, 1914–1915. English translation: “On the field of a single centre in Einstein’s theory of gravitation”. *Proc. Acad. Sci. Amsterdam* 17(1915)998–1011.

- [220] J. Droste. Het veld van twee bolvormige restunde centra in einstein's theorie der zwaartekracht. *Proc. Kon. Ned. Akad. Wetensch. Amsterdam*, 24:749–757, 1915–1916. English translation: “On the field of two Spherical Fixed Centres in Einstein's Theory of Gravitation”. *Proc. Acad. Sci. Amsterdam* 18(1916)760–769.
- [221] J. Droste. Het veld van een enkel centrum in einstein's theorie der zwaartekracht, en de beweging van een stoffelijk punt in dat veld. *Proc. Kon. Ned. Akad. Wet. Amsterdam*, 25:163–180, 1916–1917. English translation: “The Field of a Single Centre in Einstein's Theory of Gravitation, and the Motion of a Particle in That Field”. *Proc. Acad. Sci. Amsterdam* 19(1917)197–215. Reprinted, with historical comments, in *Gen. Rel. Grav.* 34(2002)1545.
- [222] M. O. Katanaev and I. V. Volovich. String model with dynamical geometry and torsion. *Phys. Lett.*, 175B(4):413–416, 1986.
- [223] T. Klösch and T. Strobl. Classical and quantum gravity in $1 + 1$ dimensions: II. The universal coverings. *Class. Quantum Grav.*, 13:2395–2421, 1996.
- [224] M. O. Katanaev, W. Kummer, and H. Liebl. Geometric interpretation and classification of global solutions in generalized dilaton gravity. *Phys. Rev. D*, 53(10):5609–5618, 1996.
- [225] M. O. Katanaev, W. Kummer, and H. Liebl. On the completeness of the black hole singularity in 2d dilaton theories. *Nucl. Phys.*, B486:353–370, 1997.
- [226] M. O. Katanaev. Global solutions in gravity: Euclidean signature. In D. Vassilevich D. Grumiller, A. Rebhan, editor, *In “Fundamental Interactions. A Memorial Volume for Wolfgang Kummer”*, pages 249–266, Singapore, 2010. World Scientific. gr-qc/0808.1559.
- [227] B. Carter. Black hole equilibrium states. In C. DeWitt and B. C. DeWitt, editors, *Black Holes*, pages 58–214, New York, 1973. Gordon & Breach.
- [228] H. Reissner. Über die Eigengravitation des elektrischen Feldes nach der Einsteinschen Theorie. *Ann. Physik (Leipzig)*, 50:106–120, 1916.
- [229] G Nordström. On the energy of the gravitational field in einstein's theory. *Proc. Kon. Ned. Akad. Wet.*, 20:1238–1245, 1918.
- [230] A. S. Eddington. A comparison of Whitehead's and Einstein's formulae. *Nature*, 113:192, 1924.
- [231] D. Finkelstein. Past-future asymmetry of the gravitational field of a point particle. *Phys. Rev.*, 110(4):965–967, 1958.
- [232] И. В. Волович and М. О. Катанаев. Квантовые струны с динамической геометрией. *Письма в ЖЭТФ*, 43(5):212–213, 1986.
- [233] M. O. Katanaev and I. V. Volovich. Two-dimensional gravity with dynamical torsion and strings. *Ann. Phys.*, 197(1):1–32, 1990.
- [234] М. О. Катанаев. Новая интегрируемая модель – двумерная гравитация с динамическим кручением. *ДАН СССР*, 309(3):591–593, 1989.

- [235] M. O. Katanaev. Complete integrability of two-dimensional gravity with dynamical torsion. *J. Math. Phys.*, 31(4):882–891, 1990.
- [236] M. O. Katanaev. Conformal invariance, extremals, and geodesics in two-dimensional gravity with torsion. *J. Math. Phys.*, 32(9):2483–2496, 1991.
- [237] M. O. Katanaev, T. Klösch, and W. Kummer. Global properties of warped solutions in general relativity. *Ann. Phys.*, 276:191–222, 1999.
- [238] M. O. Katanaev. Canonical quantization of the string with dynamical geometry and anomaly free nontrivial string in two dimensions. *Nucl. Phys. B*, 416:563–605, 1994.
- [239] W. Kummer, H. Liebl, and D. V. Vassilevich. Exact path integral quantization of generic 2D dilaton gravity. *Nucl. Phys.*, B493:491–502, 1997.
- [240] W. Kummer and G. Tieber. Universal conservation law and modified Noether symmetry in 2d models of gravity with matter. *Phys. Rev. D*, 59:044001, 1998.
- [241] M. O. Katanaev. New integrable model – two-dimensional gravity with dynamical torsion. *Sov. Phys. Dokl.*, 34(3):982–984, 1989.
- [242] W. Kummer and D. J. Schwarz. General analytic solution of R^2 -gravity with dynamical torsion in two dimensions. *Phys. Rev. D*, 45(8):3628–3635, 1992.
- [243] S. N. Solodukhin. Black-hole solution in 2d gravity with torsion. *JETP Lett.*, 57(6):329–334, 1993.
- [244] E. W. Mielke, F. Gronwald, Yu. N. Obukhov, R. Tresguerres, and F. W. Hehl. Towards complete integrability of two dimensional Poincaré gauge gravity. *Phys. Rev. D*, 48(8):3648–3662, 1993.
- [245] P. Schaller and T. Strobl. Canonical quantization of non-Einsteinian gravity and the problem of time. *Class. Quantum Grav.*, 11:331–346, 1994.
- [246] W. Kummer and P. Widerin. Conserved quasilocal quantities and general covariant theories in two dimensions. *Phys. Rev.*, D52(12):6965–6975, 1995.
- [247] B. M. Barbashov, V. V. Nesterenko, and A. M. Chervjakov. Solitons in some geometrical field theories. *Theor. Math. Phys.*, 40(1):15–27, 1979.
- [248] R. Jackiw. Liouville field theory: a two-dimensional model for gravity. In S. Christensen, editor, *Quantum theory of gravity*, pages 403–420, Bristol, 1984. Hilger.
- [249] C. Teitelboim. Gravitation and Hamiltonian structure in two spacetime dimensions. *Phys. Lett.*, 126B(1,2):41–45, 1983.
- [250] N. Ikeda and K.-J. Izawa. General form of dilaton gravity and nonlinear gauge theory. *Prog. Theor. Phys.*, 89(5):237–246, 1993.
- [251] P. Schaller and T. Strobl. Poisson structure induced (topological) field theories. *Mod. Phys. Lett.*, A9(33):3129–3136, 1994.

- [252] D. Kramer, H. Stephani, M. MacCallum, and E. Herlt. *Exact Solutions of the Einsteins Field Equations*. Deutscher Verlag der Wissenschaften, Berlin, 1980.
- [253] G. D. Birkhoff. *Relativity and modern physics*. Cambridge, Harvard University Press, Cambridge, 1923.
- [254] J. T. Jebsen. Über die allgemeinen kugelsymmetrischen losungen der einsteinschen gravitationsgleichungen im vakuum. *Ark. Mat. Ast. Fys.*, 15(18):1–9, 1921.
- [255] F. Kottler. Über die physikalischen Grundlagen der Einsteinschen Gravitationstheorie. *Ann. Physik (Leipzig), ser. 4*, 56(14):401–462, 1918.
- [256] A. Einstein and N Rosen. The particle problem in the general theory of relativity. *Phys. Rev.*, 43:73–77, 1935.
- [257] L. Flamm. Beiträge zur Einsteinschen Gravitationstheorie. *Physik. Z.*, 17:448, 1916.
- [258] P. Painlevé. La mécanique classique el la theorie de la relativité. *C. R. Acad. Sci. (Paris)*, 173:677–680, 1921.
- [259] A. Gullstrand. Allegemeine lösung des statischen einkörper-problems in der Einsteinschen gravitations theorie. *Arkiv. Mat. Astron. Fys.*, 16(8):1–15, 1922.
- [260] M. O. Katanaev. Point massive particle in general relativity. *Gen. Rel. Grav.*, 45(10):1861–1875, 2013. DOI:10.1007/s10714-013-1564-3 [arXiv:1207.3481].
- [261] И. М. Гельфанд and Г. Е. Шилов. *Обобщенные функции и действия над ними*. Физ.-мат. лит., Москва, 1958.
- [262] E. Kasner. The impossibility of Einstein fields immersed in flat space of five dimensions. *Am. J. Math.*, 43(2):126–129, 1921.
- [263] С. А. Пастон and А. А. Шейкин. Вложения для решений уравнений Эйнштейна. *ТМФ*, 175(3):429–441, 2013.
- [264] C. Fronsdal. Completion and embedding of the schwarzschild solution. *Phys. Rev.*, 116(3):778–781, 1959.
- [265] E. Kasner. Finite representation of the solar gravitational field if flat space of six dimensions. *Am. J. Math.*, 43(2):130–133, 1921.
- [266] T. Fujitani, M. Ikeda, and M. Matsumoto. On the embedding of the Schwarzschild. *J. Math. Kyoto Univ.*, 1(1):43–61, 1961.
- [267] M. Visser. Heuristic approach to the Schwarzschild geometry. *Int. J. Mod. Phys.*, D14:2051–2068, 2005.
- [268] P. Painlevé. La mecánique classique el la theorie de la relavité. *C. R. Acad. Sci. (Paris)*, 173:677–680, 1921.
- [269] G. Lemaitre. L´Univers en expansion. *Ann. Soc. Sci. (Bruxelles)*, A53:51–85, 1933. English translation: *The Expanding Universe*. *Gen. Rel. Grav.* 29(1997)641–680.

- [270] R. J. Michell. On the means of discovering the distance, magnitude, etc. of the fixed stars, in consequence of the diminution of the velocity of their light, in case such a diminution should be found to take place in any of them, and such other data should be procured from observations, as would be farther necessary for that purpose. *Philosophical Transactions of the Royal Society of London*, 74:35–57, 1784.
- [271] P. S. M. de Laplace. Exposition du systeme du monde. 1796.
- [272] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space–Time*. Cambridge University Press, Cambridge, 1973. Перевод: С. Хокинг, Дж. Эллис. *Крупномасштабная структура пространства-времени*. М.: Мир, 1977.
- [273] D. Lynden-Bell. Why do disks form jets? In J. H. Knapen, J. E. Beckman, I. Schlosman, and T. J. Mahoney, editors, *The central kilo-parsec of starbursts and AGN: the La Palma Connection, ASP conference series*, **249**, 2001. arXiv:astro-ph/0203480.
- [274] A. K. Raychaudhuri. Relativistic cosmology. I. *Phys. Rev.*, 98(4):1123–1126, 1955. Reprinted, with historical comments: *Gen. Rel. Grav.* 32(2000) 743.
- [275] A. K. Raychaudhuri. Singular state in relativistic cosmology. *Phys. Rev.*, 106:172, 1957.
- [276] J. Ehlers. Beiträge zur relativistischen Mechanik kontinuierlicher Medien [contributions to the relativistic mechanics of continuous media]. *Abhandlungen der Mathematisch-Naturwissenschaftlichen Klasse der Akademie der Wissenschaften und Literatur Mainz*, 11, 1961. English translation: *Gen. Rel. Grav.* 25, no 12, (1993)1225–1266.
- [277] A. A. Penzias and R. W. Wilson. A measurement of excess antenna temperature at 4080 mc/s. *Astrophys. J.*, 142(3):419–421, 1965.
- [278] A. Friedmann. Über die Krümmung des Raumes. *Zs. Phys.*, 10:377–386, 1922.
- [279] A. Friedmann. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zs. Phys.*, 21:326–332, 1924.
- [280] S. Perlmutter and et al. Measurements of ω and λ from 42 high-redshift supernovae. *Astrophys. J.*, 517(2):565–586, 1999.
- [281] A. G. Riess and et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116(3):1009–1038, 1998.
- [282] J. H. Jeans. The motions of stars in a kapteyn-universe. *MNRAS*, 82:122–132, 1922.
- [283] J. C. Kapteyn. First attempt at a theory of the arrangement and motion of the sidereal system. *Astrophys. J.*, 55:302–328, 1922.
- [284] A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsber. preuss. Akad. Wiss.*, 1:142–152, 1917. Перевод: Альберт Эйнштейн. Собрание научных трудов. Т. 1. М.: Наука, 1965. С. 601–612.
- [285] P. Hájíček. *An Introduction to the Relativistic Theory of Gravitation*. Springer, Heidelberg, 2008.

- [286] H. P. Robertson. Relativistic cosmology. *Rev. Mod. Phys.*, 5:62–90, 1933.
- [287] E. Kasner. Geometrical theorems on Einstein's cosmological equations. *Am. J. Math.*, 43(4):217–221, 1921.
- [288] K. Kondo. On the geometrical and physical foundations of the theory of yielding. In *Proc. 2nd Japan Nat. Congr. Applied Mechanics*, pages 41–47, Tokyo, 1952.
- [289] J. F. Nye. Some geometrical relations in dislocated media. *Acta Metallurgica*, 1:153, 1953.
- [290] B. A. Bilby, R. Bullough, and E. Smith. Continuous distributions of dislocations: a new application of the methods of non-Riemannian geometry. *Proc. Roy. Soc. London*, A231:263–273, 1955.
- [291] E. Kröner. *Kontinums Theories der Versetzungen und Eigenspannungen*. Spriger-Verlag, Berlin – Heidelberg, 1958.
- [292] E. Cartan. Sur une generalisation de la notion de courbure de Riemann et les espaces a torsion. *Compt. Rend. Acad. Sci. (Paris)*, 174:593–595, 1922. Перевод: Э. Картан. "Об обобщении понятия римановой кривизны и о пространствах с кручением". В сб. "Альберт Эйнштейн и теория гравитации М.: Мир, 1979, с. 535–537.
- [293] F. C. Frank. On the theory of liquid crystals. *Discussions Farad. Soc.*, 25:19–28, 1958.
- [294] M. O. Katanaev and I. V. Volovich. Theory of defects in solids and three-dimensional gravity. *Ann. Phys.*, 216(1):1–28, 1992.
- [295] М. О. Катанаев. Клиновая дислокация в геометрической теории дефектов. *ТМФ*, 135(2):338–352, 2003.
- [296] М. О. Катанаев. Топологически нетривиальные одномерные решения в модели Скирма. *ТМФ*, 138(2):193–208, 2004.
- [297] Л. Д. Ландау and Е. М. Лифшиц. *Теория упругости*. Наука, Москва, 1987.
- [298] А. М. Косевич. *Физическая механика реальных кристаллов*. Наукова думка, Киев, 1981.
- [299] В. Е. Захаров, С. В. Манаков, С. П. Новиков, and Л. П. Питаевский. *Теория солитонов: Метод обратной задачи*. Наука, Москва, 1980.
- [300] R. Rajaraman. *Solitons and Instantons in Quantum Field Theory*. North-Holland, Amsterdam, 1982. Перевод: Раджараман, Р. *Солитоны и инстантоны в квантовой теории поля*. М.: Мир, 1985.
- [301] Л. А. Тахтаджян and Л. Д. Фаддеев. *Гамильтонов подход в теории солитонов*. Наука, Москва, 1986.
- [302] W. J. Zakrzewski. *Low Dimensional Sigma Models*. Adam Hilger, Bristol – Philadelphia, 1989.

- [303] Ю. П. Рыбаков and В. И. Санюк. *Многомерные солитоны*. Издательство Российского университета дружбы народов, Москва, 2001.
- [304] Т. Н. R. Skyrme. Nonlinear field theory. *Proc. Roy. Soc. London*, A260:127–138, 1961.
- [305] Л. Д. Ландау and Е. М. Лифшиц. *Электродинамика сплошных сред*. Наука, Москва, второе edition, 1982.
- [306] E. Cosserat and F. Cosserat. *Théorie des corps déformables*. Hermann, Paris, 1909.
- [307] W. Nowacki. *Theory of asymmetric elasticity*. Pergamon Press, Oxford, 1985.
- [308] N. Sandru. On some problems of the linear theory of the asymmetric elasticity. *Int. J. Eng. Sci.*, 4(1):??, 1966.
- [309] A. Staruszkiewicz. Gravitational theory in three-dimensional space. *Acta Phys. Polon.*, 24(6(12)):735–740, 1963.
- [310] G. Clément. Field-theoretic particles in two space dimensions. *Nucl. Phys.*, B114:437–448, 1976.
- [311] S. Deser, R. Jackiw, and G. 't Hooft. Three-dimensional Einstein gravity: Dynamics of flat space. *Ann. Phys.*, 152(1):220–235, 1984.
- [312] V. Guillemin and S. Sternberg. *Geometric Asymptotics*. American Mathematical Society, Providence, Rhode Island, 1977. Перевод: Гийомин В., Стернберг С. *Геометрические асимптотики*. М.: Мир, 1981.
- [313] A. Vilenkin. Gravitational field of vacuum domain walls and strings. *Phys. Rev.*, D23(4):852–857, 1981.
- [314] A. Vilenkin and E. Shellard. *Cosmic Strings and Other Topological Defects*. Cambridge University Press, Cambridge, 1994.
- [315] S. Deser and R. Jackiw. String sources in 2+1-dimensional gravity. *Ann. Phys.*, 192:352–367, 1989.
- [316] A. S. Ebessen (Editor). *Carbon Nanotubes. Preparation and Properties*. CRC Press, Boca Raton–New York–London–Tokyo, 1997.
- [317] R. Saito, G. Dresselhaus, and M. S. Dresselhaus. *Physical Properties of Carbon Nanotubes*. Imperial College, London, 1998.
- [318] P. J. F. Harris. *Carbon Nanotubes and Related Structures: New materials for the Twenty-First Century*. Cambridge University Press, Cambridge, 1999.
- [319] E. Janke, F. Emde, and F. Lösch. *Tafeln Höherer Funktionen*. B. G. Teubner Verlagsgesellschaft, Stuttgart, sechste edition, 1960. Перевод: Янке Е., Эмде Ф., Леш Ф. *Специальные функции*. М.: Наука, 1977. 342 с.
- [320] J. N. Sneddon and D. S. Berry. *The classical Theory of Elasticity*. Springer-Verlag, Heidelberg, 1958.

- [321] M. B. Hindmarsh and T. W. B. Kibble. Cosmic strings. *Rep. Prog. Phys.*, 58:477, 1995.
- [322] H. F. Trotter. On the product of semi-groups of operators. *Proc. Amer. Math. Soc.*, 10(4):545–551, 1959.
- [323] И. С. Градштейн and И. М. Рыжик. *Таблицы интегралов, сумм, рядов и произведений*. Наука, Москва, 1971.
- [324] L. Schwartz. *Méthodes Mathématiques Pour Les Sciences Physiques*. Hermann, Paris, 1961. Перевод: Л. Шварц. *Математические методы для физических наук*. М.: Мир, 1965.

Предметный указатель

- 1-форма
 - левоинвариантная (left invariant 1-form), 360
- 1-форма (1-form), 167, 172
- 1-формы линейно независимые (linearly independent 1-forms) 208
- 2-форма неособая (nonsingular 2-form), 747
- 4-группа Клейна (Klein 4-group), 93
- 4-импульс (4-momentum), 144
- 4-скорость (4-velocity), 143
- 4-ускорение (4-acceleration), 143
- L_1 -норма (L_1 -norm), 45
- L_2 -норма (L_2 -norm), 45
- L_∞ -норма (L_∞ -norm), 45
- L_p -норма (L_p -norm), 45
- PCT -теорема (PCT theorem), 1103
- $\mathbb{G}$ -инвариантная функция ($\mathbb{G}$ -invariant function), 434
- $\mathbb{G}$ -инвариантная структура ($\mathbb{G}$ -invariant structure), 434
- $\mathbb{G}$ -инвариантные поля ($\mathbb{G}$ -invariant fields), 733
- $\mathbb{G}$ -инвариантное тензорное поле ($\mathbb{G}$ -invariant tensor field), 436
- $\mathbb{G}$ -многообразие ($\mathbb{G}$ -manifold), 424
- $\mathbb{k}$ -модуль ($\mathbb{k}$ -module), 34
- $\mathbb{k}$ -модуль унитарный (unital $\mathbb{k}$ -module), 34
- $\mathbb{k}$ -модуль унитарный (unitary $\mathbb{k}$ -module), 34
- P -произведение (P -product), 519
- δ -функция (δ -function), 1517
- γ -матрица (γ -matrix), 1086
- m связная область, 450
- r -форма горизонтальная (horizontal r -form), 503
- r -форма псевдотензорная (pseudotensorial r -form), 503
- r -форма тензорная (pseudotensorial r -form), 503
- (Псевдо)риманово симметрическое пространство 679
- АДМ параметризация метрики (ADM parametrization of metric), 919
- Абелева алгебра Ли (Abelian Lie algebra), 391
- Абелева группа (Abelian group), 5
- Аберрация света (light aberration), 141
- Аддитивная группа (additive group), 8
- Адиабатический предел (adiabatic limit), 545
- Адо теорема (Ado theorem), 419
- Аффинная геометрия (affine geometry), 312
- Аффинная группа (affine group), 67
- Аффинная связность (affine connection), 295
- Аффинная связность, согласованная с симплектической структурой 699
- Аффинный параметр (affine parameter), 597
- Аффинное локально симметрическое пространство 675
- Аффинное подпространство (affine subspace), 69
- Аффинное пространство (affine space), 66
- Аффинное симметрическое пространство (affine symmetric space), 675
- Аксиальная калибровка (axial gauge), 978
- Алфавит (alphabet), 450
- Алгебра (algebra), 41, 1022
- Алгебра Грассмана (Grassmann algebra), 1031
- Алгебра Клиффорда (Clifford algebra), 1025
- Алгебра Ли $\mathfrak{gl}(n, \mathbb{C})$ ($\mathfrak{gl}(n, \mathbb{C})$ Lie algebra), 404
- Алгебра Ли $\mathfrak{sl}(n, \mathbb{C})$ ($\mathfrak{sl}(n, \mathbb{C})$ Lie algebra), 406
- Алгебра Ли (Lie algebra), 191, 358, 391
- Алгебра Ли Симметрическая (symmetric Lie algebra), 677
- Алгебра Ли абелева (Abelian Lie algebra), 391
- Алгебра Ли компактная (compact Lie algebra), 391
- Алгебра Ли нильпотентная (nilpotent Lie algebra), 400
- Алгебра Ли присоединенная (adjoint Lie algebra), 396
- Алгебра Ли простая (simple Lie algebra), 401
- Алгебра Ли разрешимая (solvable Lie algebra), 399
- Алгебра Пуанкаре (Poincaré algebra), 113
- Алгебра ассоциативная (associative algebra), 41, 1022
- Алгебра фильтрованная (filtered algebra), 43
- Алгебра градуированная (graded algebra), 43
- Алгебра коммутативная (commutative algebra), 41, 1022
- Алгебра тензорная (tensor algebra), 41, 194
- Алгебра унитарная (unital algebra), 1022
- Алгебра внешняя (exterior algebra), 234
- Алгебраически замкнутое поле (algebraically closed field), 42, 57
- Аннулятор (annihilator), 53, 58
- Анти-де Ситтера группа (anti-de Sitter group), 1372
- Анти-де Ситтера метрика (anti-de Sitter metric), 1372
- Антиамодульная 2-форма (antiselfdual 2-form), 661
- Антидискретная топология (indiscrete topology), 19
- Антиэрмитова матрица (anti-Hermitian matrix), 1508
- Антилинейное отображение (antilinear map), 65
- Анзац (ansatz), 733
- Аргумент (argument), 71
- Асимптотически плоское пространство-время (asymptotically flat space-time), 951
- Ассоциативная алгебра (associative algebra), 41, 1022
- Ассоциированное расслоение (associated fiber bundle), 479
- Ассоциированное расслоение тривиальное (trivial associated fiber bundle), 485
- Атлас (atlas), 149
- Атлас полный (complete atlas), 149
- Автомодельная метрика (self-similar metric), 593
- Автомодельное решение (self-similar solution), 594
- Автоморфизм алгебры Ли (automorphism of Lie algebras), 394
- Автоморфизм главного расслоения (automorphism of a principal fiber bundle), 487
- Автоморфизм группы Ли (automorphism of a Lie group), 375
- Автоморфизм инволютивный (involutive automorphism), 394
- Автоморфизм расслоения (automorphism of a fiber bundle), 486
- Автоморфизм связности (connection automorphism), 522
- Автоморфизм вертикальный (vertical automorphism), 487
- Автоморфизм внешний (outer automorphism), 389
- Автоморфизм внутренний (inner automorphism), 389
- Азимутальный угол (azimuth angle), 340, 1133
- Банахово пространство (Banach space), 44
- Барицентрическая комбинация точек (barycentric combination of points), 69
- Барицентрическая система координат (barycentric coordinate system), 70
- Баротропная жидкость (borotropic fluid), 888
- База накрытия (base of a covering), 459
- База расслоения (base of a fiber bundle), 161, 469
- База топологии (base of topology), 18
- Базис дуальный (dual basis), 53
- Базис голономный (holonomic basis), 177, 334
- Базис координатный (coordinate basis), 177
- Базис неголономный (nonholonomic basis), 335
- Базис ортонормальный (orthonormal basis), 48, 1024
- Базис распределения (basis of distribution), 210

- Базис векторного пространства (basis of vector space), 32
Белая дыра (white hole), 1174
Бернулли лемниската (Bernoulli lemniscate), 205
Бернулли уравнение (Bernoulli equation), 892
Берри фаза (Berry phase), 554
Бесселя неравенство (Bessel's inequality), 49
Бетти число (Betti number), 257
Бианки тождества (Bianchi identities), 293, 339
Биекция (bijective (фр.) map, bijection), 71
Бихарактеристика (bicharacteristic), 617
Билинейный (bilinear), 39
Бинарная операция (binary operation), 5
Бине–Коши формула (Binet–Cauchy formula), 1503
Бином Ньютона (binomial formula), 1501
Биномиальный коэффициент (binomial coefficient), 1501
Бинормаль к кривой (binormal to a curve), 641
Бирхгофа теорема (Birkhoff's theorem), 1244, 1272
Бивектор (bivector), 661
Блок конформный (conformal block), 1161
Блок жорданов (Jordan box), 1507
Большое сжатие (big crunch), 1332
Большой взрыв (big bang), 1332
Бозон голдстоуновский (Goldstone boson), 967
Букет окружностей (union of circles), 445
Булево кольцо (Boolean ring), 9
Буст (boost), 113
Буст галилеев (Galilean boost), 115
Бутылка Клейна (Klein bottle), 163, 481
Бюргерса вектор (Burgers vector), 1405
Центр (center), 1022
Центр алгебры Ли (center of a Lie algebra), 401
Центр группы (group center), 8
Цепь (chain), 4
Циклическая группа (cyclic group), 6
Циклическая координата (cyclic coordinate), 742
Цилиндр (cylinder), 68, 481
Цилиндрические координаты (cylindrical coordinates), 345
Частица пробная (test particle), 836, 863
Частица свободная (free particle), 129, 132, 861
Частицы горизонт (particle horizon), 1382
Частично упорядоченное множество (partially ordered set, poset), 4
Частота (frequency), 620
Черепашьи координаты (tortoise coordinates), 1247, 1278
Черная дыра (black hole), 1174
Числа сингулярные (singular numbers), 1505
Число Бетти (Betti number), 257
Число собственное (eigenvalue), 409
Число степеней свободы (the number of degrees of freedom), 737
Чистая калибровка (pure gauge), 531, 1008
Дарбу координаты (Darboux coordinates), 250, 697
Дарбу теорема (Darboux theorem), 250
Де Ситтера метрика (de Sitter metric), 1366
Де Ситтера пространство (de Sitter space), 1367
Деформаций тензор (deformation tensor), 1403
Деформация отображения (deformation of a map), 444
Действие (action), 712
Действие Гильберта–Эйнштейна (Hilbert–Einstein action), 844
Действие эффективное (effective action), 732
Действие группы эффективное (effective action of a group), 425
Действие группы слева (left action of a group), 425
Действие группы справа (right action of a group), 424
Действие группы свободное (free action of a group), 425
Действие группы транзитивное (transitive action of a group), 427
Действие группы тривиальное (trivial action of a group), 425
Действие обобщенное (extended action), 785
Действие полное (total action), 794
Действие редуцированное (reduced action), 733
Действие слой-транзитивное (cross section transitive action), 539
Действие укороченное (truncated action), 753
Действие-угол переменные (action-angle variables), 761
Действительная ось (real axis), 11
Действия функционал (action functional), 712
Действия функция (action function), 754
Декартова система координат (Cartesian coordinate system), 14
Декартовы координаты (Cartesian coordinates), 13, 14, 315
Декартово произведение (Cartesian product), 4
Делитель нормальный (normal subgroup), 7
Детерминантное расслоение (determinant fiber bundle), 483
Девиации скорость (deviation velocity), 601
Девиации ускорение (deviation acceleration), 601
Девиации вектор (deviation vector), 601
Девиация геодезических (geodesic deviation), 602
Дыра белая (white hole), 1174
Дыра черная (black hole), 1174
Диада (diad), 334, 1122
Диагональная калибровка (diagonal gauge), 1124
Диаграмма Картера–Пенроуза (Carter–Penrose diagram), 1172
Диаграмма коммутативная (commutative diagram), 73
Диффеоморфизм (diffeomorphism), 84, 203
Диффеоморфизмов группа (group of diffeomorphisms), 84
Дифференциал ковариантный (covariant differential), 310
Дифференциал отображения (differential of a map), 201
Дифференциальная форма степени r (differential form of degree r), 238
Дифференцирование (differentiation), 174
Дифференцирование тензорной алгебры (differentiation in tensor algebra), 227
Дифференцирование в алгебре Ли (differentiation in Lie algebra), 395
Дифференцирование вдоль векторного поля (differentiation along a vector field), 174
Дифференцирование внешнее (exterior differentiation), 240
Дифференцирование внутреннее (internal differentiation), 237
Дифференцируемая функция (differentiable function), 164
Дифференцируемая кривая (differentiable curve), 15
Дифференцируемая структура (differentiable structure), 150
Дифференцируемое многообразие (differentiable manifold), 150
Дифференцируемое отображение (differentiable map), 200
Дикритический узел (critical knot), 184
Дилатация (dilatation), 117
Дилатация (dilation), 1404, 1489
Дилатон (dilaton), 1212
Дилатонное поле (dilaton field), 1235
Динамическая фаза (dynamical phase), 554
Дирака матрица (Dirac matrix), 1086
Дирака представление (Dirac representation), 1091
Дирака пуассонова структура (Dirac Poisson structure), 787
Дирака скобка (Dirac bracket), 787
Дирака спинор (Dirac spinor), 1096
Дирака уравнение (Dirac equation), 1104
Диск (disc), 17
Дисклинация (disclination), 1409
Дискретная топология (discrete topology), 18
Дискретное подмножество (discrete subset), 22
Дислокация коническая цилиндрическая (conical cylindrical dislocation), 1478
Дислокация краевая (edge dislocation), 1405
Дислокация винтовая (screw dislocation), 1405
Дисперсия (dispersion), 620
Дисторсии тензор (distortion tensor), 1406
Дивергенция (divergence), 255
Дивергенция векторного поля (divergence of a vector field), 325
Длина экстремали (length of an extremal), 608
Длина кривой (length of a curve), 15

- Длина слова (word length), 450
 Длина вектора (length of a vector), 44, 48, 270, 414
 Доминантное энергетическое условие (dominant energy condition), 889
 Доплера эффект (Doppler effect), 141
 Дополнение множества (complement of a set), 3
 Дополнение ортогональное (orthogonal complement), 53
 Допустимая калибровка (admissible gauge), 283
 Дуальная форма (dual form), 251
 Дуальный базис (dual basis), 53
 Дуальное пространство (dual space), 52
 Дуальное вращение (dual rotation), 996
 Движение (movement), 568
 Движение равноускоренное (uniformly accelerated motion), 145
 Движение собственное (proper motion), 68
 Двойственное пространство (dual space), 52
 Двуполостный гиперboloид (two-sheeted hyperboloid), 1136
 Двусторонний идеал (two-sided ideal), 43
 Единица (identity), 5
 Единица кольца (identity of a ring), 9
 Естественная топология (natural topology), 19
 Евклидова метрика (Euclidean metric), 5, 13
 Эддингтона–Финкельштейна координаты (Eddington–Finkelstein coordinates), 1182
 Эффект Доплера (Doppler effect), 141
 Эффективный гамильтониан (effective Hamiltonian), 790
 Эффективное действие (effective action), 732
 Эффективное действие группы (effective action of a group), 425
 Эйконал (eikonal), 619
 Эйконала уравнение (eikonal equation), 619
 Эйкональное приближение (eikonal approximation), 619
 Эйлера уравнение (Euler equation), 887
 Эйлера уравнения вращения твердого тела (Euler equations of rigid body rotations), 706
 Эйлера–Лагранжа уравнения (Euler–Lagrange equations), 714
 Эйлерова характеристика (Euler characteristic), 258
 Эйнштейна правило суммирования (Einstein summation rule), 13
 Эйнштейна пространство (Einstein space), 844
 Эйнштейна тензор (Einstein tensor), 842
 Эйнштейна уравнения (Einstein’s equations), 842
 Эйнштейна–Кауфман преобразование, 672
 Экспоненциал матрицы (matrix exponent), 409
 Экспоненциальное отображение (exponential map), 188, 379
 Экспонента внешняя (exterior exponent), 236
 Экстремаль (extremal), 603
 Экстремаль функционала (extremal of a functional), 713
 Экстремаль изотропная (isotropic extremal), 605
 Экстремаль полная (complete extremal), 606
 Эквивалентная норма (equivalent norm), 44
 Эквивалентные связи (equivalent constraints), 783
 Эквивалентный путь (equivalent path), 440
 Эквивалентности принцип (equivalence principle), 838
 Эквивалентность гомотопическая (homotopic equivalence), 445
 Эквивалентность конформная (conformal equivalence), 655
 Эквивариантное отображение (equivariant map), 437
 Электрическая компонента тензора Вейля 666
 Электрический заряд (electric charge), 972
 Электрический заряд квантование (electric charge quantization), 972
 Электродинамика скалярная (scalar electrodynamics), 987
 Электромагнитный потенциал (electromagnetic potential), 500
 Электромагнитное поле (electromagnetic field), 969, 973
 Электромагнитное поле напряженность (electromagnetic field strength), 969
 Электромагнитное поле потенциал (electromagnetic field potential), 969
 Элемент множества (element of a set), 1
 Элемент нулевой (zero element), 8
 Элемент объема (volume element), 1081
 Элемент обратный (inverse element), 5
 Элемент однородный (homogeneous element), 37
 Элемент площади гиперповерхности (area element of a hypersurface), 266
 Элемент сопряженный (conjugate element), 7
 Эндоморфизм (endomorphism), 389
 Энергетическое условие (energy condition), 889
 Энергии-импульса ковектор (energy-momentum covector), 725
 Энергии-импульса псевдотензор (energy-momentum pseudotensor), 948
 Энергии-импульса тензор (energy-momentum tensor), 725
 Энергии-импульса тензор материи (energy-momentum tensor of matter), 842
 Энергия кинетическая (kinetic energy), 745
 Энергия механической системы (energy of mechanical system), 741
 Энергия полная (total energy), 952
 Энергия потенциальная (potential energy), 745
 Энергия темная (dark energy), 1358
 Энтальпия (enthalpy), 891
 Эрмитова форма (Hermitian form), 51, 1023
 Эрмитова матрица (Hermitian matrix), 416, 1508
 Эрмитова матрица положительно определенная (positive definite Hermitian matrix), 416
 Эрмитово пространство (Hermitian space), 51
 Эрмитово скалярное произведение (Hermitian scalar product), 414
 Факторалгебра Ли (quotient Lie algebra), 394
 Факторгруппа (factor group, quotient group), 7
 Факторпространство (coset space), 427
 Фактортопология (quotient topology), 77
 Фаза Берри (Berry phase), 554
 Фаза динамическая (dynamical phase), 554
 Фаза геометрическая (geometric phase), 554
 Фазовый поток (phase flow), 743
 Фазовое пространство (phase space), 738
 Фазовое пространство расширенное (extended phase space), 749
 Ферма принцип (Fermat principle), 757
 Фиксирование калибровки (gauge fixing), 282, 978
 Фильтрация (filtration), 38
 Фильтрованная алгебра (filtered algebra), 43
 Финитная функция (compactly supported function), 157, 874
 Физическая степень свободы (physical degree of freedom), 790
 Физический гамильтониан (physical Hamiltonian), 790
 Физическое поле (physical field), 719
 Флаг (flag), 38
 Фламма параболоид (Flamm paraboloid), 1284
 Фокус (focus), 184
 Форма Киллинга (Killing form), 570
 Форма Киллинга–Картана (Killing–Cartan form), 367, 402
 Форма Лиувилля (Liouville form), 696, 750
 Форма Пфаффа (Pfaffian form), 167, 238
 Форма Вейля (Weyl form), 313
 Форма дуальная (dual form), 251
 Форма эрмитова (Hermitian form), 51, 1023
 Форма гармоническая (harmonic form), 256
 Форма инвариантная (invariant form), 402
 Форма каноническая левоинвариантная (left invariant canonical form), 361
 Форма коточная (coexact form), 256
 Форма козамкнутая (coclosed form), 256
 Форма кривизны (curvature form), 506
 Форма кривизны локальная (local curvature form), 292, 303, 507
 Форма кручения локальная (local torsion form), 303
 Форма квадратичная (quadratic form), 1024
 Форма линейная (linear form), 52
 Форма неоднородная (inhomogeneous form), 234

- Форма объема (volume form), 262
 Форма однородная (homogeneous form), 234
 Форма полуторалинейная (sesquilinear form), 1023
 Форма разложимая (decomposable form), 246
 Форма симплектическая (symplectic form), 695
 Форма степени r (form of degree r , r -form), 230
 Форма связности (connection form), 494
 Форма связности локальная (local connection form), 291, 303, 498
 Форма точная (exact form), 242
 Форма замкнутая (closed form), 242
 Формализм первого порядка (first-order formalism), 847, 993
 Формула Бине–Коши (Binet–Cauchy formula), 1503
 Формула Гаусса–Остроградского (Gauss–Ostrogradskii formula), 267
 Формула Гаусса–Вейнгартена (Gauss–Weingarten formula), 929
 Формула Грина (Green formula), 267
 Формула Хаусдорфа (Hausdorff formula), 1503
 Формула Маурера–Картана (Maurer–Cartan formula), 360
 Формула Ньютона–Лейбница (Newton–Leibnitz formula), 266
 Формула Сохоцкого (Sokhotski formulae), 1519
 Формула Стокса (Stokes formula), 266, 267
 Формула гомотопии основная (main homotopy formula), 244
 Формулы Френе (Frenet formulae), 641
 Формулировка гамильтонова (Hamiltonian formulation), 807
 Формулировка каноническая (canonical formulation), 807
 Формулировка лагранжева (Lagrangian formulation), 807
 Франка вектор (Frank vector), 1409, 1410
 Френе формулы (Frenet formulae), 641
 Фреше производная (Fréchet derivative), 46
 Фреше пространство (Fréchet space), 46
 Фридмана метрика (Friedmann metric), 1347
 Фробениуса теорема (Frobenius theorem), 211
 Фронт волновой (wave front), 757
 Фундаментальная группа (fundamental group), 448
 Фундаментальная область (fundamental domain), 467
 Фундаментальная последовательность (fundamental sequence), 16
 Фундаментальное представление группы $\mathbb{O}(1, 1)$ (fundamental representation of $\mathbb{O}(1, 1)$ group), 94
 Фундаментальное векторное поле (fundamental vector field), 491
 Функции координатные (coordinate functions), 165
 Функции носитель (support of a function), 157
 Функции структурные (structure functions), 700, 793
 Функционал действия (action functional), 712
 Функционал линейный (linear functional), 52
 Функциональная зависимость (functional dependence), 82
 Функция $\mathbb{G}$ -инвариантная ($\mathbb{G}$ -invariant function), 434
 Функция (function), 71
 Функция Гамильтона (Hamilton function), 739
 Функция Гамильтона (Hamiltonian function), 808
 Функция Хевисайда (Heaviside function), 1518
 Функция Казимира (Casimir function), 702
 Функция действия (action function), 754
 Функция действия укороченная (truncated action function), 756
 Функция дифференцируемая (differentiable function), 164
 Функция финитная (compactly supported function), 157, 874
 Функция хода (lapse function), 920
 Функция композиции (composition function), 352
 Функция модулярная (modular function), 383
 Функция обобщенная (generalized function, distribution), 874
 Функция основная (test function), 874
 Функция перехода (transition function), 81, 285, 474
 Функция пробная (test function), 874
 Функция сдвига (shift function), 920
 Функция склейки (sewing function), 81, 149, 474, 483
 Функция выпуклая (convex function), 735
 Функция волновая (wave function), 542
 Фурье преобразование (Fourier transformation), 1523
 Фурье ряд (Fourier series), 1522
 Галилеев буст (Galilean boost), 115
 Галилеево пространство-время (Galilean space-time), 128
 Галилея группа (Galilean group), 115
 Гамильтона функция (Hamilton function), 739
 Гамильтона функция (Hamiltonian function), 808
 Гамильтона–Якоби уравнение (Hamilton–Jacobi equation), 755
 Гамильтониан (Hamiltonian), 542, 739, 808
 Гамильтониан эффективный (effective Hamiltonian), 790
 Гамильтониан физический (physical Hamiltonian), 790
 Гамильтониан обобщенный (extended Hamiltonian), 785
 Гамильтониан полный (total Hamiltonian), 794
 Гамильтонова формулировка (Hamiltonian formulation), 807
 Гамильтонова плотность (Hamiltonian density), 808
 Гамильтонова система интегрируемая (integrable Hamiltonian system), 760
 Гамильтоново векторное поле (Hamiltonian vector field), 704
 Гармоническая форма (harmonic form), 256
 Гармонические координаты (harmonic coordinates), 621
 Гармоничности условие (harmonic condition), 622
 Гато производная (Gâteaux derivative), 47
 Гаусса–Остроградского формула (Gauss–Ostrogradskii formula), 267
 Гаусса–Петерсона–Кодацци уравнения (Gauss–Peterson–Codazzi equations), 930
 Гаусса–Вейнгартена формула (Gauss–Weingarten formula), 929
 Гауссова кривизна (Gaussian curvature), 331, 683, 1134
 Гауссова система координат (Gaussian coordinate system), 901
 Газ излучения (gas of radiation), 886
 Гейне–Бореля–Лебега теорема (Heine–Borel–Lebesgue theorem), 26
 Гейзенберга группа (Heisenberg group), 384
 Гельмгольца уравнение (Helmholtz equation), 1521
 Генератор алгебры Клиффорда (Clifford algebra generator), 1030
 Генератор группы Ли (generator of Lie group), 358
 Генератор преобразования симметрии (generator of a symmetry transformation), 722
 Геодезическая (geodesic), 595
 Геодезическая полная (complete geodesic), 600
 Геодезически выпуклое многообразие (geodesically convex manifold), 599
 Геодезические координаты (geodesic coordinates), 626
 Геометрическая фаза (geometric phase), 554
 Геометрическая оптика (geometric optic), 619
 Геометрический смысл кривизны (geometrical meaning of curvature), 320
 Геометрия Римана (Riemannian geometry), 314
 Геометрия Римана–Картана (Riemann–Cartan geometry), 313
 Геометрия Римана–Картана–Вейля (Riemann–Cartan–Weyl geometry), 313
 Геометрия Римана–Вейля (Riemann–Weyl geometry), 314
 Геометрия аффинная (affine geometry), 312
 Геометрия локально евклидова (locally Euclidean geometry), 315
 Геометрия псевдориманова (pseudo-Riemannian geometry), 314
 Гессиан (Hessian), 739
 Гидродинамика релятивистская (relativistic hydrodynamics), 888
 Гильберта–Эйнштейна действие (Hilbert–Einstein action), 844
 Гильбертово пространство (Hilbert space), 51
 Гиперболическая плоскость (hyperbolic plane), 1136
 Гиперболическая спираль (hyperbolic spiral), 840

- Гиперболический угол вращения (hyperbolic rotational angle), 91
- Гиперболюид двуполостный (two-sheeted hyperboloid), 1136
- Гиперболюид однополостный (one-sheet hyperboloid), 1141
- Гиперплоскость (hyperplane), 53
- Гиперповерхности параллельные (parallel hypersurfaces), 905
- Гиперповерхность (hypersurface), 189, 207
- Гиперповерхность изотропная (isotropic hypersurface), 911
- Гиперповерхность неизотропная (non-isotropic hypersurface), 911
- Гладкая кривая (smooth curve), 15
- Гладкое отображение (smooth map), 200
- Главная нормаль (principle normal), 641
- Главные направления поверхности (principal directions of a surface), 332
- Главное однородное пространство (principal homogeneous space), 430
- Главное расслоение (principal fiber bundle), 469
- Главное расслоение тривиальное (trivial principal fiber bundle), 471
- Глобальное решение (global solution), 840, 1155
- Глобальное сечение (global cross section), 162
- Глобальное сечение расслоения (global fiber bundle cross section), 470
- Глобальность (globality), 156
- Голая сингулярность (naked singularity), 1175, 1248
- Голдстоуна теорема (Goldstone theorem), 965
- Голдстоуновский бозон (Goldstone boson), 967
- Голономии группа (holonomy group), 516
- Голономии расслоение (holonomy fiber bundle), 528
- Голономная связь (holonomic constraint), 719
- Голономный базис (holonomic basis), 177, 334
- Голубое смещение (blue shift), 142
- Гомеоморфизм (homeomorphism), 75
- Гомоморфизм групп Ли (homomorphism of a Lie groups), 375
- Гомоморфизм индуцированный (induced homomorphism), 452
- Гомоморфизм расслоений (homomorphism of fiber bundles), 486
- Гомоморфизмом алгебр Ли (homomorphism of Lie algebras), 394
- Гомотетия (homothety), 117
- Гомотопическая эквивалентность (homotopic equivalence), 445
- Гомотопический тип (homotopic type), 445
- Гомотопия (homotopy), 444
- Гомотопия относительная (relative homotopy), 447
- Гомотопные отображения (homotopic maps), 444
- Горизонт (horizon), 1171, 1174, 1270
- Горизонт частицы (particle horizon), 1382
- Горизонт событий (event horizon), 1382
- Горизонтальная r -форма (horizontal r -form), 503
- Горизонтальная кривая (horizontal curve), 514, 526
- Горизонтальный лифт кривой (horizontal lift of a curve), 514, 526
- Горизонтальный лифт векторного поля (horizontal lift of a vector field), 496
- Горизонтальный подъем кривой (horizontal lift of a curve), 514
- Горизонтальный вектор (horizontal vector), 493
- Горизонтальное подпространство (horizontal subspace), 493, 525
- Горизонтальное пространство (horizontal space), 302
- Горизонтальное сечение (horizontal cross section), 293, 527
- Горячая материя (hot matter), 1360
- Градиент (gradient), 255
- Градиент скалярного поля (gradient of a scalar field), 325
- Градиентное преобразование (gradient transformation), 970
- Градуированная алгебра (graded algebra), 43
- Градуированное линейное пространство (graded linear space), 37
- Градуированное векторное пространство (graded vector space), 37
- График отображения (graph of a map), 71
- Граница множества (boundary of a set), 22
- Граничная точка (boundary point), 23
- Грассмана алгебра (Grassmann algebra), 1031
- Гравитационная масса (gravitational mass), 131
- Гравитационная постоянная (gravitational constant), 130, 842
- Гравитационный радиус (gravitational radius), 1270
- Гравитация унимодулярная (unimodular gravity), 860
- Грина формула (Green formula), 267
- Группа $GL(n, \mathbb{C})$ ($GL(n, \mathbb{C})$ group), 411
- Группа $GL(n, \mathbb{R})$ ($GL(n, \mathbb{R})$ group), 411
- Группа $O(n)$ ($O(n)$ group), 412
- Группа $O(p, q)$ ($O(p, q)$ group), 413
- Группа $PIN(V, Q)$ ($PIN(V, Q)$ group), 1064
- Группа $SL(n, \mathbb{C})$ ($SL(n, \mathbb{C})$ group), 412
- Группа $SL(n, \mathbb{R})$ ($SL(n, \mathbb{R})$ group), 412
- Группа $SPIN(V, Q)$ ($SPIN(V, Q)$ group), 1064
- Группа $SU(n)$ ($SU(n)$ group), 415
- Группа $U(n)$ ($U(n)$ group), 414
- Группа (group), 5
- Группа Галилея (Galilean group), 115
- Группа Галилея однородная (Homogeneous Galilean group), 115
- Группа Гейзенберга (Heisenberg group), 384
- Группа Клиффорда (Clifford group), 1065
- Группа Ли (Lie group), 348
- Группа Ли компактная (compact Lie group), 349
- Группа Ли локальная (local Lie group), 352
- Группа Ли простая (simple Lie group), 401
- Группа Ли унимодулярная (unimodular Lie group), 383
- Группа Липшица (Lipschitz group), 1065
- Группа Лоренца (Lorentz group), 109
- Группа Лоренца собственная ортохронная (proper orthochronous Lorentz group), 111
- Группа Мёбиуса (Möbius group), 7
- Группа Пуанкаре (Poincaré group), 109
- Группа абелева (Abelian group), 5
- Группа аддитивная (additive group), 8
- Группа аффинных преобразований (affine transformation group), 67
- Группа анти-де Ситтера (anti-de Sitter group), 1372
- Группа циклическая (cyclic group), 6
- Группа диффеоморфизмов (group of diffeomorphisms), 84
- Группа фундаментальная (fundamental group), 448
- Группа голономии (holonomy group), 516
- Группа голономии инфинитезимальная (infinitesimal holonomy group), 534
- Группа голономии локальная (local holonomy group), 532
- Группа голономии суженная (restricted holonomy group), 516
- Группа изотропии (isotropy group), 428
- Группа калибровочная (gauge group), 500
- Группа коммутативная (commutative group), 5
- Группа комплексная симплектическая (complex symplectic group), 692
- Группа конечная (finite group), 6
- Группа конформная (conformal group), 116
- Группа левых преобразований (group of left transformations), 425
- Группа накрывающая (covering group), 387
- Группа неабелева (non-Abelian group), 5
- Группа общих линейных преобразований (group of general linear transformations), 411
- Группа ортохорная (orthochorous group), 1059
- Группа правых преобразований (group of right transformations), 424
- Группа преобразований (transformation group), 425
- Группа преобразований однопараметрическая (one-parameter transformation group), 188
- Группа преобразований собственно разрывная (properly discontinuous transformation group), 464

- Группа псевдоунитарная (pseudounitary group), 415
 Группа симплектическая вещественная (real symplectic group), 689
 Группа скольжений (sliding group), 468
 Группа собственная ортохронная (proper orthochronous group), 1059
 Группа собственных вращений (proper rotation group), 98
 Группа специальных линейных преобразований 412
 Группа трансляций (translational group), 98, 109
 Группа вращений неоднородная (inhomogeneous rotation group), 98
 Группа вращений собственная (proper rotational group), 1059
 Групповая скорость (group velocity), 620
 Гука закон (Hooke law), 1403
 Гёльдера неравенство (Hölder inequality), 1525
 Гёльдера норма (Hölder norm), 45
 Гюйгенса принцип (Huygens principle), 615, 757
 Хаара мера (Haar measure), 383
 Хаббла параметр (Hubble parameter), 1353
 Хаббла закон (Hubble's law), 1353
 Характеристическая система r -формы (characteristic system for r -form), 248
 Характеристический многочлен (characteristic polynomial), 667
 Характеристический полином (characteristic polynomial), 56
 Характеристическое подпространство r -формы (characteristic subspace for r -form), 248
 Характеристическое уравнение (secular equation), 409
 Характеристика 1-формы (characteristic of 1-form), 748
 Характеристика (characteristic), 179, 615
 Характеристика эйлера (Euler characteristic), 258
 Характеристика поля (characteristic of a field), 12
 Характеристика уравнения Гамильтона–Якоби (characteristic of the Hamilton–Jacobi equation), 755
 Хаусдорфа формула (Hausdorff formula), 1503
 Хаусдорфово пространство (Hausdorff space), 28
 Хевисайда функция (Heaviside function), 1518
 Хиггса механизм (Higgs mechanism), 990
 Хода функция (lapse function), 920
 Ходжа оператор (Hodge operator), 251
 Ходжа разложение (Hodge decomposition), 257
 Холодная материя (cold matter), 1360
 Хопфа окружность (Hopf circle), 476
 Хопфа поток (Hopf flow), 476
 Хопфа расслоение (Hopf fiber bundle), 476
 Хопфа–Ринова теорема (Hopf–Rinow theorem), 638
 Хордова метрика (chord metric), 12
 Хронологическое произведение (chronological product), 519
 Идеал (ideal), 43, 393
 Идеал двусторонний (two-sided ideal), 43
 Идеал левый (left ideal), 43
 Идеал правый (right ideal), 43
 Идеальная жидкость (ideal fluid), 891
 Импульс (momentum), 144
 Импульс обобщенный (generalized momentum), 808
 Индекс контравариантный (contravariant index), 54, 192
 Индекс ковариантный (covariant index), 54, 192
 Индекс немой (damn index), 14
 Индекс свободный (free index), 14
 Индикатриса (indicatrix), 757
 Индуцированная пуассонова структура (induced Poisson structure), 706
 Индуцированная связность (induced connection), 295, 523
 Индуцированная топология (induced topology), 20
 Индуцированный гомоморфизм (induced homomorphism), 452
 Индуцированное расслоение (induced fiber bundle), 490
 Инерциальная система координат (inertial coordinate system), 106, 128, 132
 Инерциальная система отсчета (inertial coordinate system), 89
 Инерциальные координаты (inertial coordinates), 115
 Инертная масса (inertial mass), 130
 Инфинитезимальная группа голономии (infinitesimal holonomy group), 534
 Инфинитезимальная изометрия (infinitesimal isometry), 570
 Инфляция (inflation), 1353
 Инъекция расслоения (injection of a fiber bundle), 487
 Инъективное отображение (injective (fp.) map), 71
 Интеграл (integral), 259
 Интеграл Кристоффеля–Шварца (Christoffel–Schwarz integral), 1436
 Интеграл движения (integral of motion), 742
 Интеграл первый (first integral), 760
 Интеграл по многообразию (integral over a manifold), 260
 Интегральная кривая (integral curve), 181
 Интегральный инвариант (integral invariant), 765
 Интегральный инвариант Пуанкаре относительный (relative integral Poincaré invariant), 696
 Интегральный инвариант Пуанкаре–Картана (Poincaré–Cartan integral invariant), 749
 Интегральное многообразие (integral manifold), 210
 Интегральное многообразие максимальное (maximal integral manifold), 210
 Интегрируемая гамильтонова система (integrable Hamiltonian system), 760
 Интегрируемая связность (integrable connection), 328
 Интервал (interval), 13, 17, 271
 Инвариант интегральный (integral invariant), 765
 Инвариант топологический (topological invariant), 77
 Инвариантная форма (invariant form), 402
 Инвариантная подгруппа (invariant subgroup), 7
 Инвариантная связность (invariant connection), 522
 Инвариантное подпространство (invariant subspace), 56
 Инвариантное тензорное поле (invariant tensor field), 229
 Инвариантность края (invariance of a boundary), 76
 Инвариантность размерности (invariance of dimensionality), 76
 Инверсия пространства (space inversion), 1100
 Инволюция (involution), 760, 793, 1054
 Инволюция градуировки (grade involution), 1054
 Инволютивный автоморфизм (involutive automorphism), 394
 Инволютивное преобразование (involutive transformation), 736
 Инволютивное распределение (involutive distribution), 210
 Излучение (radiation), 837, 886
 Измельчение (refinement), 23
 Изолированная точка (isolated point), 23
 Изометрическое вложение (isometric embedding), 280
 Изометрия (isometry), 568
 Изоморфизм алгебр Ли (isomorphism of Lie algebras), 394
 Изоморфизм групп Ли (isomorphism of a Lie groups), 375
 Изоморфизм локальных групп Ли (isomorphism of local Lie groups), 352
 Изоморфизм расслоений (isomorphism of fiber bundles), 486
 Изотермические координаты (isothermal coordinates), 1125
 Изотопическое пространство (isotopic space), 1010
 Изотопическое вращение (isotopic rotation), 1010
 Изотропии группа (isotropy group), 428
 Изотропии подгруппа (isotropy subgroup), 538
 Изотропная экстремаль (isotropic extremal), 605
 Изотропная гиперповерхность (isotropic hypersurface), 911
 Изотропные координаты (isotropic coordinates), 1287
 Изотропное пространство (isotropic space), 578
 Изотропное векторное поле (isotropic vector field), 275
 Калибровка (gauge), 282
 Калибровка Лоренца (Lorenz gauge), 978
 Калибровка аксиальная (axial gauge), 978
 Калибровка чистая (pure gauge), 531, 1008
 Калибровка диагональная (diagonal gauge), 1124
 Калибровка допустимая (admissible gauge), 283
 Калибровка каноническая (canonical gauge), 797
 Калибровка конформная (conformal gauge), 1125
 Калибровка кулоновская (Coulomb gauge), 978
 Калибровка радиационная (radiative gauge), 882

- Калибровка светового конуса (light cone gauge), 825, 911
 Калибровка временн́ая (temporal gauge), 821, 901, 978
 Калибровочная группа (gauge group), 500
 Калибровочная модель (gauge model), 796
 Калибровочное поле (gauge field), 500, 1005
 Калибровочное преобразование (gauge transformation), 500, 728, 796
 Калибровочное условие (gauge condition), 797
 Каноническая формулировка (canonical formulation), 807
 Каноническая инвариантная связность (canonical invariant connection), 540
 Каноническая калибровка (canonical gauge), 797
 Каноническая комплексная структура (canonical complex structure), 62
 Каноническая левоинвариантная форма (left invariant canonical form), 361
 Каноническая линейная форма (canonical linear form), 696
 Каноническая ориентация (canonical orientation), 262
 Каноническая ориентация края (canonical orientation of a boundary), 160
 Каноническая плоская связность (canonical flat connection), 529
 Каноническая пуассонова структура (canonical Poisson structure), 704
 Каноническая симплектическая форма (canonical symplectic form), 689
 Канонический параметр (canonical parameter), 15, 597, 604
 Канонический тензор энергии-импульса (canonical energy-momentum tensor), 725
 Каноническое билинейное отображение (canonical bilinear map), 40
 Каноническое преобразование (canonical transformation), 706
 Карта (chart), 149
 Картана переменные (Cartan variables), 304, 334
 Картана структурное уравнение (E. Cartan's structure equation), 507
 Картана теорема (Cartan theorem), 386
 Картана–Ботта периодичность (Cartan–Bott periodicity), 1047
 Картана–Дьюдоне теорема (Cartan–Dieudonné theorem), 1068
 Картера–Пенроуза диаграмма (Carter–Penrose diagram), 1172
 Касательный вектор (tangent vector), 170, 181
 Касательное пространство (tangent space), 170
 Касательное расслоение (tangent fiber bundle), 170, 482
 Казимира функция (Casimir function), 702
 Казнера метрика (Kasner metric), 1398
 Киллинга траектория (Killing trajectory), 570
 Киллинга уравнение (Killing equation), 570
 Киллинга векторное поле (Killing vector field), 570
 Киллинга–Картана форма (Killing–Cartan form), 367, 402
 Кинетическая энергия (kinetic energy), 745
 Киральная симметрия (chiral symmetry), 1109
 Киральный ток (chiral current), 1110
 Киральное представление (Chiral representation), 1087
 Класс r -формы (class of r -form), 248
 Класс ориентации (class of orientation), 453
 Клейна 4-группа (Klein 4-group), 93
 Клейна бутылка (Klein bottle), 163, 481
 Клейна–Гордона–Фока уравнение (Klein–Gordon–Fock equation), 955
 Клетка жорданова (Jordan box), 1507
 Клиффорда алгебра (Clifford algebra), 1025
 Клиффорда группа (Clifford group), 1065
 Клиффорда сопряжение (Clifford conjugation), 1054
 Коцикл склеиваний (sewing cocycle), 474
 Коэффициент Ламе (Lame coefficient), 1403
 Коэффициент биномиальный (binomial coefficient), 1501
 Кокасательное пространство (cotangent space), 173
 Кокасательное расслоение (cotangent fiber bundle), 172
 Кокасательное векторное поле (cotangent vector field), 172
 Кокасательное векторное пространство (cotangent vector space), 172
 Кокручения тензор (contorsion tensor), 313
 Кольцо (ring), 9
 Кольцо Ли (Lie ring), 191
 Кольцо булево (Boolean ring), 9
 Кольцо коммутативное (commutative ring), 9
 Кольцо рациональное (rational ring), 10
 Кольцо с единицей (ring with identity), 9
 Коммутативная алгебра (commutative algebra), 41, 1022
 Коммутативная диаграмма (commutative diagram), 73
 Коммутативная группа (commutative group), 5
 Коммутативное кольцо (commutative ring), 9
 Коммутатор (commutator), 42
 Коммутатор векторных полей (commutator of vector fields), 190
 Компакт (compact set), 24
 Компактификация (compactification), 78
 Компактная алгебра Ли (compact Lie algebra), 391
 Компактная группа Ли (compact Lie group), 349
 Компактная симплектическая группа (compact symplectic group), 693
 Компактное подмножество (compact subset), 24
 Компактное пространство (compact space), 24
 Комплексификация (complexification), 59, 63
 Комплексная плоскость (complex plane), 11
 Комплексная плоскость расширенная (extended complex plane), 11, 1135
 Комплексная проективная прямая (complex projective line), 476
 Комплексная прямая (complex line), 476
 Комплексная сфера (complex sphere), 11
 Комплексная симплектическая группа (complex symplectic group), 692
 Комплексная структура (complex structure), 61
 Комплексная структура каноническая (canonical complex structure), 62
 Комплексная структура сопряженная (conjugate complex structure), 64
 Комплексный поворот (complex rotation), 107
 Компонента топологического пространства 29
 Компоненты неголономности (nonholonomicity components), 335
 Компоненты тензорного поля (components of a tensor field), 193
 Композиции функция (composition function), 352
 Композиция (composition), 5
 Композиция отображений (map composition), 72
 Конец пути (end of a curve), 439
 Конечная группа (finite group), 6
 Конечное множество (finite set), 2
 Конечность локальная (local finiteness), 27
 Конфигурационное пространство (configuration space), 737
 Конфигурационное пространство расширенное (extended configuration space), 754
 Конформная эквивалентность (conformal equivalence), 655
 Конформная группа (conformal group), 116
 Конформная калибровка (conformal gauge), 1125
 Конформный блок (conformal block), 1161
 Конформный множитель (conformal factor), 281, 653, 1125
 Конформно евклидова метрика (conformally Euclidean metric), 281
 Конформно эквивалентные метрики (conformally equivalent metrics), 271
 Конформно минковское пространство-время 282
 Конформное преобразование координат 652
 Конформное преобразование специальное (special conformal transformation), 117
 Конформное векторное поле Киллинга (conformal Killing vector field), 653
 Конформной кривизны тензор (conformal curvature tensor), 646
 Коническая цилиндрическая дислокация (conical cylindrical dislocation), 1478

- Консервативная система (conservative system), 741
 Константы структурные (structure constants), 354
 Контракция (contraction), 114
 Контравариантный индекс (contravariant index), 54, 192
 Контравариантное векторное поле (contravariant vector field), 166, 192
 Конус световой (light cone), 90, 107
 Координата циклическая (cyclic coordinate), 742
 Координата обобщенная (generalized coordinate), 808
 Координаты Дарбу (Darboux coordinates), 250, 697
 Координаты Эддингтона–Финкельштейна (Eddington–Finkelstein coordinates), 1182
 Координаты Крускала–Секереша (Cruskal–Sekeresh coordinates), 1186
 Координаты Шварцшильда (Schwarzschild coordinates), 1159
 Координаты барицентрические (barycentric coordinates), 70
 Координаты цилиндрические (cylindrical coordinates), 345
 Координаты черепахи (tortoise coordinates), 1247, 1278
 Координаты декартовы (Cartesian coordinates), 13, 14, 315
 Координаты гармонические (harmonic coordinates), 621
 Координаты геодезические (geodesic coordinates), 626
 Координаты инерциальные (inertial coordinates), 115
 Координаты изотермические (isothermal coordinates), 1125
 Координаты изотропные (isotropic coordinates), 1287
 Координаты локальные (local coordinates), 149
 Координаты нормальные (normal coordinates), 626
 Координаты полярные (polar coordinates), 86
 Координаты преобразование (coordinate transformation), 81
 Координаты римановы (Riemann coordinates), 626
 Координаты сферические (spherical coordinates), 340
 Координаты стереографические (stereographic coordinates), 686, 1134, 1144
 Координаты светового конуса (light cone coordinates), 89, 1130
 Координатная окрестность (coordinate neighborhood), 149
 Координатная прямая (coordinate line), 13
 Координатные функции (coordinate functions), 165
 Координатный базис (coordinate basis), 177
 Координатное покрытие (coordinate covering), 149
 Коразмерность слоения (codimensionality of a foliation), 216
 Корефер (coframe), 294, 334
 Космическая скорость вторая (second cosmic velocity), 1320
 Космологическая постоянная (cosmological constant), 842
 Космологический принцип (cosmological principle), 1344
 Космологическое время (cosmological time), 1347
 Коши последовательность (Cauchy sequence), 16
 Коши–Буняковского неравенство (Cauchy–Bunyakovskii inequality), 49, 1525
 Коточная форма (coexact form), 256
 Коттона тензор (Cotton tensor), 656
 Коулмана принцип (Coleman principle), 734
 Ковариантная производная (covariant derivative), 291
 Ковариантная производная вдоль кривой (covariant derivative along a curve), 311
 Ковариантная производная вдоль векторного поля (covariant derivative along a vector field), 290
 Ковариантная производная внешняя (external covariant derivative), 504
 Ковариантный дифференциал (covariant differential), 310
 Ковариантный индекс (covariant index), 54, 192
 Ковариантное дифференцирование внешнее (external covariant differentiation), 504
 Ковариантное векторное поле (covariant vector field), 167, 192
 Ковектор энергии-импульса (energy-momentum covector), 725
 Ковектор нормальной медлительности фронта, 758
 Ковекторное поле (covector field), 167
 Козамкнутая форма (coclosed form), 256
 Краевая дислокация (edge dislocation), 1405
 Краевая точка многообразия (boundary point of a manifold), 159
 Край многообразия (boundary of a manifold), 158
 Красное смещение (red shift), 142
 Кристоффеля символы (Christoffel's symbols), 314
 Кристоффеля–Шварца интеграл (Christoffel–Schwarz integral), 1436
 Критическая точка (critical point), 713
 Кривая (curve), 15, 149
 Кривая дифференцируемая (differentiable curve), 15
 Кривая гладкая (smooth curve), 15
 Кривая горизонтальная (horizontal curve), 526
 Кривая интегральная (integral curve), 181
 Кривая кусочно дифференцируемая (piecewise differentiable curve), 15
 Кривая кусочно гладкая (piecewise smooth curve), 15
 Кривая нулевая (null curve), 108
 Кривая простая (simple curve), 15
 Кривая простая замкнутая (closed simple curve), 440
 Кривая пространственноподобная (space-like curve), 108
 Кривая светоподобная (light-like curve), 108
 Кривая времениподобная (time-like curve), 108
 Кривая замкнутая (closed curve), 15
 Кривая жорданова (Jordan curve), 440
 Кривизна (curvature), 327
 Кривизна гауссова (Gaussian curvature), 331, 683, 1134
 Кривизна гиперповерхности внешняя (external curvature of hypersurface), 928
 Кривизна кривой (curvature of a curve), 641
 Кривизна скалярная (scalar curvature), 331
 Кривизны форма (curvature form), 506
 Кривизны тензор (curvature tensor), 327
 Кривой перепараметризация (curve reparameterization), 15
 Кронекера символ (Kronecker symbol), 197
 Кротовая нора (wormhole), 1284
 Кручение (torsion), 297
 Кручение группы (group torsion), 6
 Кручение кривой (torsion of a curve), 641
 Кручения тензор (torsion tensor), 297
 Круг (circle, disk), 17
 Крускала–Секереша координаты (Cruskal–Sekeresh coordinates), 1186
 Куб (cube), 17
 Кулоновская калибровка (Coulomb gauge), 978
 Кусочно дифференцируемая кривая (piecewise differentiable curve), 15
 Кусочно гладкая кривая (piecewise smooth curve), 15
 Квадрат вектора (square of a vector), 270
 Квадратичная форма (quadratic form), 1024
 Квантование электрического заряда (electric charge quantization), 972
 Кватернион (quaternion), 1511
 Лагранжа множитель (Lagrange multiplier), 720
 Лагранжа скобка (Lagrange bracket), 768
 Лагранжева формулировка (Lagrangian formulation), 807
 Лагранжева плотность (Lagrangian density), 712, 807
 Лагранжиан (Lagrangian), 712, 807
 Ламе коэффициент (Lame coefficient), 1403
 Лапласа уравнение (Laplace equation), 1521
 Лапласа–Бельтрами оператор (Laplace–Beltrami operator), 253, 326
 Лебега лемма (Lebesgue lemma), 26
 Лемма Лебега (Lebesgue lemma), 26
 Лемма Пуанкаре (Poincaré lemma), 243
 Лемма Стокса (Stokes lemma), 748
 Лемма о перестройке (Reorganization lemma), 1089
 Лемниската Бернулли (Bernoulli lemniscate), 205
 Левый идеал (left ideal), 43
 Левый смежный класс (left coset), 7
 Левый спинор (left spinor), 1098
 Леви–Чивиты связность (Levi–Civita connection), 314
 Левое действие группы (left action of a group), 425
 Левоинвариантная 1-форма (left invariant 1-form), 360

- Левоинвариантная каноническая форма (left invariant canonical form), 361
 Левоинвариантное векторное поле (left invariant vector field), 355
 Лежандра преобразование (Legendre transformation), 735
 Ли алгебра (Lie algebra), 191, 358, 391
 Ли факторалгебра (quotient Lie algebra), 394
 Ли группа (Lie group), 348
 Ли кольцо (Lie ring), 191
 Ли подалгебра (Lie subalgebra), 393
 Ли подгруппа (Lie subgroup), 378
 Ли производная (Lie derivative), 225
 Ли–Пуассона скобка (Lie–Poisson bracket), 705
 Лифт кривой (lift of a curve), 514, 526
 Лифт векторного поля (lift of a vector field), 496
 Линейная форма (linear form), 52
 Линейная форма каноническая (canonical linear form), 696
 Линейный функционал (linear functional), 52
 Линейный оператор (linear operator), 33
 Линейно независимые 1-формы (linearly independent 1-forms), 208
 Линейно связное топологическое пространство (arcwise connected, path connected topological space), 441
 Линейно упорядоченное множество (linearly ordered set), 4
 Линейное отображение (linear map), 46
 Линейное пространство (linear space), 31
 Линейное пространство градуированное (graded linear space), 37
 Линия мировая частицы (particle world line), 861
 Линия прямая (straight line), 4, 16, 36, 69
 Линия тока (flow line), 1325
 Линия уровня (level line, level curve), 189
 Липшица группа (Lipschitz group), 1065
 Липшица условие (Lipschitz condition), 180
 Лист Мебиуса (Möbius band, Möbius strip), 159, 481
 Лист локальный (local leaf), 216
 Лист накрытия (sheet of a covering), 459
 Лист слоения (foliation leaf), 216
 Лиувилля форма (Liouville form), 696, 750
 Лиувилля теорема (Liouville's theorem), 765
 Лиувилля уравнение (Liouville equation), 1126, 1129, 1148
 Лобачевского плоскость (Lobachevsky plane), 1136
 Локальная форма кривизны (local curvature form), 292, 303, 507
 Локальная форма кручения (local torsion form), 303
 Локальная форма связности (local connection form), 291, 303, 498
 Локальная группа Ли (local Lie group), 352
 Локальная группа голономии (local holonomy group), 532
 Локальная конечность (local finiteness), 27
 Локальная модель (local model), 721
 Локальная тривиальность аффинной связности 328
 Локальные координаты (local coordinates), 149
 Локальный изоморфизм групп Ли (local isomorphism of Lie groups), 353
 Локальный лист (local leaf), 216
 Локально евклидова геометрия (locally Euclidean geometry), 315
 Локально евклидово топологическое пространство (locally Euclidean topological space), 151
 Локально компактное топологическое пространство (locally compact topological space), 27
 Локально линейно связное топологическое пространство (locally arcwise connected, locally path connected topological space), 441
 Локально связное пространство (locally connected space), 30
 Локально тривиальная связность (locally trivial connection), 328
 Локальное сечение (local cross section), 162
 Локальное сечение нулевое (zero local cross section), 496
 Локальное сечение расслоения (local fiber bundle cross section), 470
 Локальность (locality), 156
 Лоренца группа (Lorentz group), 109
 Лоренцева калибровка (Lorenz gauge), 978
 Лоренцева метрика (Lorentzian metric), 106
 Лоренцева сигнатура (Lorentzian signature), 274
 Лоренцева связность (Lorentz connection), 1113
 Лоренцево многообразие (Lorentzian manifold), 274
 Лоренцево сокращение (Lorentz contraction), 137
 Луч (ray), 758
 Магнитная компонента тензора Вейля 666
 Майораны представление (Majorana representation), 1111
 Майорановский спинор (Majorana spinor), 1110
 Максимально продолженное многообразие (maximally extended manifold), 840
 Максимальное интегральное многообразие (maximal integral manifold), 210
 Максвелла уравнения (Maxwell's equations), 993
 Масса гравитационная (gravitational mass), 131
 Масса инертная (inertial mass), 130
 Масштабный множитель (scale factor), 1346
 Материя горячая (hot matter), 1360
 Материя холодная (cold matter), 1360
 Материя пылевидная (dust matter), 886
 Материя темная (dark matter), 1358
 Матрица Дирака (Dirac matrix), 1086
 Матрица Паули (Pauli matrix), 1509
 Матрица Якоби (Jacobi matrix), 80
 Матрица антиэрмитова (anti-Hermitian matrix), 1508
 Матрица эрмитова (Hermitian matrix), 416, 1508
 Матрица невырожденная (nondegenerate matrix), 405
 Матрица ортогональная (orthogonal matrix), 84
 Матрица подобная (similar matrix), 409, 1506
 Матрица регулярная (regular matrix), 405
 Матрица сопряженная (conjugate matrix), 409
 Матрица жорданова (Jordan matrix), 1507
 Маурера–Картана формула (Maurer–Cartan formula), 360
 Мебиуса лист (Möbius band, Möbius strip), 159, 481
 Механизм Хиггса (Higgs mechanism), 990
 Мера Хаара (Haar measure), 383
 Метод факторизации (factorization method), 754
 Метрическая полнота (metric completeness), 16
 Метрическая связность (metrical connection), 313
 Метрическое пространство (metric space), 16
 Метричность (metricity), 314
 Метрика (metric), 5, 268
 Метрика Фридмана (Friedmann metric), 1347
 Метрика Казнера (Kasner metric), 1398
 Метрика Лоренца (Lorentzian metric), 106
 Метрика анти-де Ситтера (anti-de Sitter metric), 1372
 Метрика автомодельная (self-similar metric), 593
 Метрика де Ситтера (de Sitter metric), 1366
 Метрика евклидова (Euclidean metric), 5, 13
 Метрика хордовая (chord metric), 12
 Метрика конформно евклидова (conformally Euclidean metric), 281
 Метрика обратная (inverse metric), 269
 Метрика полная (complete metric), 16
 Метрика псевдориманова (pseudo-Riemannian metric), 268
 Метрика риманова (Riemannian metric), 268
 Метрика топологическая (topological metric), 16
 Метрика вейлевски евклидова (Weyl Euclidean metric), 281
 Метрики конформно эквивалентные (conformally equivalent metrics), 271
 Метрики сигнатура (signature of a metric), 273
 Метризуемое топологическое пространство (metrizable topological space), 20
 Метризуемое топологическое векторное пространство (metrizable topological vector space), 46
 Минимальная подстановка (minimal substitution), 845, 851, 971, 1010
 Минимальный многочлен (minimal polynomial), 58, 667
 Минковского неравенство (Minkowski inequality), 1525
 Минковского плоскость (Minkowskian plane), 89

- Минковского пространство (Minkowskian space), 106, 131
 Мировая линия частицы (particle world line), 861
 Мнимая единица (imaginary unit), 10
 Мнимая ось (imaginary axis), 11
 Многочлен характеристический (characteristic polynomial), 667
 Многочлен минимальный (minimal polynomial), 58, 667
 Многообразие (manifold), 147
 Многообразие дифференцируемое (differentiable manifold), 150
 Многообразие геодезически выпуклое (geodesically convex manifold), 599
 Многообразие лоренцево (Lorentzian manifold), 274
 Многообразие максимально продолженное (maximally extended manifold), 840
 Многообразие неориентируемое (nonorientable manifold), 152
 Многообразие односвязное (simply connected manifold), 449
 Многообразие ориентированное (oriented manifold), 152
 Многообразие ориентируемое (orientable manifold), 152, 453
 Многообразие параллелизуемое (parallelizable manifold), 455
 Многообразие полное (complete manifold), 840
 Многообразие псевдориманово (pseudo-Riemannian manifold), 274
 Многообразие пуассоново (Poisson manifold), 700
 Многообразие риманово (Riemannian manifold), 274
 Многообразие с краем (manifold with boundary), 158
 Многообразие с краем ориентируемое (orientable manifold with boundary), 159
 Многообразие симплектическое (symplectic manifold), 695
 Многообразие тривиальное (trivial manifold), 152
 Многообразие замкнутое (closed manifold), 160
 Многообразия ориентация (orientation of a manifold), 456
 Многоугольник жорданов (Jordan polygon), 440
 Множества непересекающиеся (nonoverlapping sets), 3
 Множество (set), 1
 Множество частично упорядоченное (partially ordered set, poset), 4
 Множество конечное (finite set), 2
 Множество линейно упорядоченное (linearly ordered set), 4
 Множество несчетное (uncountable set), 2
 Множество открытое (open set), 18
 Множество счетное (countable set), 2, 75
 Множество выпуклое (convex set), 36
 Множество всюду плотное (everywhere dense set), 27
 Множество замкнутое (closed set), 21
 Множитель Лагранжа (Lagrange multiplier), 720
 Множитель конформный (conformal factor), 281, 653, 1125
 Множитель масштабный (scale factor), 1346
 Модели нелокальные (nonlocal model), 721
 Модель калибровочная (gauge model), 796
 Модель локальная (local model), 721
 Модель релятивистская (relativistic model), 115
 Модель с высшими производными (model with higher derivatives), 721
 Модель топологическая (topological model), 800, 1229
 Модель вырожденная (degenerate model), 802, 810
 Модуль (module), 8
 Модулярная функция (modular function), 383
 Момент орбитальный (orbital momentum), 726
 Момент спиновый (spin momentum), 726
 Момент количества движения тензор (angular-momentum tensor), 726
 Моноид (monoid), 5
 Мопертюи принцип наименьшего действия (Maupertius principle of least action), 753
 Мощность (cardinality), 75
 Мёбиуса группа (Möbius group), 7
 Наблюдаемая скорость (observed velocity), 129, 143, 861
 Наблюдаемое ускорение (observed acceleration), 143
 Начало координат (origin of coordinates), 13
 Начало отсчета (origin), 4, 12
 Начало пути (beginning of a curve), 439
 Накрытие (covering), 459
 Накрытие универсальное (universal covering), 459
 Накрытия лист (sheet of a covering), 459
 Накрывающая группа (covering group), 387
 Накрывающее пространство (covering space), 459
 Направления поверхности главные (principal directions of a surface), 332
 Направленности предел (limit of a directedness), 25
 Направленность (directedness), 25
 Напряженность электромагнитного поля (electromagnetic field strength), 510, 969
 Напряженность калибровочного поля (gauge field strength), 510
 Напряженность поля Янга–Миллса (Yang–Mills field strength), 293
 Напряженность поля Янга–Миллса (Yang–Mills field strength), 510
 Нарушение симметрии спонтанное (spontaneous symmetry breaking), 965
 Неабелева группа (non-Abelian group), 5
 Нефизическая степень свободы (unphysical degree of freedom), 790
 Нефизическое поле (unphysical field), 719
 Неголономная связь (nonholonomic constraint), 719
 Неголономный базис (nonholonomic basis), 335
 Неголономности компоненты (nonholonomicity components), 335
 Неизотропная гиперповерхность (non-isotropic hypersurface), 911
 Нелокальная модель (nonlocal model), 721
 Неметричности тензор (nonmetricity tensor), 312
 Немой индекс (damn index), 14
 Неоднородная форма (inhomogeneous form), 234
 Неоднородная группа вращений (inhomogeneous rotation group), 98
 Неориентируемое многообразие (nonorientable manifold), 152
 Несобая 2-форма (nonsingular 2-form), 747
 Непересекающиеся множества (nonoverlapping sets), 3
 Неподвижная точка (fixed point), 428
 Непрерывное отображение (continuous map), 75
 Непрерывности уравнение (continuity equation), 886
 Неприводимое тензорное поле (irreducible tensor field), 195
 Непродолжаемое риманово пространство (non-extendable Riemannian space), 640
 Неравенство Бесселя (Bessel's inequality), 49
 Неравенство Гёльдера (Hölder inequality), 1525
 Неравенство Коши–Буняковского (Cauchy–Bunyakovskii inequality), 49, 1525
 Неравенство Минковского (Minkowski inequality), 1525
 Неравенство Шварца (Schwartz inequality), 49
 Неравенство Юнга (Young inequality), 736
 Неравенство треугольника (triangle inequality), 16
 Несчетное множество (uncountable set), 2
 Нефакратимое слово (reduced word), 450
 Несвязное топологическое пространство (disconnected topological space), 29
 Невырожденная матрица (nondegenerate matrix), 405
 Невырожденная особая точка (nondegenerate singular point), 182
 Невырожденное отображение (nondegenerate map), 203
 Невырожденное преобразование координат (nondegenerate coordinate transformation), 82
 Независимость функциональная (functional independence), 82
 Нильпотентная алгебра Ли (nilpotent Lie algebra), 400
 Нора кротовая (wormhole), 1284
 Норма L_1 (L_1 -norm), 45
 Норма L_2 (L_2 -norm), 45
 Норма L_∞ (L_∞ -norm), 45
 Норма L_p (L_p -norm), 45
 Норма Гёльдера (Hölder norm), 45

- Норма эквивалентная (equivalent norm), 44
 Норма отображения (norm of a map), 1067
 Норма вектора (norm of a vector), 44
 Нормальная подгруппа (normal subgroup), 7
 Нормальная система координат (normal coordinate system), 637
 Нормальные координаты (normal coordinates), 626
 Нормальный делитель (normal subgroup), 7
 Нормированное векторное пространство (normed vector space), 44
 Носитель функции (support of a function), 157
 Ньютона бином (binomial formula), 1501
 Ньютона уравнения (Newton equations), 743
 Ньютона второй закон (second Newton's law), 869
 Ньютона закон гравитации (Newton's gravitational law), 130
 Ньютона–Лейбница формула (Newton–Leibnitz formula), 266
 Ньютонов потенциал (Newton potential), 1521
 Нулевая кривая (null curve), 108
 Нулевое векторное поле (null vector field), 275, 747
 Нулевой элемент (zero element), 8
 Нулевой вектор (null vector), 30, 50
 Объединение множеств (union of sets), 3
 Объем (volume), 264
 Объем элемента (volume element), 1081
 Объемная форма (volume form), 262
 Область (domain), 29
 Область фундаментальная (fundamental domain), 467
 Область определения функции (domain of definition of a function), 72
 Область выпуклая (convex domain), 445
 Область значений функции (range of function values), 72
 Обмотка тора всюду плотная (everywhere dense winding of torus), 205
 Обобщенная функция (generalized function, distribution), 874
 Обобщенная координата (generalized coordinate), 808
 Обобщенная скорость (generalized velocity), 737, 808
 Обобщенный гамильтониан (extended Hamiltonian), 785
 Обобщенный импульс (generalized momentum), 808
 Обобщенное действие (extended action), 785
 Обращение времени (time inversion), 92, 275, 1102
 Обратная метрика (inverse metric), 269
 Обратный элемент (inverse element), 5
 Обратный путь (inverse path), 447
 Обратное отображение (inverse map), 72
 Образ множества (set image), 72
 Образующая алгебры Клиффорда (Clifford algebra generator), 1030
 Образующие алгебры Ли (generators of Lie algebra), 356
 Однопараметрическая группа преобразований (one-parameter transformation group), 188
 Однопараметрическая подгруппа (one-parameter subgroup), 379
 Однополостный гиперболоид (one-sheet hyperboloid), 1141
 Однородная форма (homogeneous form), 234
 Однородная и изотропная вселенная (homogeneous and isotropic universe), 1346
 Однородная компонента вектора (homogeneous vector component), 37
 Однородная вселенная (homogeneous universe), 1390
 Однородный элемент (homogeneous element), 37
 Однородное пространство (homogeneous space), 430, 578
 Односвязное многообразие (simply connected manifold), 449
 Одновременные события (simultaneous events), 128
 Ограничение пуассоновой структуры (restriction of a Poisson structure), 706
 Ограниченное подмножество (bounded subset), 18
 Окрестность (neighborhood), 21
 Окрестность координатная (coordinate neighborhood), 149
 Окружность Хопфа (Hopf circle), 476
 Операция бинарная (binary operation), 5
 Оператор $*$ ($*$ -operator), 251
 Оператор δ (coboundary operator), 252
 Оператор Ходжа (Hodge operator), 251
 Оператор Лапласа–Бельтрами (Laplace–Beltrami operator), 253, 326
 Оператор Якоби (Jacobi operator), 610
 Оператор четности (parity operator), 85, 92
 Оператор диагонализующий (diagonalizable operator), 56
 Оператор кограницы (coboundary operator), 252
 Оператор линейный (linear operator), 33
 Оператор обращения времени (time inversion operator), 92
 Оператор проекционный (projection operator), 37, 279
 Оператор пространственных отражений (space reflection operator), 98
 Оператор внешнего дифференцирования (operator of exterior differentiation), 168
 Орбита действия группы (orbit of a group action), 426
 Орбитальный момент (orbital momentum), 726
 Ориентации класс (class of orientation), 453
 Ориентация (orientation), 84
 Ориентация каноническая (canonical orientation), 262
 Ориентация края каноническая (canonical orientation of a boundary), 160
 Ориентация многообразия (orientation of a manifold), 456
 Ориентированное многообразие (oriented manifold), 152
 Ориентируемое многообразие (orientable manifold), 152, 453
 Ориентируемое многообразие с краем (orientable manifold with boundary), 159
 Ортогональная матрица (orthogonal matrix), 84
 Ортогональные векторы (orthogonal vectors), 47, 270, 275
 Ортогональное дополнение (orthogonal complement), 53
 Ортогональное представление (orthogonal representation), 385
 Ортохорная группа (orthochorous group), 1059
 Ортохронное преобразование (orthochronous transformation), 92
 Ортохронное преобразование Лоренца (orthochronous Lorentz transformation), 110
 Ортонормальный базис (orthonormal basis), 48, 1024
 Ортонормальный репер (orthonormal frame), 336, 489
 Основная формула гомотопии (main homotopy formula), 244
 Основная функция (test function), 874
 Особая точка векторного поля (singular point of a vector field), 182
 Ось действительная (real axis), 11
 Ось мнимая (imaginary axis), 11
 Открытая вселенная (open universe), 160, 841
 Открытое множество (open set), 18
 Открытое отображение (open map), 76
 Открытое подмногообразие (open submanifold), 207
 Открытое покрытие (open covering), 23
 Относительная гомотопия (relative homotopy), 447
 Относительный интегральный инвариант Пуанкаре (relative integral Poincaré invariant), 696, 750
 Относительно компактное подмножество (relatively compact subset), 24
 Относительности принцип (relativity principle), 132
 Отображение антилинейное (antilinear map), 65
 Отображение дифференцируемое (differentiable map), 200
 Отображение экспоненциальное (exponential map), 188, 379
 Отображение гладкое (smooth map), 200
 Отображение инъективное (injective (фр.) map), 71
 Отображение линейное (linear map), 46
 Отображение множеств (set mapping), 71
 Отображение непрерывное (continuous map), 75
 Отображение невырожденное (nondegenerate map), 203
 Отображение обратное (inverse map), 72
 Отображение открытое (open map), 76
 Отображение полулинейное (semilinear map), 65
 Отображение пуассоново (Poisson map), 706
 Отображение следования (succession map), 75

- Отображение сюръективное (surjective (фр.) map), 71
 Отображение тождественное (identity map), 71
 Отображение трансверсальное (transversal map), 209
 Отображение взаимно однозначное (one-to-one map), 71
 Отображение замкнутое (closed map), 76
 Отображение, дифференцируемое на подмножестве (map differentiable on a subset), 485
 Отображения деформация (deformation of a map), 444
 Отображения дифференциал (differential of a map), 201
 Отображения гомотопные (homotopic maps), 444
 Отображения поднятие (lift of a map), 462
 Отображения возврат (pullback), 55, 202
 Отображения, гомотопные относительно подмножества, 447
 Овеществление (realification), 59
 Параболоид Фламма (Flamm paraboloid), 1284
 Паракомпактность (paracompactness), 27
 Параллелизуемое многообразие (parallelizable manifold), 455
 Параллелограмма правило (parallelogram law), 48
 Параллельная прямая (parallel straight line), 69
 Параллельные гиперповерхности (parallel hypersurfaces), 905
 Параллельный перенос (parallel transport), 70, 294, 316, 317
 Параллельный перенос слоя (parallel transport of a fiber), 514
 Параллельное сечение (parallel cross section), 293, 527
 Параметр Хаббла (Hubble parameter), 1353
 Параметр аффинный (affine parameter), 597
 Параметр канонический (canonical parameter), 15, 597, 604
 Параметризация метрики АДМ (ADM parametrization of metric), 919
 Паули матрица (Pauli matrix), 1509
 Перехода функция (transition function), 81, 285, 474
 Переменные Картана (Cartan variables), 304, 334
 Переменные действие-угол (action-angle variables), 761
 Перенос класса ориентации (transport of a class of orientation), 454
 Перенос параллельный (parallel transport), 70, 294, 316, 317
 Перепараметризация кривой (curve reparameterization), 15
 Пересечение множеств (intersection of sets), 3
 Перестановки знак (sign of permutation), 199
 Периодичность 8 (periodicity 8), 1047
 Периодичность Картана–Ботта (Cartan–Bott periodicity), 1047
 Перпендикулярные векторы (perpendicular vectors), 270, 275
 Первая теорема Нереп (first Noether's theorem), 722
 Первая вариация (first variation), 609
 Первый интеграл (first integral), 760
 Первый закон Ньютона (first Newton's law), 130
 Первичная связь (primary constraint), 805
 Первого порядка формализм (first-order formalism), 847
 Петля (loop), 439
 Петля Вильсона (Wilson loop), 521
 Пылевидная материя (dust matter), 886
 Пыль (dust), 886
 Пфаффовая форма (Pfaffian form), 167, 238
 Пфаффова система уравнений (Pfaffian system of equations), 212
 Пфаффова система уравнений вполне интегрируемая (totally integrable Pfaffian system of equations), 213
 Пифагора теорема (Pythagorean theorem), 48
 Планарное решение (planar solution), 1253
 Плоская связность (flat connection), 529
 Плоская связность каноническая (canonical flat connection), 529
 Плоское подмножество (planar subset), 36
 Плоскость (plane), 13
 Плоскость Лобачевского (Lobachevsky plane), 1136
 Плоскость Минковского (Minkowskian plane), 89
 Плоскость гиперболическая (hyperbolic plane), 1136
 Плоскость комплексная (complex plane), 11
 Плотность гамильтонова (Hamiltonian density), 808
 Плотность лагранжева (Lagrangian density), 712, 807
 Плотность метрики (metric density), 856
 Плотность степени p (density of weight p), 166
 Плотность тензорная (tensor density), 197
 Подалгебра Ли (Lie subalgebra), 393
 Подгруппа (subgroup), 7
 Подгруппа Ли (Lie subgroup), 378
 Подгруппа Ли замкнутая (closed Lie subgroup), 378
 Подгруппа инвариантная (invariant subgroup), 7
 Подгруппа изотропии (isotropy subgroup), 538
 Подгруппа нормальная (normal subgroup), 7
 Подгруппа однопараметрическая (one-parameter subgroup), 379
 Подгруппа сопряженная (conjugate subgroup), 7
 Подмногообразие (submanifold), 204
 Подмногообразие открытое (open submanifold), 207
 Подмногообразие пуассоново (Poisson submanifold), 706
 Подмногообразие регулярное (regular submanifold), 206
 Подмногообразие замкнутое (closed submanifold), 207
 Подмногообразия трансверсальные (transversal submanifolds), 209
 Подмножество (subset), 2
 Подмножество дискретное (discrete subset), 22
 Подмножество компактное (compact subset), 24
 Подмножество ограниченное (bounded subset), 18
 Подмножество относительно компактное (relatively compact subset), 24
 Подмножество плоское (planar subset), 36
 Подмножество собственное (proper subset), 2
 Подмножество всюду плотное (everywhere dense subset), 23
 Поднятие отображения (lift of a map), 462
 Подобия преобразование (similarity transformation), 7, 1506
 Подобная матрица (similar matrix), 409, 1506
 Подпокрытие (subcovering), 23
 Подпространство аффинное (affine subspace), 69
 Подпространство горизонтальное (horizontal subspace), 493, 525
 Подпространство инвариантное (invariant subspace), 56
 Подпространство собственное (proper subspace), 56
 Подпространство вертикальное (vertical subspace), 299, 492, 525
 Подпространство, ассоциированное с r -формой, 246
 Подрасслоение (sub fiber bundle), 487
 Подстановка минимальная (minimal substitution), 845, 851, 971, 1010
 Погружение многообразия (immersion of a manifold), 204
 Поинтинга вектор (Pointing vector), 997
 Покрытие (covering), 23
 Покрытие координатное (coordinate covering), 149
 Покрытие открытое (open covering), 23
 Поле (field), 10
 Поле Янга–Миллса (Yang–Mills field), 293, 500, 1005
 Поле алгебраически замкнутое (algebraically closed field), 42, 57
 Поле дилатона (dilaton field), 1235
 Поле электромагнитное (electromagnetic field), 969, 973
 Поле физическое (physical field), 719
 Поле калибровочное (gauge field), 500, 1005
 Поле комплексных чисел (complex number field), 10
 Поле ковекторное (covector field), 167
 Поле материи (matter field), 293, 837
 Поле нефизическое (unphysical field), 719
 Поле скалярное (scalar field), 164
 Поле спинорное (spinor field), 1084
 Поле векторное (vector field), 166, 170
 Поле векторное контравариантное (contravariant vector field), 166
 Поле векторное ковариантное (covariant vector field), 167
 Полилинейность (multilinearity), 231
 Полином характеристический (characteristic polynomial), 56

- Поливектор (multivector), 199
 Полная энергия (total energy), 952
 Полная метрика (complete metric), 16
 Полный атлас (complete atlas), 149
 Полный гамильтониан (total Hamiltonian), 794
 Полный интеграл уравнения Гамильтона–Якоби (complete integral of Hamilton–Jacobi equation), 772
 Полное действие (total action), 794
 Полное многообразие (complete manifold), 840
 Полное векторное поле (complete vector field), 181
 Полностью антисимметричный тензор (totally antisymmetric tensor), 198
 Полнота экстремали (completeness of an extremal), 606
 Полнота геодезической (completeness of a geodesic), 600
 Полнота метрическая (metric completeness), 16
 Положительно определенная эрмитова матрица (positive definite Hermitian matrix), 416
 Полугеодезическая система координат (semigeodesic coordinate system), 901
 Полугруппа (semigroup), 5
 Полулинейное отображение (semilinear map), 65
 Полупрямая сумма (semidirect sum), 398
 Полупрямое произведение групп (semidirect product of groups), 389
 Полуторалинейная форма (sesquilinear form), 1023
 Полуторалинейное скалярное произведение (sesquilinear scalar product), 414
 Поляризация (polarization), 1024
 Полярные координаты (polar coordinates), 86
 Полярный угол (polar angle), 340, 345, 1133
 Поперечная составляющая векторного поля (transverse component of a vector field), 976
 Поправки постньютоновские (post-Newtonian corrections), 872
 Порождающие свободной группы (generators of a free group), 450
 Порядок группы (order of a group), 6
 Последовательность Коши (Cauchy sequence), 16
 Последовательность фундаментальная (fundamental sequence), 16
 Постньютоновские поправки (post-Newtonian corrections), 872
 Постоянная Хаббла (Hubble constant), 1353
 Постоянная гравитационная (gravitational constant), 130, 842
 Потенциал электромагнитный (electromagnetic potential), 500
 Потенциал электромагнитного поля (electromagnetic field potential), 969
 Потенциал ньютонов (Newton potential), 1521
 Потенциальная энергия (potential energy), 745
 Потенциальная составляющая векторного поля (potential component of a vector field), 976
 Поток Хопфа (Hopf flow), 476
 Поток фазовый (phase flow), 743
 Поток векторного поля (vector field flow), 187
 Поверхность (surface), 149
 Поверхность связей (constraint surface), 783
 Поворот комплексный (complex rotation), 107
 Правый идеал (right ideal), 43
 Правый смежный класс (right coset), 7
 Правый спинор (right spinor), 1098
 Правило параллелограмма (parallelogram law), 48
 Правило суммирования Эйнштейна (Einstein summation rule), 13
 Правое действие группы (right action of a group), 424
 Предел адиабатический (adiabatic limit), 545
 Предел направленности (limit of a directedness), 25
 Предел последовательности (limit of a sequence), 16
 Предельная точка (limit point), 21
 Представление Дирака (Dirac representation), 1091
 Представление Майораны (Majorana representation), 1111
 Представление Вейля (Weyl representation), 1087
 Представление алгебры Ли (representation of a Lie algebra), 394
 Представление алгебры Ли точное (exact representation of a Lie algebra), 394
 Представление группы Ли (representation of a Lie group), 376
 Представление группы Ли точное (exact representation of a Lie group), 376
 Представление киральное (Chiral representation), 1087
 Представление ортогональное (orthogonal representation), 385
 Представление присоединенное (adjoint representation), 363
 Представление присоединенное алгебры Ли (adjoint representation of a Lie algebra), 396
 Представление спинорное (spinor representation), 125, 1084
 Представление спинорное группы $\mathrm{SO}(3)$ (spinor representation of $\mathrm{SO}(3)$ group), 105
 Представление стандартное (standard representation), 1091
 Представление унитарное (unitary representation), 385
 Преобразование Эйнштейна–Кауфман, 672
 Преобразование Фурье (Fourier transformation), 1523
 Преобразование Лежандра (Legendre transformation), 735
 Преобразование Лоренца ортохронное (orthochronous Lorentz transformation), 110
 Преобразование Лоренца собственное (proper Lorentz transformation), 110
 Преобразование Вейля (Weyl transformation), 271, 653, 659
 Преобразование четности (parity transformation), 1100
 Преобразование градиентное (gradient transformation), 970
 Преобразование инволютивное (involutive transformation), 736
 Преобразование калибровочное (gauge transformation), 500, 728, 796
 Преобразование каноническое (canonical transformation), 706
 Преобразование координат (coordinate transformation), 81
 Преобразование координат активное (active coordinate transformation), 81
 Преобразование координат невырожденное (nondegenerate coordinate transformation), 82
 Преобразование координат пассивное (passive coordinate transformation), 81
 Преобразование ортохронное (orthochronous transformation), 92
 Преобразование подобия (similarity transformation), 7, 1506
 Преобразование симметрии генератор (generator of a symmetry transformation), 722
 Преобразование точечное (pointwise transformation), 746
 Преобразований группа (transformation group), 425
 Преобразования координат конформные 652
 Приближение эйконала (eikonal approximation), 619
 Приближение геометрической оптики (geometric optic approximation), 619
 Примесь (impurity), 1409
 Принцип Ферма (Fermat principle), 757
 Принцип Гюйгенса (Huygens principle), 615, 757
 Принцип Коулмана (Coleman principle), 734
 Принцип эквивалентности (equivalence principle), 838
 Принцип космологический (cosmological principle), 1344
 Принцип наименьшего действия Мопертюи (Maupertius principle of least action), 753
 Принцип относительности (relativity principle), 132
 Приращение функции (increment of a function), 221
 Присоединенная алгебра Ли (adjoint Lie algebra), 396
 Присоединенное представление (adjoint representation), 363
 Присоединенное представление алгебры Ли (adjoint representation of a Lie algebra), 396
 Пробная частица (test particle), 836, 863
 Пробная функция (test function), 874
 Продольная составляющая векторного поля (longitudinal component of a vector field), 976
 Продолжение функции (extension of a function), 72, 209
 Продолжение отображения (extension of a map), 72
 Проекционный оператор (projection operator), 37, 279

- Проекция (projection), 161, 469
 Проективное пространство (projective space), 154
 Проектор (projector), 37
 Произведение (product), 5
 Произведение декартово (Cartesian product), 4
 Произведение форм скалярное (scalar product of forms), 252
 Произведение групп полупрямое (semidirect product of groups), 389
 Произведение групп прямое (direct product of groups), 8
 Произведение хронологическое (chronological product), 519
 Произведение отображений (composition of transformations), 72
 Произведение прямое (direct product), 4
 Произведение скалярное (scalar product), 47, 51
 Произведение сплетенное (warped product), 1235
 Произведение тензорное (tensor product), 39, 54
 Произведение топологическое (topological product), 19
 Произведение упорядоченное (ordered product), 519
 Производная Ли (Lie derivative), 225
 Производная ковариантная (covariant derivative), 291
 Производная по Фреше (Frechét derivative), 46
 Производная по Гато (Gâteaux derivative), 47
 Производная сильная (strong derivative), 47
 Производная слабая (weak derivative), 47
 Производная вариационная (variational derivative), 714
 Производная вдоль векторного поля (derivative along a vector field), 311
 Производящая функция канонического преобразования (generating function of a canonical transformation), 770
 Прообраз множества (inverse image of a set), 72
 Простая алгебра Ли (simple Lie algebra), 401
 Простая группа Ли (simple Lie group), 401
 Простая кривая (simple curve), 15
 Простая замкнутая кривая (closed simple curve), 440
 Простой спектр (simple spectrum), 56
 Пространственная инверсия (space inversion), 1100
 Пространственноподобная кривая (space-like curve), 108
 Пространственноподобное векторное поле (space-like vector field), 275
 Пространство (псевдо)риманово симметрическое 679
 Пространство (space), 2, 128
 Пространство Эйнштейна (Einstein space), 844
 Пространство Фреше (Frechét space), 46
 Пространство Минковского (Minkowskian space), 106, 131
 Пространство Риччи плоское (Ricci flat space), 843
 Пространство Симметричное (symmetric space), 676
 Пространство аффинное (affine space), 66
 Пространство аффинное локально симметрическое 675
 Пространство аффинное симметрическое (affine symmetric space), 675
 Пространство банахово (Banach space), 44
 Пространство де Ситтера (de Sitter space), 1367
 Пространство дуальное (dual space), 52
 Пространство двойственное (dual space), 52
 Пространство эрмитово (Hermitian space), 51
 Пространство фазовое (phase space), 738
 Пространство гильбертово (Hilbert space), 51
 Пространство главное однородное (principal homogeneous space), 430
 Пространство горизонтальное (horizontal space), 302
 Пространство хаусдорфово (Hausdorff space), 28
 Пространство изотопическое (isotopic space), 1010
 Пространство касательное (tangent space), 170
 Пространство кокасательное (cotangent space), 173
 Пространство компактное (compact space), 24
 Пространство конфигурационное (configuration space), 737
 Пространство линейное (linear space), 31
 Пространство локально связное (locally connected space), 30
 Пространство метрическое (metric space), 16
 Пространство покрывающее (covering space), 459
 Пространство орбит (orbit space), 427
 Пространство постоянной кривизны (constant curvature space), 675
 Пространство проективное (projective space), 154
 Пространство расслоения (fiber bundle space), 161, 469
 Пространство расслоенное (fiber bundle), 161
 Пространство риманово непродолжаемое (non-extendable Riemannian space), 640
 Пространство сепарабельное (separable space), 27
 Пространство сопряженное (conjugate space), 52
 Пространство стягиваемое (contractible space), 445
 Пространство штеткелево (Stäckel space), 615
 Пространство топологическое (topological space), 18
 Пространство векторное (vector space), 31
 Пространство векторное топологическое (topological vector space), 46
 Пространство-мишень (target space), 71, 200
 Пространство-время (space-time), 275
 Пространство-время асимптотически плоское (asymptotically flat space-time), 951
 Пространство-время галилеево (Galilean space-time), 128
 Пространство-время пустое (empty space-time), 843
 Пространство-время стационарное (stationary space-time), 573
 Прямая комплексная (complex line), 476
 Прямая комплексная проективная (complex projective line), 476
 Прямая координатная (coordinate line), 13
 Прямая линия (straight line), 4, 16, 36, 69
 Прямая параллельная (parallel straight line), 69
 Прямая сумма расслоений (direct sum of fiber bundles), 287
 Прямая сумма векторных пространств (direct sum of vector spaces), 36
 Прямое произведение (direct product), 4
 Прямое произведение групп (direct product of groups), 8
 Псевдоевклидова геометрия (pseudo-Euclidean geometry), 315
 Псевдогруппа преобразований (pseudogroup of transformations), 83
 Псевдориманова геометрия (pseudo-Riemannian geometry), 314
 Псевдориманова метрика (pseudo-Riemannian metric), 268
 Псевдориманово многообразие (pseudo-Riemannian manifold), 274
 Псевдориманово векторное расслоение (pseudo-Riemannian vector bundle), 289
 Псевдосфера (pseudosphere), 684
 Псевдоскаляр (pseudoscalar), 93
 Псевдотензор (pseudotensor), 93
 Псевдотензор энергии-импульса (energy-momentum pseudotensor), 948
 Псевдотензорная r -форма (pseudotensorial r -form), 503
 Псевдоунитарная группа (pseudounitary group), 415
 Псевдовектор (pseudovector), 93
 Пуанкаре группа (Poincaré group), 109
 Пуанкаре лемма (Poincaré lemma), 243
 Пуанкаре относительный интегральный инвариант (relative integral Poincaré invariant), 750
 Пуанкаре–Картана интегральный инвариант (Poincaré–Cartan integral invariant), 749
 Пуассона скобка (Poisson bracket), 698, 700, 740
 Пуассонова сигма модель (Poisson sigma-model), 1225
 Пуассонова структура (Poisson structure), 700
 Пуассонова структура Дирака (Dirac Poisson structure), 787
 Пуассонова структура индуцированная (induced Poisson structure), 706
 Пуассонова структура каноническая (canonical Poisson structure), 704
 Пуассоново многообразие (Poisson manifold), 700
 Пуассоново отображение (Poisson map), 706
 Пуассоново подмногообразие (Poisson submanifold), 706
 Пунктирный спинор (dotted spinor), 1096
 Пустое множество (empty set), 2

- Пустое пространство-время (empty space-time), 843
 Путь (path), 15
 Путь эквивалентный (equivalent path), 440
 Путь обратный (inverse path), 447
 Путь тождественный (identical path), 440
 Путь замкнутый (closed path), 439
 Рациональное кольцо (rational ring), 10
 Радиационная калибровка (radiative gauge), 882
 Радикал алгебры Ли (radical of a Lie algebra), 401
 Радиус Шварцшильда (Schwarzschild radius), 1270
 Радиус гравитационный (gravitational radius), 1270
 Радиус-вектор (radius vector), 30
 Райчаудхури уравнение (Raychaudhuri equation), 1329
 Ранг r -формы (rank of r -form), 246
 Ранг отображения (rank of a map), 203
 Ранг пуассоновой структуры (rank of a Poisson structure), 702
 Ранг свободной группы (rank of a free group), 450
 Ранг тензорного поля (rank of a tensor field), 192
 Распределение (distribution), 210
 Распределение инволютивное (involutive distribution), 210
 Распределение векторных полей (distribution of vector fields), 211
 Распределение вполне интегрируемое (totally integrable distribution), 210
 Расслоение (fiber bundle), 161
 Расслоение Хопфа (Hopf fiber bundle), 476
 Расслоение ассоциированное (associated fiber bundle), 479
 Расслоение детерминантное (determinant fiber bundle), 483
 Расслоение главное (principal fiber bundle), 469
 Расслоение голономии (holonomy fiber bundle), 528
 Расслоение индуцированное (induced fiber bundle), 490
 Расслоение касательное (tangent fiber bundle), 170, 482
 Расслоение кокасательное (cotangent fiber bundle), 172
 Расслоение ортонормальных реперов (fiber bundle of orthonormal frames), 489
 Расслоение редуцированное (reduced fiber bundle), 487
 Расслоение реперов (frame fiber bundle), 297
 Расслоение тензорное (tensor fiber bundle), 482
 Расслоение тензоров типа (r, s) (fiber bundle of tensors of type (r, s)), 192
 Расслоение тривиальное (trivial fiber bundle), 162
 Расслоение векторное (vector fiber bundle), 285
 Расслоение векторное псевдориманово (pseudo-Riemannian vector bundle), 289
 Расслоение векторное риманово (Riemannian vector bundle), 289
 Расслоенное пространство (fiber bundle), 161
 Расстояние (distance), 16, 637
 Расширенная комплексная плоскость (extended complex plane), 11, 1135
 Расширенное фазовое пространство (extended phase space), 749
 Расширенное конфигурационное пространство (extended configuration space), 754
 Равенство слабое (weak equality), 874
 Равноускоренное движение (uniformly accelerated motion), 145
 Разбиение единицы (partition of unity), 157
 Разделение переменных (separation of variables), 778
 Разложение Ходжа (Hodge decomposition), 257
 Разложимая r -форма (decomposable r -form), 246
 Разложимый тензор (decomposable tensor), 40
 Разложимое тензорное поле (decomposable tensor field), 195
 Размерность аффинного пространства (dimensionality of an affine space), 67
 Размерность алгебры (dimensionality of an algebra), 41, 1022
 Размерность евклидова пространства (dimensionality of the Euclidean space), 13, 19
 Размерность группы Ли (dimensionality of a Lie group), 349
 Размерность многообразия (dimensionality of a manifold), 147
 Размерность слоения (dimensionality of a foliation), 216
 Размерность векторного пространства (dimensionality of vector space), 32
 Разность множеств (difference of sets), 3
 Разрешимая алгебра Ли (solvable Lie algebra), 399
 Развертка кривой, 538
 Редуцированная связность (reduced connection), 522
 Редуцированное действие (reduced action), 733
 Редуцированное расслоение (reduced fiber bundle), 487
 Редукции теорема (reduction theorem), 527
 Редукция расслоения (reduction of a fiber bundle), 487
 Регулярная матрица (regular matrix), 405
 Регулярное подмногообразие (regular submanifold), 206
 Регулярное вложение (regular embedding), 206
 Релятивистская гидродинамика (relativistic hydrodynamics), 888
 Релятивистская модель (relativistic model), 115
 Репер (frame), 289, 297, 334
 Репер ортонормальный (orthonormal frame), 336, 489
 Реперов расслоение (frame fiber bundle), 297
 Решение автомодельное (self-similar solution), 594
 Решение глобальное (global solution), 840, 1155
 Решение планарное (planar solution), 1253
 Реверс (reversion), 1054
 Риччи плоское пространство (Ricci flat space), 843
 Риччи тензор (Ricci tensor), 331
 Римана сфера (Riemann sphere), 11, 1135
 Римана–Картана геометрия (Riemann–Cartan geometry), 313
 Римана–Картана–Вейля геометрия (Riemann–Cartan–Weyl geometry), 313
 Римана–Кристоффеля тензор (Riemann–Christoffel tensor), 327
 Римана–Вейля геометрия (Riemann–Weyl geometry), 314
 Риманова геометрия (Riemannian geometry), 314
 Риманова метрика (Riemannian metric), 268
 Римановы координаты (Riemann coordinates), 626
 Риманово многообразие (Riemannian manifold), 274
 Риманово пространство непродолжаемое (non-extendable Riemannian space), 640
 Риманово векторное расслоение (Riemannian vector bundle), 289
 Росток (germ), 173
 Ротор (curl), 256
 Ряд Фурье (Fourier series), 1522
 Самодуальная 2-форма (selfdual 2-form), 661
 Счетное множество (countable set), 2, 75
 Сдвиг (translation), 66
 Сдвига функция (shift function), 920
 Сечение (cross section), 162
 Сечение глобальное (global cross section), 162
 Сечение горизонтальное (horizontal cross section), 293, 527
 Сечение локальное (local cross section), 162
 Сечение параллельное (parallel cross section), 293, 527
 Сечение расслоения (fiber bundle cross section), 470
 Сечение, параллельное вдоль кривой (section parallel along a curve), 294
 Седло (saddle point), 182
 Семейство (family), 2
 Сен–Венана условия интегрируемости (Saint–Venant integrability conditions), 1405
 Сепарабельное пространство (separable space), 27
 Сфера (sphere), 17
 Сфера Римана (Riemann sphere), 11, 1135
 Сфера комплексная (complex sphere), 11
 Сферические координаты (spherical coordinates), 340
 Сглаживание многообразия (smoothing of a manifold), 151
 Сигма модель пуассонова (Poisson sigma-model), 1225
 Сигнатура лоренцева (Lorentzian signature), 274
 Сигнатура метрики (signature of a metric), 273
 Сильная производная (strong derivative), 47

- Сильное энергетическое условие (strong energy condition), 889
- Симметрическая алгебра Ли (symmetric Lie algebra), 677
- Симметричное пространство (symmetric space), 676
- Симметрия киральная (chiral symmetry), 1109
- Симметрия в точке (symmetry at a point), 674
- Симплекс (simplex), 70
- Симплекс вырожденный (degenerate simplex), 70
- Симплекс замкнутый (closed simplex), 70
- Симплектическая форма (symplectic form), 695
- Симплектическая форма каноническая (canonical symplectic form), 689
- Симплектическая группа компактная (compact symplectic group), 693
- Симплектическая группа комплексная (complex symplectic group), 692
- Симплектическая группа вещественная (real symplectic group), 689
- Симплектическое многообразие (symplectic manifold), 695
- Символ Кронекера (Kronecker symbol), 197
- Символы Кристоффеля (Christoffel's symbols), 314
- Сингулярная точка (singular point), 840
- Сингулярные числа (singular numbers), 1505
- Сингулярность голая (naked singularity), 1175, 1248
- Синхронная система координат (synchronous coordinate system), 901
- Система консервативная (conservative system), 741
- Система координат декартова (Cartesian coordinate system), 14
- Система координат гауссова (Gaussian coordinate system), 901
- Система координат инерциальная (inertial coordinate system), 106, 128, 132
- Система координат нормальная (normal coordinate system), 637
- Система координат полугеодезическая (semigeodesic coordinate system), 901
- Система координат синхронная (synchronous coordinate system), 901
- Система координат сопутствующая (comoving coordinate system), 896, 899
- Система отсчета инерциальная (inertial coordinate system), 89
- Система уравнений пфафхова (Pfaffian system of equations), 212
- Система, ассоциированная с r -формой, 246
- Скаляр (scalar), 31
- Скаляр расширения (scalar of expansion), 1324, 1326
- Скалярная электродинамика (scalar electrodynamics), 987
- Скалярная кривизна (scalar curvature), 331
- Скалярное поле (scalar field), 164
- Скалярное поле тривиальное (trivial scalar field), 164
- Скалярное произведение (scalar product), 47, 51
- Скалярное произведение эрмитово (Hermitian scalar product), 414
- Скалярное произведение форм (scalar product of forms), 252
- Скалярное произведение полуторалинейное (sesquilinear scalar product), 414
- Скалярное произведение векторных полей (scalar product of vector fields), 268
- Склеивающий коцикл (sewing cocycle), 474
- Скобка Дирака (Dirac bracket), 787
- Скобка Лагранжа (Lagrange bracket), 768
- Скобка Ли векторных полей (Lie bracket of vector fields), 190
- Скобка Ли–Пуассона (Lie–Poisson bracket), 705
- Скобка Пуассона (Poisson bracket), 698, 700, 740
- Скольжений группа (sliding group), 468
- Скорость (velocity), 143
- Скорость девиации (deviation velocity), 601
- Скорость групповая (group velocity), 620
- Скорость наблюдаемая (observed velocity), 129, 143
- Скорость обобщенная (generalized velocity), 737, 808
- Скорость собственная (proper velocity), 861
- Слабая производная (weak derivative), 47
- Слабое энергетическое условие (weak energy condition), 889
- Слабое равенство (weak equality), 874
- След (trace), 1056
- След аффинной связности (trace of an affine connection), 324
- След тензора (trace of a tensor), 195
- Слоение (foliation), 216
- Слоить (foliate), 216
- Слой (fiber), 161, 471
- Слой типичный (typical fiber), 161
- Слой-транзитивное действие (cross section transitive action), 539
- Слово (word), 450
- Сложение (addition), 8
- Смещение голубое (blue shift), 142
- Смещение красное (red shift), 142
- Смещения вектор (displacement vector), 221
- Смежный класс (coset), 7
- Событие (event), 128
- События одновременные (simultaneous events), 128
- Собственная группа вращений (proper rotational group), 1059
- Собственная ортохронная группа (proper orthochronous group), 1059
- Собственная ортохронная группа Лоренца (proper orthochronous Lorentz group), 93, 111
- Собственная скорость (proper velocity), 861
- Собственных вращений группа (proper rotation group), 98
- Собственный вектор (eigenvector), 56
- Собственно разрывная группа преобразований (properly discontinuous transformation group), 464
- Собственное число (eigenvalue), 409
- Собственное движение (proper motion), 68
- Собственное подмножество (proper subset), 2
- Собственное подпространство (proper subspace), 56
- Собственное преобразование Лоренца (proper Lorentz transformation), 110
- Собственное время (proper time), 608, 861
- Собственное значение (eigenvalue), 1504
- Собственное значение оператора (operator eigenvalue), 56
- Сохоцкого формулы (Sokhotski formulae), 1519
- Сохраняющийся ток (conserved current), 723
- Сохраняющийся заряд (conserved charge), 724
- Сокращение лоренцево (Lorentz contraction), 137
- Соленоидальная составляющая векторного поля (solenoid component of a vector field), 976
- Соотношение Шалля (Chasles relation), 66
- Сопряжение клиффордово (Clifford conjugation), 1054
- Сопряжение по Дираку (Dirac conjugation), 1098
- Сопряжение в алгебре Ли (conjugation in Lie algebra), 395
- Сопряжение зарядовое (charge conjugation), 1101
- Сопряженная комплексная структура (conjugate complex structure), 64
- Сопряженная матрица (conjugate matrix), 409
- Сопряженная подгруппа (conjugate subgroup), 7
- Сопряженные точки (conjugate points), 611
- Сопряженный элемент (conjugate element), 7
- Сопряженный спинор (conjugate spinor), 1098
- Сопряженное комплексное векторное пространство (conjugate complex vector space), 64
- Сопряженное направление, 759
- Сопряженное пространство (conjugate space), 52
- Сопутствующая система координат (comoving coordinate system), 896, 899
- Специальное конформное преобразование (special conformal transformation), 117
- Спектр (spectrum), 56
- Спектр простой (simple spectrum), 56
- Спинор (spinor), 1084
- Спинор Дирака (Dirac spinor), 1096
- Спинор Вейля (Weyl spinor), 1098

- Спинор левый (left spinor), 1098
 Спинор майорановский (Majorana spinor), 1110
 Спинор правый (right spinor), 1098
 Спинор пунктирный (dotted spinor), 1096
 Спинор сопряженный (conjugate spinor), 1098
 Спинорный ток (spinor current), 1105
 Спинорное поле (spinor field), 1084
 Спинорное представление (spinor representation), 125, 1084
 Спинорное представление группы $\mathbb{SO}(3)$ (spinor representation of $\mathbb{SO}(3)$ group), 105
 Спиновая структура (spin structure), 1410
 Спиновая связность (Spin connection), 1113
 Спиновый момент (spin momentum), 726
 Спираль гиперболическая (hyperbolic spiral), 840
 Спиральность (helicity), 883
 Сплетенное произведение (warped product), 1235
 Спонтанное нарушение симметрии (spontaneous symmetry breaking), 965
 Стационарная точка (fixed point), 428
 Стационарная точка (stationary point), 713
 Стационарное пространство-время (stationary space-time), 573
 Стандартное представление (standard representation), 1091
 Степень свободы физическая (physical degree of freedom), 790
 Степень свободы нефизическая (unphysical degree of freedom), 790
 Степень внешняя (external power), 236
 Стереографические координаты (stereographic coordinates), 686, 1134, 1144
 Стокса формула (Stokes formula), 266, 267
 Стокса лемма (Stokes lemma), 748
 Структура $\mathbb{G}$ -инвариантная ($\mathbb{G}$ -invariant structure), 434
 Структура дифференцируемая (differentiable structure), 150
 Структура комплексная (complex structure), 61
 Структура пуассона (Poisson structure), 700
 Структура спиновая (spin structure), 1410
 Структурные функции (structure functions), 700, 793
 Структурные константы (structure constants), 354
 Структурные уравнения (structure equations), 301
 Структурное уравнение Картана (E. Cartan's structure equation), 507
 Ступенька (step function), 1518
 Стягиваемое пространство (contractible space), 445
 Субмерсия (submersion), 219
 Сумма полупрямая (semidirect sum), 398
 Сумма прямая векторных пространств (direct sum of vector spaces), 36
 Суперпозиция отображений (composition of transformations), 72
 Сужение функции (restriction of a function), 72, 209
 Сужение отображения (restriction of map), 72
 Суженная группа голономии (restricted holonomy group), 516
 Свертка тензоров (contraction of tensors), 195
 Светоподобная кривая (light-like curve), 108
 Светоподобный вектор (light-like vector), 50
 Светоподобное векторное поле (light-like vector field), 275
 Светового конуса калибровка (light cone gauge), 825, 911
 Светового конуса координаты (light cone coordinates), 89, 1130
 Световой конус (light cone), 90, 107
 Световой конус будущего (future light cone), 107
 Световой конус прошлого (past light cone), 107
 Свободная частица (free particle), 129, 132, 861
 Свободная группа (free group), 450
 Свободный индекс (free index), 14
 Свободное действие группы (free action of a group), 425
 Свойство топологическое (topological property), 77
 Связей поверхность (constraint surface), 783
 Связи эквивалентные (equivalent constraints), 783
 Связи первого рода (first class constraints), 793
 Связи второго рода (second class constraints), 784
 Связное топологическое пространство (connected topological space), 29
 Связности форма (connection form), 494
 Связность Леви-Чивиты (Levi-Civita connection), 314
 Связность аффинная (affine connection), 295
 Связность индуцированная (induced connection), 295, 523
 Связность интегрируемая (integrable connection), 328
 Связность инвариантная (invariant connection), 522
 Связность каноническая инвариантная (canonical invariant connection), 540
 Связность каноническая плоская (canonical flat connection), 529
 Связность локально тривиальная (locally trivial connection), 328
 Связность метрическая (metrical connection), 313
 Связность на главном расслоении (connection on a principal fiber bundle), 493
 Связность на векторном расслоении (connection on a vector bundle), 289
 Связность плоская (flat connection), 529
 Связность редуцированная (reduced connection), 522
 Связь (constraint), 719, 782
 Связь голономная (holonomic constraint), 719
 Связь неголономная (nonholonomic constraint), 719
 Связь первичная (primary constraint), 805
 Связь вторичная (secondary constraint), 806
 Сюръективное отображение (surjective (fp.) map), 71
 Сжатие большой (big crunch), 1332
 Шаля соотношение (Chasles relation), 66
 Шар (ball), 17, 638
 Шредингера уравнение (Schrödinger equation), 542
 Штекелево пространство (Stäckel space), 615
 Шварца неравенство (Schwartz inequality), 49
 Шварцшильда координаты (Schwarzschild coordinates), 1159
 Шварцшильда радиус (Schwarzschild radius), 1270
 Тахион (tachyon), 140
 Тело (division ring), 10
 Темная энергия (dark energy), 1358
 Темная материя (dark matter), 1358
 Тензор (tensor), 40, 41
 Тензор Эйнштейна (Einstein tensor), 842
 Тензор Коттона (Cotton tensor), 656
 Тензор Риччи (Ricci tensor), 331
 Тензор Римана-Кристоффеля (Riemann-Christoffel tensor), 327
 Тензор Вейля (Weyl tensor), 646
 Тензор деформаций (deformation tensor), 1403
 Тензор дисторсии (distortion tensor), 1406
 Тензор энергии-импульса (energy-momentum tensor), 725
 Тензор энергии-импульса канонический (canonical energy-momentum tensor), 725
 Тензор энергии-импульса материя (energy-momentum tensor of matter), 842
 Тензор изгиба-кручения (twist tensor), 1421
 Тензор кокручения (contorsion tensor), 313
 Тензор конформной кривизны (conformal curvature tensor), 646
 Тензор кривизны (curvature tensor), 327
 Тензор кручения (torsion tensor), 297
 Тензор момента количества движения (angular-momentum tensor), 726
 Тензор моментных напряжений (torque stress tensor), 1421
 Тензор неметричности (nonmetricity tensor), 312
 Тензор полностью антисимметричный (totally antisymmetric tensor), 198
 Тензор разложимый (decomposable tensor), 40
 Тензор сдвига (shear tensor), 1324, 1326
 Тензор вращения (rotation tensor), 1324, 1326
 Тензора ранг (rank of a tensor), 192
 Тензориальная r -форма (pseudotensorial r -form), 503
 Тензорная алгебра (tensor algebra), 41, 194
 Тензорная плотность (tensor density), 197
 Тензорное поле $\mathbb{G}$ -инвариантное ($\mathbb{G}$ -invariant tensor field), 436

- Тензорное поле инвариантное (invariant tensor field), 229
Тензорное поле неприводимое (irreducible tensor field), 195
Тензорное поле разложимое (decomposable tensor field), 195
Тензорное поле типа (r, s) (tensor field of type (r, s)), 192
Тензорное произведение (tensor product), 39, 54
Тензорное произведение расслоений (tensor product of fiber bundles), 288
Тензорное произведение тензорных полей (tensor product of tensor fields), 194
Тензорное расслоение (tensor fiber bundle), 482
Теорема *PCT* (*PCT* theorem), 1103
Теорема Адо (Ado theorem), 419
Теорема Бирхгоффа (Birkhoff's theorem), 1244, 1272
Теорема Дарбу (Darboux theorem), 250
Теорема Фробениуса (Frobenius theorem), 211
Теорема Гейне–Бореля–Лебера (Heine–Borel–Lebesgue theorem), 26
Теорема Голдстоуна (Goldstone theorem), 965
Теорема Хопфа–Ринова (Hopf–Rinov theorem), 638
Теорема Картана (Cartan theorem), 386
Теорема Картана–Дьюдоне (Cartan–Dieudonné theorem), 1068
Теорема Лиувилля (Liouville's theorem), 765
Теорема Нетер первая (first Noether's theorem), 722
Теорема Нетер вторая (second Noether's theorem), 728
Теорема Пифагора (Pythagorean theorem), 48
Теорема Тихонова (Tychonoff theorem), 26
Теорема о невозможности причесать ежа (you can't comb the hair of a hedgehog), 446
Теорема редукции (reduction theorem), 527
Теплопроводности уравнение (heat equation), 1522
Тетрада (tetrad), 334
Тип гомотопический (homotopic type), 445
Типичный слой (typical fiber), 161
Точечное преобразование (pointwise transformation), 746
Точка граничная (boundary point), 23
Точка изолированная (isolated point), 23
Точка критическая (critical point), 713
Точка многообразия краевая (boundary point of a manifold), 159
Точка многообразия внутренняя (interior point of a manifold), 159
Точка неподвижная (fixed point), 428
Точка предельная (limit point), 21
Точка пространства (space point), 2
Точка сингулярная (singular point), 840
Точка стационарная (fixed point), 428
Точка стационарная (stationary point), 713
Точка внешняя (external point), 23
Точка внутренняя (internal point), 23
Точки сопряженные (conjugate points), 611
Точная форма (exact form), 242
Точное представление алгебры Ли (exact representation of a Lie algebra), 394
Точное представление группы Ли (exact representation of a Lie group), 376
Ток киральный (chiral current), 1110
Ток сохраняющийся (conserved current), 723
Ток спинорный (spinor current), 1105
Топологическая метрика (topological metric), 16
Топологическая модель (topological model), 800, 1229
Топологический инвариант (topological invariant), 77
Топологическое многообразие (topological manifold), 149
Топологическое произведение (topological product), 19
Топологическое пространство (topological space), 18
Топологическое пространство линейно связное (arcwise connected, path connected topological space), 441
Топологическое пространство локально евклидово (locally Euclidean topological space), 151
Топологическое пространство локально компактное (locally compact topological space), 27
Топологическое пространство локально линейно связное (locally arcwise connected, locally path connected topological space), 441
Топологическое пространство метризуемое (metrizable topological space), 20
Топологическое пространство несвязное (disconnected topological space), 29
Топологическое пространство счетное в бесконечности (locally compact σ -compact topological space), 157
Топологическое пространство связное (connected topological space), 29
Топологическое свойство (topological property), 77
Топологическое векторное пространство (topological vector space), 46
Топологическое векторное пространство метризуемое (metrizable topological vector space), 46
Топологии база (base of topology), 18
Топология (topology), 18
Топология антидискретная (indiscrete topology), 19
Топология более грубая (coarser topology), 19
Топология более сильная (stronger topology), 19
Топология более слабая (coarser topology), 19
Топология более тонкая (stronger topology), 19
Топология дискретная (discrete topology), 18
Топология естественная (natural topology), 19
Топология идентификации (identification topology), 77
Топология индуцированная (induced topology), 20
Топология тривиальная (trivial topology), 19
Топ (torus), 68
Тождества Бианки (Bianchi identities), 293, 301, 339
Тождества Якоби (Jacobi identities), 42, 190, 354
Тождественный путь (identical path), 440
Тождественное отображение (identity map), 71
Тождество Фирца (Fierz identity), 1093
Траектория Киллинга (Killing trajectory), 570
Траектория частицы (particle trajectory), 738, 861
Трансляция группа (translational group), 98
Трансляция (translation), 67
Трансверсальные подмногообразия (transversal submanifolds), 209
Трансверсальные векторные пространства (transversal vector spaces), 35
Трансверсальное отображение (transversal map), 209
Трансверсальности условия (transversality conditions), 718
Транзитивное действие группы (transitive action of a group), 427
Третий закон Ньютона (third Newton's law), 130
Треугольника неравенство (triangle inequality), 16
Триада (triad), 334
Тривиальная топология (trivial topology), 19
Тривиальное ассоциированное расслоение (trivial associated fiber bundle), 485
Тривиальное действие группы (trivial action of a group), 425
Тривиальное главное расслоение (trivial principal fiber bundle), 471
Тривиальное многообразие (trivial manifold), 152
Тривиальное расслоение (trivial fiber bundle), 162
Тривиальное скалярное поле (trivial scalar field), 164
Тривиальное векторное поле (trivial vector field), 167
Трубка характеристик (tube of characteristics), 748
Угол азимутальный (azimuth angle), 340, 1133
Угол между векторами (angle between vectors), 270
Угол полярный (polar angle), 340, 345, 1133
Угол вращения (rotational angle), 85
Угол вращения гиперболический (hyperbolic rotational angle), 91
Укороченная функция действия (truncated action function), 756
Укороченное действие (truncated action), 753
Укороченное уравнение Гамильтона–Якоби (truncated Hamilton–Jacobi equation), 756
Ультрарелятивистский предел (ultrarelativistic limit), 876
Умножение внешнее (exterior product, wedge product), 232
Умножение внутреннее (internal multiplication), 237
Унимодулярная гравитация (unimodular gravity), 860

- Унимодулярная группа Ли (unimodular Lie group), 383
- Унитарная алгебра (unital algebra), 1022
- Унитарный $\mathfrak{k}$ -модуль (unital $\mathfrak{k}$ -module), 34
- Унитарный $\mathfrak{k}$ -модуль (unitary $\mathfrak{k}$ -module), 34
- Унитарное представление (unitary representation), 385
- Унитарность (unitarity), 414
- Универсальное накрытие (universal covering), 459
- Упорядоченное произведение (ordered product), 519
- Уравнение Бернулли (Bernoulli equation), 892
- Уравнение Даламбера (d'Alembert equation), 1522
- Уравнение Дирака (Dirac equation), 1104
- Уравнение Эйлера (Euler equation), 887
- Уравнение Гамильтона–Якоби (Hamilton–Jacobi equation), 755
- Уравнение Гамильтона–Якоби укороченное (truncated Hamilton–Jacobi equation), 756
- Уравнение Гельмгольца (Helmholtz equation), 1521
- Уравнение Киллинга (Killing equation), 570
- Уравнение Клейна–Гордона–Фока (Klein–Gordon–Fock equation), 955
- Уравнение Лапласа (Laplace equation), 1521
- Уравнение Лиувилля (Liouville equation), 1126, 1129, 1148
- Уравнение Райчаудхури (Raychaudhuri equation), 1329
- Уравнение Шредингера (Schrödinger equation), 542
- Уравнение Якоби (Jacobi equation), 611
- Уравнение девиации геодезических (geodesic deviation equation), 602
- Уравнение эйконала (eikonal equation), 619
- Уравнение характеристическое (secular equation), 409
- Уравнение непрерывности (continuity equation), 886
- Уравнение теплопроводности (heat equation), 1522
- Уравнение вековое (secular equation), 409
- Уравнение волновое (wave equation), 615
- Уравнения Эйлера вращения твердого тела (Euler equations of rigid body rotations), 706
- Уравнения Эйлера–Лагранжа (Euler–Lagrange equations), 714
- Уравнения Эйнштейна (Einstein's equations), 842
- Уравнения Эйнштейна вакуумные (vacuum Einstein's equations), 843
- Уравнения Гаусса–Петерсона–Кодацци (Gauss–Peterson–Codazzi equations), 930
- Уравнения Максвелла (Maxwell's equations), 993
- Уравнения Ньютона (Newton equations), 743
- Уравнения Вейля (Weyl equations), 1110
- Уравнения движения (equations of motion), 722
- Уравнения структурные (structure equations), 301
- Ускорение (acceleration), 143, 861
- Ускорение девиации (deviation acceleration), 601
- Ускорение кривой (acceleration of a curve), 597
- Ускорение наблюдаемое (observed acceleration), 143
- Условие Липшица (Lipschitz condition), 180
- Условие доминантное энергетическое (dominant energy condition), 889
- Условие энергетическое (energy condition), 889
- Условие гармоничности (harmonic condition), 622
- Условие калибровочное (gauge condition), 797
- Условие метричности (metricity condition), 314
- Условие положительности Риччи (Ricci positivity condition), 889
- Условие сильное энергетическое (strong energy condition), 889
- Условие слабое энергетическое (weak energy condition), 889
- Условие унитарности (unitarity condition), 414
- Условия интегрируемости Сен–Венана (Saint–Venant integrability conditions), 1405
- Условия трансверсальности (transversality conditions), 718
- Узел (knot), 183
- Вакансия (vacancy), 1409
- Вакуумные уравнения Эйнштейна (vacuum Einstein's equations), 843
- Вариационная производная (variational derivative), 714
- Вариация аффинной связности (variation of an affine connection), 311
- Вариация формы (form variation), 221
- Вариация первая (first variation), 609
- Вариация вторая (second variation), 609
- Вейлевская эквивалентность (Weyl equivalence), 655
- Вейлевски евклидова метрика (Weyl Euclidean metric), 281
- Вейля форма (Weyl form), 313
- Вейля представление (Weyl representation), 1087
- Вейля преобразование (Weyl transformation), 271, 653, 659
- Вейля спинор (Weyl spinor), 1098
- Вейля тензор (Weyl tensor), 646
- Вейля уравнения (Weyl equations), 1110
- Вековое уравнение (secular equation), 409
- Вектор (vector), 30, 31
- Вектор Бюргерса (Burgers vector), 1405
- Вектор Франка (Frank vector), 1409, 1410
- Вектор Пойнтинга (Pointing vector), 997
- Вектор девиации (deviation vector), 601
- Вектор горизонтальный (horizontal vector), 493
- Вектор касательный (tangent vector), 170, 181
- Вектор нулевой (null vector), 30, 50
- Вектор скорости (velocity vector), 15
- Вектор смещения (displacement vector), 221
- Вектор светоподобный (light-like vector), 50
- Вектор вертикальный (vertical vector), 493
- Вектор волновой (wave vector), 620
- Вектор вращения (rotational vector), 100
- Векторы ортогональные (orthogonal vectors), 47, 270, 275
- Векторы перпендикулярные (perpendicular vectors), 270, 275
- Векторные пространства трансверсальные (transversal vector spaces), 35
- Векторное поле (vector field), 166, 170, 177
- Векторное поле Киллинга (Killing vector field), 570
- Векторное поле Киллинга конформное (conformal Killing vector field), 653
- Векторное поле фундаментальное (fundamental vector field), 491
- Векторное поле гамильтоново (Hamiltonian vector field), 704
- Векторное поле изотропное (isotropic vector field), 275
- Векторное поле кокасательное (cotangent vector field), 172
- Векторное поле контравариантное (contravariant vector field), 192
- Векторное поле ковариантное (covariant vector field), 192
- Векторное поле левоинвариантное (left invariant vector field), 355
- Векторное поле нулевое (null vector field), 275, 747
- Векторное поле полное (complete vector field), 181
- Векторное поле пространственноподобное (space-like vector field), 275
- Векторное поле смещений (displacement vector field), 1402
- Векторное поле светоподобное (light-like vector field), 275
- Векторное поле тривиальное (trivial vector field), 167
- Векторное поле времениподобное (time-like vector field), 275
- Векторное поле якобиево (Jacobi vector field), 611
- Векторное представление группы $\mathbb{O}(1, 1)$ (vector representation of $\mathbb{O}(1, 1)$ group), 94
- Векторное пространство (vector space), 31
- Векторное пространство градуированное (graded vector space), 37
- Векторное пространство кокасательное (cotangent vector space), 172
- Векторное пространство нормированное (normed vector space), 44
- Векторное пространство топологическое (topological vector space), 46
- Векторное расслоение (vector fiber bundle), 285
- Векторного поля поток (vector field flow), 187
- Вертикальный автоморфизм (vertical automorphism), 487
- Вертикальный вектор (vertical vector), 493

- Вертикальное подпространство (vertical subspace), 299, 492, 525
- Вещественная симплектическая группа (real symplectic group), 689
- Выпуклая функция (convex function), 735
- Выпуклая область (convex domain), 445
- Выпуклое множество (convex set), 36
- Вырожденная модель (degenerate model), 802, 810
- Вырожденный симплекс (degenerate simplex), 70
- Вильсона петля (Wilson loop), 521
- Винтовая дислокация (screw dislocation), 1405
- Вложение изометрическое (isometric embedding), 280
- Вложение многообразия (embedding of a manifold), 204
- Вложение расслоения (embedding of a fiber bundle), 487
- Вложение регулярное (regular embedding), 206
- Внешнее дифференцирование (exterior differentiation), 240
- Внешнее ковариантное дифференцирование (external covariant differentiation), 504
- Внешнее умножение (exterior product, wedge product), 232
- Внешний автоморфизм (outer automorphism), 389
- Внешняя алгебра (exterior algebra), 234
- Внешняя экспонента (exterior exponent), 236
- Внешняя ковариантная производная (external covariant derivative), 504
- Внешняя кривизна гиперповерхности (external curvature of hypersurface), 928
- Внешняя степень (external power), 236
- Внешняя точка (external point), 23
- Внутреннее дифференцирование (internal differentiation), 237
- Внутреннее умножение (internal multiplication), 237
- Внутренний автоморфизм (inner automorphism), 389
- Внутренность множества (set interior), 22
- Внутренняя точка (internal point), 23
- Внутренняя точка многообразия (interior point of a manifold), 159
- Волновая функция (wave function), 542
- Волновое уравнение (wave equation), 615
- Волновой фронт (wave front), 757
- Волновой вектор (wave vector), 620
- Возврат отображения (pullback), 55, 202
- Вполне интегрируемая пфафхова система уравнений (totally integrable Pfaffian system of equations), 213
- Вполне интегрируемое распределение (totally integrable distribution), 210
- Вращение дуальное (dual rotation), 996
- Вращение изотопическое (isotopic rotation), 1010
- Вращение против часовой стрелки (counterclockwise rotation), 88
- Времени обращение (time inversion), 275, 1102
- Времениподобная кривая (time-like curve), 108
- Времениподобное векторное поле (time-like vector field), 275
- Временная калибровка (temporal gauge), 821, 901, 978
- Время (time), 128, 275
- Время космологическое (cosmological time), 1347
- Время собственное (proper time), 608, 861
- Вселенная однородная (homogeneous universe), 1390
- Вселенная однородная и изотропная (homogeneous and isotropic universe), 1346
- Вселенная открытая (open universe), 160, 841
- Вселенная замкнутая (closed universe), 160, 841
- Всюду плотная обмотка тора (everywhere dense winding of torus), 205
- Всюду плотное множество (everywhere dense set), 27
- Всюду плотное подмножество (everywhere dense subset), 23
- Вторая аксиома счетности (second axiom of countability), 21
- Вторая космическая скорость (second cosmic velocity), 1320
- Вторая теорема Нетер (second Nether's theorem), 728
- Вторая вариация (second variation), 609
- Вторичная связь (secondary constraint), 806
- Второй закон Ньютона (second Newton's law), 130, 869
- Взаимно однозначное отображение (one-to-one map), 71
- Взрыв большой (big bang), 1332
- Ядро гомоморфизма (kernel of a homomorphism), 375, 394
- Ядро группы преобразований (kernel of a transformation group), 426
- Якоби матрица (Jacobi matrix), 80
- Якоби оператор (Jacobi operator), 610
- Якоби тождества (Jacobi identities), 42, 190, 354
- Якобиан (Jacobian), 80
- Якобиево векторное поле (Jacobi vector field), 611
- Янга–Миллса поле (Yang–Mills field), 293, 500, 1005
- Юнга неравенство (Young inequality), 736
- Закон Гука (Hooke law), 1403
- Закон Хаббла (Hubble's law), 1353
- Закон Ньютона первый (first Newton's law), 130
- Закон Ньютона третий (third Newton's law), 130
- Закон Ньютона второй (second Newton's law), 130
- Закон ассоциативности (associativity law), 5
- Закон гравитации Ньютона (Newton's gravitational law), 130
- Закон композиции (composition rule), 5
- Закон всемирного тяготения (Newton's gravitational law), 130, 872
- Законы дистрибутивности (distributivity rules), 9
- Замедление времени (clock retardation), 136
- Замыкание множества (closure of a set), 22
- Замкнутая форма (closed form), 242
- Замкнутая кривая (closed curve), 15
- Замкнутая подгруппа Ли (closed Lie subgroup), 378
- Замкнутая вселенная (closed universe), 160, 841
- Замкнутый путь (closed path), 439
- Замкнутый симплекс (closed simplex), 70
- Замкнутое многообразие (closed manifold), 160
- Замкнутое множество (closed set), 21
- Замкнутое отображение (closed map), 76
- Замкнутое подмногообразие (closed submanifold), 207
- Заряд электрический (electric charge), 972
- Заряд сохраняющийся (conserved charge), 724
- Зарядовое сопряжение (charge conjugation), 1101
- Зависимость функциональная (functional dependence), 82
- Значение собственное (eigenvalue), 1504
- Знак перестановки (sign of permutation), 199
- Жидкость баротропная (barotropic fluid), 888
- Жидкость идеальная (ideal fluid), 891
- Жорданов блок (Jordan box), 1507
- Жорданов многоугольник (Jordan polygon), 440
- Жорданов узел (Jordan knot), 183
- Жорданова клетка (Jordan box), 1507
- Жорданова кривая (Jordan curve), 440
- Жорданова матрица (Jordan matrix), 1507