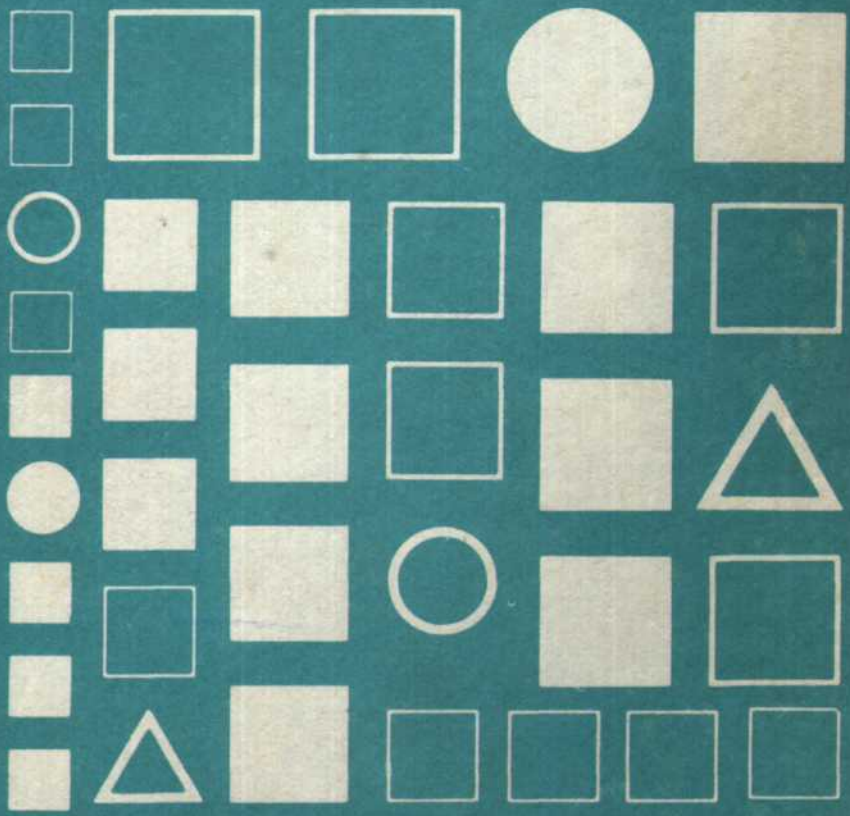
22.1 л 73
Л93

В. А. Любецкий

ОСНОВНЫЕ ПОНЯТИЯ ШКОЛЬНОЙ МАТЕМАТИКИ



В.А. Любецкий

ОСНОВНЫЕ ПОНЯТИЯ ШКОЛЬНОЙ МАТЕМАТИКИ

Допущено
Министерством просвещения СССР
в качестве учебного пособия для студентов
педагогических институтов
по специальности № 2104 «Математика»

МОСКВА
«ПРОСВЕЩЕНИЕ» 1987

ББК 22.1
Л93

Рецензенты:
кафедра высшей алгебры, элементарной математики и методики преподавания математики МОПИ им. Н. К. Крупской (зав. кафедрой — доцент М. М. Рассудовская);
доктор физ.-мат. наук, профессор М. М. Постников (МГУ им. М. В. Ломоносова)

Василий Александрович Любецкий

ОСНОВНЫЕ ПОНЯТИЯ
ШКОЛЬНОЙ МАТЕМАТИКИ

Зав. редакцией *Р. А. Хабиб*
Редакторы *Т. В. Автономова, Н. Р. Брумберг*
Младшие редакторы *Л. Е. Козырева, Е. А. Сафронова*
Художник *В. П. Трифонов*
Художественный редактор *Е. Н. Карасик*
Технический редактор *С. С. Якушкина*
Корректор *Т. А. Воробьева*

ИБ № 9715

Сдано в набор 04.12.85. Подписано к печати 02.02.87. Формат 60×90¹/₁₆. Бум. офс. № 2. Гарнитура литературная. Печать офсетная. Усл. печ. л. 25,0. Усл. кр.-отт. 50,19. Уч.-изд. л. 25,98. Тираж 58000. Заказ № 227. Цена 1 р. 10 к.

Ордена Трудового Красного Знамени издательство «Просвещение» Государственного комитета РСФСР по делам издательств, полиграфии и книжной торговли. 129846, Москва, 3-й проезд Марьиной рощи, 41.

Отпечатано с диапозитивов Смоленского полиграфкомбината Росглавполиграфпрома Государственного комитета РСФСР по делам издательств, полиграфии и книжной торговли. 214020, г. Смоленск, ул. Смольянинова, 1 на Саратовском ордена Трудового Красного Знамени полиграфическом комбинате Росглавполиграфпрома Государственного комитета РСФСР по делам издательств, полиграфии и книжной торговли. 410004, Саратов, ул. Чернышевского, 59.

Любецкий В. А.

Л93 Основные понятия школьной математики: Учеб. пособие для студентов пед. ин-тов по спец. № 2104 «Математика».— М.: Просвещение, 1987.—400 с.: ил.

В учебном пособии излагаются основные понятия школьной математики (элементарные функции, угол, вектор, плоскость, планиметрия, измерение величин, площадь и мера плоской фигуры, решение алгебраических уравнений, геометрические построения, основания понятия числа) с точки зрения математических курсов пединститута; выясняется место этих основных понятий в системе представлений высшей математики.

ББК 22.1

Л $\frac{4309000000-389}{103(03)-87}$ 35—86

© Издательство «Просвещение», 1987

ОГЛАВЛЕНИЕ

Предисловие 7
Основные термины и символы, используемые в книге 11

Глава I

Элементарные функции. Угол

Введение 19
§ 1. Линейная функция 22
1. Аксиоматическое определение линейной функции 22
2. Свойства линейной функции 22
3. Теорема существования и единственности линейной функции . . . 23
§ 2. Показательная функция 24
1. Аксиоматическое определение показательной функции 24
2. Свойства показательной функции 24
3. Теорема существования и единственности показательной функции . . 26
§ 3. Логарифмическая функция 30
1. Аксиоматическое определение логарифмической функции 30
2. Свойства логарифмических функций. Теорема существования и единственности логарифмической функции 31
§ 4. Степенная функция 32
1. Аксиоматическое определение степенной функции 32
2. Теорема существования и единственности степенной функции . . . 34
3. Свойства степенных функций 34
§ 5. Функции косинус и синус числового аргумента 35
1. Экспоненциальная функция и ее периодичность 35
2. Теоремы существования и единственности экспоненциальной функции 40
3. Функции косинус и синус числового аргумента: аксиоматические определения и свойства 45
§ 6. Угол. Функции косинус и синус углового аргумента. Измерение углов . . . 48
1. Введение 48
2. Определение угла в арифметической плоскости 49
3. Конструктивные определения функций косинус и синус углового аргумента. Свойства этих функций 53
4. Измерение углов 55
5. Обсуждение полученных результатов 60

Глава II

Вектор. Плоскость. Планиметрия

Введение 64
§ 1. Сравнение различных подходов к понятию вектора 66

1. Вектор как пара чисел. Свободный вектор. Вектор как параллельный перенос . . . 66
2. Вектор как дифференцирование. Вектор как класс касающихся кривых . . . 70
3. Вектор как тензор . . . 75

§ 2. Понятие плоскости . . . 77
1. Аффинная плоскость . . . 77
2. Школьные геометрические понятия в аффинной плоскости . . . 80
3. Плоскость с формой . . . 84
4. Проективная плоскость . . . 89

§ 3. Аксиоматический подход к определению плоскости . . . 94
1. Два типа аксиоматического определения плоскости . . . 94
2. Аксиоматическое теоретико-множественное определение плоскости . . . 95
3. Аксиоматики плоскости Евклида — Гильберта, Лобачевского и Римана . . . 98
4. Двумерные римановы многообразия как модели аксиоматических определений плоскости . . . 106

§ 4. Основные группы школьной планиметрии и их действие в плоскости . . . 113
1. Аффинные отображения . . . 113
2. Основные группы школьной планиметрии, действующие в арифметической плоскости . . . 118
3. Поднятие группы биекций в арифметической плоскости в векторную и аффинную плоскости . . . 123

§ 5. Понятие планиметрии . . . 126
1. Клейновский подход в геометрии: понятие о планиметрии данной группы . . . 126
2. Евклидова планиметрия — планиметрия ортогональной группы . . . 129

Глава III

Измерение величин. Площадь и мера плоских фигур

Введение . . . 134

§ 1. Примеры измерений и величин . . . 137
§ 2. Положительная скалярная величина . . . 140
§ 3. Измерение площади многоугольника . . . 154
1. Конструктивное определение площади многоугольника. Свойство конечной аддитивности . . . 154
2. Инвариантность функции площади относительно эквиаффинной группы . . . 158

§ 4. Сравнение конструктивного и аксиоматического определений площади многоугольника. Сравнение различных способов измерения площади многоугольника . . . 161
1. Аксиоматическое определение площади многоугольника и его сравнение с конструктивным определением . . . 161
2. Определение площади многоугольника с помощью движений . . . 165
3. Способы измерения площади многоугольника . . . 167

§ 5. Сравнение конструктивного и аксиоматического определений меры плоской фигуры. Вычисление меры простейших криволинейных фигур . . . 178
1. Измерение плоских криволинейных фигур . . . 178
2. Неизмеримые множества . . . 191
3. Аксиоматическое определение меры . . . 193
4. Сравнение конструктивного и аксиоматического определений меры . . . 202
5. Вычисление меры простейших криволинейных фигур . . . 205
6. Сравнение борелевской меры с мерами Жордана и Лебега . . . 208

Глава IV

Алгебраические уравнения степеней, меньших или равных 5, и геометрические построения

§ 1. Связь между разрешимостью алгебраических уравнений в радикалах и выполнимостью традиционных геометрических построений . . . 210
1. Кубические уравнения и квадратичные расширения . . . 210
2. Построение циркулем и линейкой . . . 212
3. Проблемы удвоения куба, трисекции угла и построения правильного семиугольника с помощью циркуля и линейки . . . 218
4. Геометрические построения, включающие операцию выбора произвольной точки в заданной фигуре . . . 221
5. Геометрические построения с помощью одного циркуля . . . 224

§ 2. Задача о разрешимости алгебраических уравнений в радикалах. Критерий разрешимости. Пример неразрешимого в радикалах алгебраического уравнения 5-й степени . . . 227
1. Постановка задачи о разрешимости алгебраического уравнения в радикалах . . . 227
2. Понятие разрешимой группы . . . 232
3. Определение симметрической и знакопеременной групп . . . 233
4. Разрешимость симметрической и знакопеременной групп . . . 236
5. Понятие группы Галуа. Формулировка теоремы Галуа . . . 241
6. Пример алгебраического уравнения, группа Галуа которого совпадает с симметрической группой 5-й степени . . . 247
7. Доказательство необходимого условия в теореме Галуа . . . 254

§ 3. Решение алгебраических уравнений степени, меньшей или равной 4, в радикалах . . . 261
1. План решения в радикалах алгебраических уравнений с разрешимой группой Галуа . . . 261
2. Разрешимость в радикалах алгебраических уравнений с циклической группой Галуа . . . 262
3. Разрешимость в радикалах квадратного уравнения . . . 266
4. Разрешимость в радикалах алгебраических уравнений с разрешимой группой Галуа . . . 268
5. Разрешимость в радикалах кубического уравнения . . . 269

Глава V

Логико-математические основания понятия числа

§ 1. Понятие натурального ряда . . . 273
1. Финитный подход к определению натурального ряда . . . 273
2. Теоретико-множественный и аксиоматический подходы к определению натурального ряда . . . 275
3. Сравнение определений целых чисел . . . 281

§ 2. Определение рационального числа как линейной функции . . . 282
§ 3. Основные подходы к определению вещественных чисел . . . 287
1. Определение вещественного числа как фундаментальной последовательности . . . 287
2. Продолжение алгебраических операций с поля на его пополнение . . . 291
3. Определение вещественного числа как сечения . . . 296
4. Определение вещественного числа как последовательности знаков . . . 301

§ 4. Основные подходы к определению комплексных чисел . . . 308
§ 5. Роль алгебраической замкнутости, локальной компактности и упорядоченности среди свойств комплексных и вещественных чисел . . . 310
§ 6. Связь полей вещественных и комплексных чисел. Продолжение линейного порядка с поля на его алгебраическое расширение и метрическое пополнение . . . 320

Приложение 1 (к главе I)

1. Группы, изоморфные прямой и окружности 324
2. Длина дуги. Определение функций косинус и синус числового аргумента на основе понятия длины дуги 332

Приложение 2 (к главе III)

Доказательство теоремы о моделях системы положительных скалярных величин 341

Приложение 3 (к главе IV)

Доказательство некоторых вспомогательных алгебраических утверждений 346

Приложение 4

Сферическая, гиперболическая и эллиптическая плоскости

§ 1. Точки, прямые и отрезки в сферической, эллиптической и гиперболической плоскостях 354
§ 2. Метрики в сферической, эллиптической и гиперболической плоскостях . 366
§ 3. Группы движений и измерении углов в сферической, эллиптической и гиперболической плоскостях 383

ПРЕДИСЛОВИЕ

Задача этой книги — показать место основных понятий школьной математики в гораздо более широкой системе представлений высшей математики и в этих рамках строго и последовательно изложить понятия школьной (элементарной) математики с точки врения высшей математики (которая отождествляется с содержанием пединститутских курсов алгебры и теории чисел, анализа, геометрии, математической логики и теории алгоритмов).

Хорошо известно, что многие выпускники пединститутов — будущие учителя, испытывают затруднения в своей профессиональной области — школьной математике. Это касается умения решать элементарные задачи и, в еще большей степени, понимания тонких вопросов элементарной математики, умения связывать те обширные математические теории, которые изучаются в течение четырех-пяти лет в пединституте, с конкретными вопросами элементарной математики. Цель пособия — помочь преодолеть две последние из отмеченных трудностей, способствуя тем самым усилению профессиональной направленности в подготовке учителя.

В первых главах рассматриваются наиболее традиционные понятия школьной математики: элементарные функции, угол, измерение углов (глава I); вектор, плоскость, планиметрия (глава II); величина, площадь и мера плоской фигуры (глава III), геометрические построения циркулем и линейкой, решение алгебраических уравнений низших степеней в радикалах (глава IV).

Менее элементарную направленность имеют глава V и приложение 4. Поэтому чуть подробнее коснемся их содержания. В § 1 главы V детально рассматривается построение системы натуральных чисел — основы всех числовых систем. В § 2 этой главы традиционный подход к понятию рационального числа сравнивается с другим подходом, в рамках которого рациональное число — функция. В § 3 рассматриваются основные способы перехода от рациональных чисел — дискретного объекта к вещественным и комплексным числам — непрерывным объектам (в § 5 эта линия изложения продолжается переходом от рациональных чисел к нечисловым p-адическим полям). В целом § 3, 4, 5 пятой главы посвящены алгебротопологическим свойствам вещественных чисел; включение этого материала связано с тем, что именно сочетание алгебраических

и топологических свойств создает вещественным и комплексным числам уникальное положение в математике. Приложение 4 содержит подробное изложение элементарных вопросов неевклидовой планиметрии. Ясное понимание евклидовой планиметрии (о которой говорится в главе II), по-видимому, предполагает для контраста знакомство с неевклидовой планиметрией.

Для согласования терминологии и обозначений после предисловия приводится материал, содержащий некоторые общие понятия высшей математики; эти понятия играют в книге подсобную роль — языка, на котором говорится о школьной математике. Правильно рассматривать их как специализированную часть русского языка, подобную языку врача, химика или биолога. Разумно обращаться к этому материалу лишь в том случае, если какие-то обозначения или термины, употребляемые в книге, оказываются для читателя новыми и их смысл не ясен из контекста.

Предполагается, что читатель знаком с основными понятиями школьной математики на том предварительном уровне понимания, который выносится из школы и первых трех курсов пединститутов. Также предполагается некоторая опытность читателя в оперировании с основными алгебраическими, топологическими и логическими понятиями из упомянутых математических курсов; однако фактическое содержание этих курсов может быть не знакомо (или почти не знакомо) читателю. Поэтому изложение в книге ведется постепенно, как правило, с полными определениями и доказательствами; от читателя в основном требуется умение не спеша разбирать временами довольно длинные построения. Параграфы 5, 6 пятой главы предъявляют более высокие требования к читателю, так как изложение в них носит обзорный характер. В книге встречаются довольно абстрактные понятия, такие, как индуцированная топология, топологическое пространство, подгруппа, гомоморфизм, связность, локальная компактность, действие, модуль, однако они употребляются исключительно для случаев $\boldsymbol{R}$, $\boldsymbol{R}_{>0}$, $\boldsymbol{Z}$, $\boldsymbol{C}$, S^1, S^2. Конечно, для таких простейших случаев эти понятия можно заменить соответствующими частными, внешне более простыми выражениями. Например, вместо локальной компактности можно говорить о наличии окрестности, являющейся отрезком или дугой, включающей концы. Такая замена вряд ли приведет к упрощению существа дела и в то же время сделает многие формулировки внешне тяжеловесными и специфически привязанными к каждому из отдельных случаев; тем более, что эти понятия рассматриваются в основных математических курсах. Для некоторых категорий читателей такая замена абстрактных терминов соответствующими элементарными выражениями может быть полезным упражнением, относящимся по существу не к математике, а к русскому языку.

Степень детальности в рассмотрении того или иного понятия школьной математики различна и зависит от внимания, которое ему уделяется в основных математических курсах.

Так, понятия элементарной функции, угла и измерения углов в их элементарных аспектах известны студенту старших курсов почти на том же уровне, что и выпускнику школы. Поэтому здесь изложение носит систематический характер.

Понятие вектора обычно определяется аксиоматически, как элемент произвольного векторного пространства. При всей важности такого аксиоматического подхода нужно представлять себе и конкретные модели аксиоматического определения вектора, в том числе и только простейшую модель вектора как направленного отрезка. Именно разнообразие этих моделей придает понятию вектора фундаментальное значение. Поэтому подробно рассматриваются различные конструктивные подходы к понятию вектора.

Понятие геометрической плоскости тщательно изучается в курсе геометрии, поэтому мы касаемся его бегло, только в плане адекватности различных определений плоскости интуитивному представлению о ней. Понятие планиметрии с аксиоматической точки зрения также подробно рассматривается в курсе геометрии, и мы касаемся его только в обзорном порядке. Однако при всей важности аксиоматического понимания планиметрии существенна и клейновская точка зрения на нее. Поэтому подробно рассматривается клейновский подход и, в частности, вычисляются все инварианты ортогональной группы, которые и образуют с этой точки зрения евклидову планиметрию.

Понятие величины подробно рассматривается в книге, так как в сущности оно отсутствует в основных математических курсах. Столь же подробно рассматриваются и сравниваются различные способы измерения площади многоугольника, и в этой связи напоминается аксиоматическое определение площади многоугольника.

Мера понимается как продолжение функции площади с множества многоугольников на более широкое множество криволинейных фигур. В то же время мера определяется аксиоматически и эти два подхода тщательно сравниваются. Затем на основе аксиоматического определения меры (без использования интегралов) вычисляются ее значения для круга, сектора, сегмента и т. п. элементарных плоских фигур. В курсе анализа рассматривается мера Лебега только на прямой, здесь по существу рассматривается мера Лебега на плоскости. Таким образом, в этом вопросе, как и в других, автор стремился обеспечить преемственность излагаемого материала по отношению к основным курсам.

Подробно рассматриваются классические задачи об удвоении объема куба, трисекции угла и построении правильного семиугольника с помощью циркуля и линейки. При этом доказывается невозможность таких построений.

Затем подробно изучается вопрос о разрешимости в радикалах алгебраических уравнений степени, меньшей или равной 5. Доказывается теорема Галуа. На ее основе находятся известные формулы для решения уравнений степени, меньшей или равной 4.

В книге большинство вопросов рассматривается с точки зрения инвариантов подходящей группы преобразований, т. е. инвариан-

тов действия подходящей группы; иными словами, с точки зрения непрерывных гомоморфизмов простейших групп. Можно надеяться, что такая точка зрения придает книге цельный, единообразный характер.

В книге можно найти материал для факультативных занятий в школе. Однако вопросы преподавания математики в школе и вопросы изложения ее в школьных учебниках здесь не рассматриваются. В этом, как и в других отношениях, автор старался следовать духу книги Ф. Клейна «Элементарная математика с точки зрения высшей». Книга Ф. Клейна своей конкретной содержательностью мало похожа на ряд современных изложений элементарной математики, в которых на первый план выдвигаются вопросы формально-логического порядка, например вопросы такого типа, как является ли элементарная функция множеством пар или отношением; кажется, что такого рода вопросы маловажны для существа дела.

Автор неоднократно читал лекционный курс, одноименный с названием книги, для слушателей факультета повышения квалификации преподавателей и студентов пятого курса математического факультета. Эти лекции были отпечатаны и после некоторой правки составили рукопись книги; поэтому особенности, терпимые в лекционном изложении, к сожалению, перешли в книгу.

Приложение 4 написано П. В. Семеновым. Автор благодарит его также за большую помощь в подготовке рукописи.

Автор глубоко признателен В. Т. Базылеву, К. И. Дуничеву, Л. Я. Куликову, В. И. Мишину, А. И. Москаленко, Р. С. Черкасову, Е. П. Шимбиревой, Е. А. Щеголькову за ценные советы и указания во время работы над курсом и книгой.

Автор посвящает книгу своим детям Василине и Елене.

ОСНОВНЫЕ ТЕРМИНЫ И СИМВОЛЫ, ИСПОЛЬЗУЕМЫЕ В КНИГЕ[1]

Знак $\rightleftharpoons$ означает «равно по определению» и «эквивалентно по определению».

1. Множество. Слова *множество, совокупность, семейство, класс* употребляются нами как синонимы. Элементы множеств сами являются множествами, элементы элементов множеств также являются множествами и т. д.

Отметим операцию объединения множеств: $\bigcup X \rightleftharpoons \{x \mid \exists y \in X (x \in y)\}$.

Не следует путать знаки $\in$ и $\subseteq$.

Когда говорят об *объекте* или *математическом объекте*, то обычно имеют в виду какое-то множество.

Основной принцип наивной теории множеств (только ее мы и будем использовать) состоит в том, что для любого свойства (высказывания, предиката) $\varphi(x)$ с переменной x существует совокупность всех тех и только тех множеств x_0, которые обладают этим свойством, т. е. для которых выполняется (истинно) $\varphi(x_0)$. Эта совокупность множеств записывается в виде $\{x \mid \varphi(x)\}$.

2. Функция. Слова *функция, отображение* употребляются нами как синонимы. Функция, как правило, обозначается одной буквой, например, f или g и записывается в виде $f(\cdot)$ или $f : X \to Y$, где X и Y некоторые множества. Иногда бывает удобно записывать функцию в виде $y = f(x)$ или даже $f(x)$; в этом случае переменная x пробегает ее область определения, а переменная y — область значений. Если функция обозначена, например f, то ее область определения обозначается $D(f)$, а область значений — $R(f)$.

Запись $f: X \to Y$ имеет одну двусмысленность: как правило, она означает, что $D(f) = X$. Однако иногда $D(f) \subset X$. Например, часто записывают $\frac{1}{x}$: $R \to R$, хотя

[1] Подробнее изложение всего ниже следующего материала читатель найдет в учебных пособиях по основным математическим курсам для пединститутов, например:

1. *Куликов Л. Я.* Алгебра и теория чисел.— М.: Высшая школа, 1979.—559 с.
2. *Базылев В. Т., Дуничев К. И., Иваницкая В. П.* Геометрия: Учеб. пособие для студентов 1 курса физ.-мат. фак-тов пед. ин-тов.— М.: Просвещение, 1974.—352 с.

Базылев В. Т., Дуничев К. И. Геометрия: Учеб. пособие для студентов физ.-мат. фак-тов пед. ин-тов.— М.: Просвещение, 1975.—368 с.

3. *Нечаев В. И.* Числовые системы.— М.: Просвещение, 1975.—192 с.
4. *Колмогоров А. Н., Драгалин А. Г.* Введение в математическую логику.— М.: Изд-во МГУ, 1982.—112 с.

В дальнейшем ссылки на эти пособия будут содержать только фамилию автора и страницы.

$D\left(\frac{1}{x}\right) = \mathbb{R}\setminus\{0\}$. Если нужно подчеркнуть, что функция f не обязательно определена на всем множестве X, т. е. $D(f)\subseteq X$, то пишут $f: X\rightsquigarrow Y$. В этом случае функцию f называют *частично определенной*. Конечно, всегда определяется соотношение $R(f)\subseteq Y$ и часто выполняется $R(f)\subset Y$. Запись $f: x\mapsto y$ означает, что функция f переводит объект x в объект y, т. е. $f(x)=y$. Запись $f|X_1$, где $X_1\subseteq D(f)$, по определению обозначает сужение функции $f: X\to Y$ на множество X_1, т. е. $(f|X_1)(x)=f(x)$, где $x\in X_1$ и $(f|X_1): X_1\to Y$.

3. Декартово произведение и множество всех подмножеств. Структура. Топология. Над множествами часто выполняют две операции: одна из них — декартово произведение семейства множеств $\{X_\alpha\}$, где индекс α пробегает некоторое множество индексов A, а вторая — «булеан».

А именно, *декартовым произведением* называется множество

$$\{f: A\to \bigcup_{\alpha\in A} X_\alpha \mid \forall\alpha\in A(f(\alpha)\in X_\alpha)\}.$$

Оно обозначается $\prod_{\alpha\in A} X_\alpha$. Если $A=\{1, 2\}$, то $\prod_{\alpha\in A} X_\alpha$ равно декартову произведению $X_1\times X_2$ двух множеств X_1 и X_2. Если $A=\{1, 2, \ldots, n\}$, то $\prod_{\alpha\in A} X_\alpha$ равно декартову произведению $X_1\times X_2\times\ldots\times X_n$ ровно n множеств $X_1, \ldots, X_n$.

Булеаном множества X (множеством всех подмножеств множества X) называется множество $\{y \mid y\subseteq X\} = \{y \mid \forall z(z\in y\Rightarrow z\in X)\}$. Оно обозначается $\mathcal{P}(X)$.

Если X — произвольное множество, то *отношением* на X называется любое подмножество множества $X^n = \underbrace{X\times\ldots\times X}_{n}$, где n произвольное натуральное число.

Иногда термин «отношение» употребляется в более широком смысле: отношением на X называется произвольное подмножество любого множества Y, которое получается из множества X последовательным применением конечное число раз, начиная с X, двух операций: декартова произведения и булеана. Например, множество Y может равняться $\mathcal{P}(X)$, или $(\mathcal{P}(X))^n$, или $X\times\mathcal{P}(X)$, или $X^k\times(\mathcal{P}(X^n))^m$, где n, m, k произвольные натуральные числа.

Структурой (системой) с носителем X называется конечное множество, состоящее из множества X и какого-то набора операций и отношений как на самом множестве X, так и, может быть, еще на некоторых вспомогательных множествах $C_1, \ldots, C_k$. Последние множества называются *константами структуры*. Структура записывается в виде набора множеств $\langle X, f_1, \ldots, f_n, P_1, \ldots, P_m, C_1, \ldots, C_k\rangle$, где X исходное множество, называемое *носителем структуры*, $f_1, \ldots, f_n$ — операции на X (или на X и вспомогательных множествах $C_1, \ldots, C_k$) и $P_1, \ldots, P_m$ — отношения на множествах $X, C_1, \ldots, C_k$. Элементы этого набора обычно удовлетворяют некоторым условиям (аксиомам). Уже то обстоятельство, что все множества f_i являются операциями, а все множества P_j — отношениями, означает, что множества (объекты) f_i и P_j удовлетворяют определенным условиям (аксиомам).

В курсе алгебры имеют дело с частным случаем понятия структуры — понятием алгебраической системы. В этом курсе приводятся многочисленные примеры алгебраических систем, которые одновременно являются и примерами структур.

Другим типичным примером структуры является любой набор множеств вида $\langle X, \mathfrak{T}\rangle$, где $\mathfrak{T}\subseteq\mathcal{P}(X)$ (и в этом смысле $\mathcal{T}$ можно назвать отношением на X), для которого выполняются следующие три аксиомы:

1) $\varnothing, X\in\mathfrak{T}$;

2) $(X_\alpha, \ldots, X_\alpha, \ldots \in\mathfrak{T})\Rightarrow(\bigcup_\alpha X_\alpha)\in\mathfrak{T}$ (или точнее: $(\{X_\alpha\}\subseteq\mathfrak{T})\Rightarrow(\bigcup_\alpha X_\alpha)\in\mathfrak{T}$);

3) $(X, Y\in\mathfrak{T})\Rightarrow X\cap Y\in\mathfrak{T}$.

Такой набор множеств $\langle X, \mathfrak{T}\rangle$ называется *топологическим пространством*, множество $\mathfrak{T}$ называется топологией этого пространства. Элементы множества $\mathfrak{T}$ называются открытыми множествами данного топологического пространства $\langle X, \mathfrak{T}\rangle$. Буква $\mathcal{O}$ употребляется в книге только для обозначения открытых множеств топологических пространств.

Например, наборы $\langle X, \{\varnothing, X\}\rangle$ и $\langle X, \mathcal{P}(X)\rangle$ являются топологическими пространствами. Первое из них называется *антидискретным*, а второе — *дискретным топологическим пространством*. Соответственно топология $\{\varnothing, X\}$ называется антидискретной, а топология $\mathcal{P}(X)$ — *дискретной*. Поэтому не дискретная топология в множестве X — это такая топология $\mathfrak{T}$, для которой $\mathfrak{T}\neq\mathcal{P}(X)$.

Пересечение любого семейства топологий в множестве X является топологией в множестве X.

4. База. Непрерывная функция. *Базой* топологии в множестве X называется такое семейство его подмножеств $\mathcal{T}$ (т. е. $\mathcal{T}\subseteq\mathcal{P}(X)$), для которого выполняется условие: $\bigcup\mathcal{T}=X$ и $\forall\mathcal{O}_1, \mathcal{O}_2\in\mathcal{T}\,\forall x\in(\mathcal{O}_1\cap\mathcal{O}_2)\exists\mathcal{O}\in\mathcal{T}(x\in\mathcal{O}\subseteq\mathcal{O}_1\cap\mathcal{O}_2)$. Любая база топологии в множестве X задает топологию в нем, определяемую как семейство вида $\mathfrak{T}=\{\bigcup Y \mid Y\subseteq\mathcal{T}\}$. Проверьте, что это семейство множеств действительно является топологией в множестве X. Любое непустое семейство $\mathcal{T}_0$ подмножеств множества X задает топологию в X, определяемую базой следующего вида

$$\mathcal{T}=\{Y_1\cap\ldots\cap Y_n \mid Y_1, \ldots, Y_n\in\mathcal{T}_0,\ n\geqslant 0\}$$ (проверьте).

Семейство $\mathcal{T}_0$ называется *предбазой* топологии в множестве X. Одна и та же топология задается, как правило, многими разными базами и предбазами.

Если в множестве X задано отношение порядка $\leqslant$, то топология в X, определяемая предбазой топологии, состоящей из всех множеств вида $\{x\in X \mid a<x\}$ и $\{x\in X \mid x<b\}$; где a, b — любые фиксированные элементы из X, называется *порядковой* топологией. Соответствующая этой предбазе база состоит из всех интервалов в X, т. е. всех множеств вида $\{x\in X \mid a<x<b\}$. В множестве вещественных чисел $\mathbb{R}$ всегда рассматривается *интервальная* топология. Совпадает ли с порядковой топологией в $\mathbb{R}$ топология в $\mathbb{R}$, определенная базой, состоящей из всех множеств вида $\{x\in\mathbb{R} \mid a<x\}$?

В метрическом пространстве $\langle X, \rho\rangle$, где ρ — метрика в X, всегда рассматривается топология, определяемая базой, состоящей из всех *шаров*, т. е. всех множеств вида $\mathcal{O}_{a,\,\xi}=\{x\in X \mid \rho(a, x)<\xi\}$, где $a\in X$ и вещественное число $\xi>0$.

Функция $f: X\to Y$, где множества X и Y имеют соответственно топологии $\mathfrak{T}_X$ и $\mathfrak{T}_Y$, называется непрерывной, если

$$\forall\mathcal{O}\in\mathfrak{T}_Y(f^{-1}(\mathcal{O})\in\mathfrak{T}_X).$$

Если X и Y метрические пространства и функция f вида $f: X\to Y$, то f непрерывна (в выше указанном смысле) тогда и только тогда, когда

$$\forall a\in X\,\forall\varepsilon>0\,\exists\delta>0(\rho_X(a, x)<\delta\Rightarrow\rho_Y(f(a), f(x))<\varepsilon)$$

(проверьте).

Интервальная топология в $\mathbb{R}$ совпадает с топологией, определяемой в $\mathbb{R}$ обычной метрикой $\rho(x, y)=|x-y|$.

Отметим, что существуют такие топологические пространства $\langle X, \mathfrak{T}\rangle$, для которых топология $\mathfrak{T}$ не определяется никакой метрикой в X. Такие топологические пространства называются *неметризуемыми*. В элементарной математике неметризуемые топологические пространства не встречаются. Поэтому можно считать, что все рассматриваемые в книге топологические пространства являются метрическими и топологии в них определяются базами, являющимися множествами всех шаров; в этом смысле понятие топологического пространства является для нас излишним.

Гомеоморфизмом топологических пространств $\langle X, \mathfrak{T}_X\rangle$ и $\langle Y, \mathfrak{T}_Y\rangle$ называется любая биекция $f: X \leftrightarrow Y$, для которой функции f и f^{-1} непрерывны.

Топологическое пространство X называется *отделимым* (хаусдорфовым), если выполняется условие:

$$\forall x,\ y \in X (x \neq y \Rightarrow \exists \mathcal{O}_1,\ \mathcal{O}_2 \in \mathfrak{T}_X (x \in \mathcal{O}_1 \wedge y \in \mathcal{O}_2 \wedge \mathcal{O}_1 \cap \mathcal{O}_2 = \varnothing)).$$

Топологическое пространство называется *компактным* (компактом), если из всякого покрытия его открытыми множествами можно выделить его конечное подпокрытие.

Если $f: X \leftrightarrow Y$ непрерывная биекция отделимого топологического пространства X в компакт Y, то f гомеоморфизм (докажите).

5. Индуцирование топологии. Часто используют следующие три способа определения (задания) топологии.

При первом способе множество X является подмножеством множества Y и в Y уже фиксирована какая-то топология $\mathfrak{T}_Y$. Тогда в X по определению рассматривается топология $\mathfrak{T}_X = \{\mathcal{O} \cap X \mid \mathcal{O} \in \mathfrak{T}_Y\}$. Она называется *топологией, индуцированной множеством* Y.

При втором способе множество X является областью определения какой-то функции f вида $f: X \to Y$ и в множестве Y уже фиксирована какая-то топология $\mathfrak{T}_Y$. Тогда в множестве X по определению рассматривается топология $\mathfrak{T}_X = \{f^{-1}(\mathcal{O}) \mid \mathcal{O} \in \mathfrak{T}_Y\}$. Она называется *топологией, индуцированной функцией* f. Например, если $X \subseteq Y$ и $f(x) = x$, $f: X \to Y$, то топологии в множестве X, индуцированные функцией f и множеством Y, совпадают.

Поэтому первый способ определения топологии является частным случаем второго способа. Если в множестве X заданы две топологии $\mathfrak{T}_1$ и $\mathfrak{T}_2$ и $\mathfrak{T}_1 \subseteq \mathfrak{T}_2$, то топология $\mathfrak{T}_1$ называется *более слабой (меньшей)*, чем топология $\mathfrak{T}_2$, а топология $\mathfrak{T}_2$ называется *более сильной (большей)*, чем топология $\mathfrak{T}_1$. Ясно, что в любом множестве X самой слабой является антидискретная топология и самой сильной — дискретная топология.

Топология, индуцированная в множество X, с помощью функции f, где $f: X \to Y$, совпадает со слабейшей (наименьшей) топологией в X, относительно которой эта функция непрерывна (проверьте).

Второй способ определения топологии можно обобщить следующим образом. Пусть $\{f_\alpha\}$ (где индекс α пробегает множество A) — семейство функций вида $f_\alpha: X \to Y_\alpha$ и в каждом из множеств Y_α фиксирована какая-то топология $\mathfrak{T}_\alpha$. Тогда в множестве X рассматривается топология, определяемая предбазой $\mathcal{T}_0 = \{f_\alpha^{-1}(\mathcal{O}) \mid \alpha \in A \wedge \mathcal{O} \in \mathfrak{T}_\alpha\}$. Она называется *топологией, индуцированной семейством функций* $\{f_\alpha \mid \alpha \in A\}$.

Топология, индуцируемая в множество X, семейством функций $\{f_\alpha \mid \alpha \in A\}$ совпадает со слабейшей (наименьшей) топологией в X, относительно которой все функции непрерывны.

Например, в декартовом произведении $X = \prod_{\alpha \in A} X_\alpha$ семейства множеств $\{X_\alpha\}$

(если во всех множествах X_α фиксированы какие-то топологии $\mathfrak{T}_\alpha$) всегда рассматривается топология, индуцируемая следующим семейством функций (называемых проекциями):

$$P_\alpha: X \to X_\alpha,\ P_\alpha(f) = f(\alpha),\ \text{где } f \in X.$$

Она называется *топологией произведения*. Базой топологии произведения является семейство всех множеств вида

$$\{x_\alpha \mid x_{\alpha_1} \in \mathcal{O}_{\alpha_1} \wedge \ldots \wedge x_{\alpha_n} \in \mathcal{O}_{\alpha_n}\},$$

где $\{\alpha_1, \ldots, \alpha_n\}$ какое-то конечное множество индексов из A и $\mathcal{O}_{\alpha_1} \in \mathfrak{T}_{\alpha_1}, \ldots, \mathcal{O}_{\alpha_n} \in \mathfrak{T}_{\alpha_n}$ и $n \geqslant 0$.

Например, в плоскости $\boldsymbol{R}^2$ всегда рассматривается топология произведения. Точка $\begin{pmatrix} a \\ b \end{pmatrix}$ из $\boldsymbol{R}^2$ отождествляется с комплексным числом $a+bi$, следовательно, множества $\boldsymbol{R}^2$ и $\boldsymbol{C}$ также отождествляются. Поэтому в множестве комплексных чисел $\boldsymbol{C}$ также всегда рассматривается топология произведения.

В подмножествах множеств $\boldsymbol{R}$ и $\boldsymbol{C}$, как правило, рассматривается топология, индуцируемая соответственно множествами $\boldsymbol{R}$ и $\boldsymbol{C}$. Поэтому топологии в множествах $F_2 = \{0, 1\}$, $F_m = \{0, 1, 2, \ldots, m-1\}$, $\boldsymbol{N}$, $\boldsymbol{Z}$, $\boldsymbol{Q}$ и $\boldsymbol{S}$ (где S — окружность единичного радиуса с центром в точке $\overline{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$) задаются по первому способу определения топологии. Эти топологии в множествах F_2, F_m, $\boldsymbol{N}$ и $\boldsymbol{Z}$ являются дискретными и совпадают с порядковыми топологиями в них. Топология, индуцируемая в $\boldsymbol{Q}$ множеством $\boldsymbol{R}$, также совпадает с интервальной топологией в $\boldsymbol{Q}$, но не является дискретной.

В множествах $\boldsymbol{R}^3$ и $\boldsymbol{R}^n$ (где $n \geqslant 1$) всегда рассматривается топология произведения, а в подмножествах множества $\boldsymbol{R}^n$ — топология, индуцируемая объемлющим множеством $\boldsymbol{R}^n$. Например, в S^2 (сфере единичного радиуса в R^3 с центром в точке $\overline{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$) всегда рассматривается индуцированная топология.

Третий способ определения топологии в множестве X обычно применяется в том случае, если X является областью значений какой-то функции f вида $f: Y \to X$ и в множестве Y уже фиксирована топология $\mathfrak{T}_Y$. Тогда в множестве X определяется топология $\mathfrak{T}_X = \{\mathcal{O} \subseteq X \mid f^{-1}(\mathcal{O}) \in \mathfrak{T}_Y\}$. Она называется топологией, индуцированной функцией f в ее область значений.

Топология, индуцируемая в множество X с помощью функции f вида $f: Y \to X$, совпадает с сильнейшей (наибольшей) топологией в X, относительно которой функция f непрерывна (проверьте).

Важным применением третьего способа определения топологии является случай, когда топология определяется в фактор-множестве $X/\sim$, т. е. в множестве X, факторизованном некоторым отношением эквивалентности $\sim$. А именно, *каноническим отображением* называется функция $\varphi: X \to (X/\sim)$, определяемая равенством $\varphi(x) = [x]$, где $[x]$ класс эквивалентности, содержащий элемент x. В множестве $X/\sim$ всегда определяется наибольшая топология, относительно которой каноническое отображение непрерывно. Иными словами, подмножество $\mathcal{O}$ множества $(X/\sim)$ по определению открыто в нем тогда и только тогда, когда $\varphi^{-1}(\mathcal{O})$ открыто в X.

Такая топология встречается, например, при изучении следующего важного фактор-множества. Положим $x \sim y$, если $(x-y) \in \boldsymbol{Z}$ и $x,\ y \in \boldsymbol{R}$. Образуем фактор-множество $T = \boldsymbol{R}/\sim$ и индуцируем в T топологию из $\boldsymbol{R}$ с помощью канонического

отображения φ: $R \to T$. Функция $f(x) = e^{2\pi i x}$ является гомеоморфизмом топологических пространств T и S, что существенно используется при построении тригонометрических функций.

6. Компактные и связные топологические пространства. Все обычные топологические пространства и, в частности, пространства, встречающиеся в элементарной математике, отделимы. Поэтому в дальнейшем все топологические пространства предполагаются отделимыми.

Любой отрезок $[a, b]$ в R с топологией, индуцированной множеством R, является *компактом*. Любой интервал $]a, b[$ в R, где $-\infty \leqslant a < b \leqslant +\infty$, с топологией, индуцированной множеством R, не является компактом. Из курса анализа известно, что подмножество $X \subseteq R^n$ с топологией, индуцированной в нем множеством R^n, является компактом в том и только том случае, если X замкнуто и ограничено. Топологическое пространство $\langle X, \mathfrak{T} \rangle$ называется *локально компактным*, если $\forall x \in X \exists \theta_x \in \mathfrak{T} (x \in \theta,\ \overline{\theta}$ — компактно); запись $\overline{\theta}$ обозначает замыкание множества θ (*Замыканием множества* Y называется множество всех его точек прикосновения). Конечно, $\overline{\theta}$ рассматривается с топологией, индуцируемой множеством X.

Как правило, топологические пространства, встречающиеся в элементарной математике, локально компактны. Например, таковы топологические пространства R, C, S, S^2, Z, N.

В любом компактном пространстве X выполняется свойство: из всякой последовательности $\{x_n\}$: $N \to X$ можно выбрать сходящуюся подпоследовательность $\{x_{n_k}\}$: $N \to X$. Здесь $\{n_k\}: N \to N$ и $\{x_{n_k}\} = \{x_n\} \circ \{n_k\}$.

Следующий абзац используется только в главе V. При первом чтении его можно пропустить.

Обозначим $C([0, 1], R)$ множество всех непрерывных функций вида f: $[0, 1] \to R$. Определим метрику $\rho(f_1, f_2) = \sup\{|f_1(x) - f_2(x)| \mid x \in [0, 1]\}$. Набор $\langle C([0, 1], R), \rho \rangle$ — метрическое пространство. В любой наперед заданной ε-окрестности нуля этого метрического пространства содержится последовательность $\varepsilon \cdot \frac{n}{1+n^2}$. При этом для нее не существует сходящейся подпоследовательности; поэтому пространство $C([0, 1], R)$ не локально компактно. Замечательная теорема Рисса утверждает: пусть в векторном пространстве $\mathfrak{X}$ над полем скаляров R или C определена топология, относительно которой обычные операции над векторами непрерывны, тогда локальная компактность $\mathfrak{X}$ влечет конечномерность $\mathfrak{X}$. Поскольку векторное пространство $C([0, 1], R)$ с покоординатно определенными операциями явно не конечномерно, то по теореме Рисса вновь получаем, что оно не локально компактно.

Топологическое пространство $\langle X, \mathfrak{T} \rangle$ называется *связным*, если оно обладает свойством:

$$\neg \exists \theta_1, \theta_2 \in \mathfrak{T} (X = \theta_1 \cup \theta_2 \wedge \theta_1 \cap \theta_2 = \varnothing).$$

Множество Y в топологическом пространстве $\langle X, \mathfrak{T} \rangle$ (т. е. $Y \subseteq X$) называется *связным*, если

$$\neg \exists \theta_1, \theta_2 \in \mathfrak{T} (Y = (\theta_1 \cap Y) \cup (\theta_2 \cap Y) \wedge ((\theta_1 \cap Y) \cap (\theta_2 \cap Y) = \varnothing)).$$

Последнее условие эквивалентно тому, что Y, как топологическое пространство с топологией, индуцированной в нем множеством X, связно.

Обозначим $R_{\geqslant 0}$ и $R_{>0}$ соответственно множества всех неотрицательных чисел и всех строго положительных чисел. Аналогичен смысл обозначений $R_{\leqslant 0}$ и $R_{<0}$.

Множества R, $R_{\geqslant 0}$, $R_{>0}$, $R_{\leqslant 0}$, $R_{<0}$, (a, b), S и S^2 связны в соответствующих топологических пространствах R и R^2, а множества $[a, b] \cup [c, d]$ и $]a, b[\cup]c, d[$, где $b < c$, не связны в пространстве R.

7. Тело. Модуль. Алгебра. В курсе алгебры определяются понятия группы, кольца, поля. Определение тела получается из определения поля (см. [1], с. 146), заменой слова «коммутативное кольцо» на «кольцо». Обозначим R^+ множество всех вещественных чисел R с операцией сложения, т. е. аддитивную группу поля вещественных чисел. Обозначим R^* множество всех вещественных чисел без нуля с операцией умножения, т. е. мультипликативную группу поля вещественных чисел. Обозначим R' множество вещественных чисел, строго больших нуля, с операцией умножения — это мультипликативная группа. Обозначим Z^+ множество целых чисел с обычной операцией сложения.

Аналогично: K^+ — обозначение аддитивной группы кольца (или векторного пространства) K. Обозначим K' множество всех обратимых элементов кольца K и K^* множество всех ненулевых элементов кольца K. В этих множествах рассматривается операция умножения индуцированная кольцом K; относительно нее множество K' является группой.

Если в определении векторного пространства (см. [1], с. 245) заменить поле скаляров F на кольцо скаляров F, то получается более широкое понятие *модуля над кольцом* F. *Модуль над полем* F — то же самое, что и векторное пространство над полем F.

Если $\mathfrak{X}$ векторное пространство над полем R и $a, b \in \mathfrak{X}$, то обозначим

$$[a, b] = \{x \in \mathfrak{X} \mid \exists \lambda (0 \leqslant \lambda \leqslant 1 \wedge x = \lambda \cdot a + (1 - \lambda) \cdot b)\} \text{ (отрезок)},$$
$$]a, b[= \{x \in \mathfrak{X} \mid \exists \lambda (0 < \lambda < 1 \wedge x = \lambda \cdot a + (1 - \lambda) \cdot b)\} \text{ (интервал)},$$
$$(a, b) = \{\lambda \cdot (b - a) \mid \lambda \in R\} \text{ (прямая)}.$$

Обозначим $\langle a, b \rangle$ упорядоченную пару точек a, b, ее часто называют вектором.

Элементы множества R^2 записываются в зависимости от ситуации одним из следующих способов: (x, y) или $\langle x, y \rangle$, или $\begin{pmatrix} x \\ y \end{pmatrix}$.

Понятие *линейной алгебры (алгебры) над полем скаляров* F определяется в [1], с. 298). Если в этом определении заменить поле F на кольцо скаляров F, то получится определение *линейной алгебры (алгебры) над кольцом* F.

Условимся записывать *матрицу* $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ в виде (ab), где $a = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ и $b = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$. Конечно, записи $\det A$, $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ и $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ обозначают одно и то же — *определитель матрицы* A.

Отношением порядка называется алгебраическая система вида $\langle X, \leqslant \rangle$, где множество $\leqslant$ является подмножеством множества X^2 и для него выполняются условия:

$$\forall x \in X (\langle x, x \rangle \in \leqslant),$$
$$\forall x, y, z \in X ((\langle x, y \rangle \in \leqslant \wedge \langle y, x \rangle \in \leqslant \Rightarrow x = y) \wedge$$
$$\wedge (\langle x, y \rangle, \langle y, z \rangle \in \leqslant \Rightarrow \langle x, z \rangle \in \leqslant)).$$

Иными словами, отношение порядка на X — это двуместный предикат на множестве X, обладающий обычными свойствами рефлексивности, антисимметричности и транзитивности.

8. Изоморфизмы. Вложение. Морфизм. Среди общематематических понятий одним из центральных является понятие изоморфизма. Если даны две структуры $\langle X, f_1, \ldots, f_n, P_1, \ldots, P_m\rangle$ и $\langle Y, g_1, \ldots, g_n, Q_1, \ldots, Q_m\rangle$, то они называются *изоморфными* в том случае, когда существует биекция ψ: $X \leftrightarrow Y$ между носителями X и Y этих структур, для которой выполняются следующие свойства (пусть для определенности f_1: $X^2 \to X$ и $g_1 : Y^2 \to Y$, и $P_1 \subseteq \mathscr{P}(X)$ и $Q_1 \subseteq \mathscr{P}(Y)$): $\forall x_1, x_2, x_3 (f_1(x_1, x_2) = x_3 \Leftrightarrow g_1(\psi(x_1), \psi(x_2)) = \psi(x_3))$, $\forall w \subseteq X (w \in P_1 \Leftrightarrow \psi(w) \in Q_1)$, где $\psi(w) \rightleftharpoons \{\psi(x) | x \in w\}$, а также аналогичные свойства выполняются для соответствующих друг другу функций $f_i \leftrightarrow g_i$ и соответствующих друг другу отношений $P_j \leftrightarrow Q_j$; здесь $i = 2, \ldots, n$ и $j = 2, \ldots, m$. Сама функция ψ называется изоморфизмом структур $\langle X, \ldots\rangle$ и $\langle Y, \ldots\rangle$. Интуитивно биекция ψ отождествляет элементы x и $\psi(x)$, а также y и $\psi^{-1}(y)$, где $x \in X$ и $y \in Y$.

Если ψ: $X \to Y$ инъективная функция и для нее выполнены свойства из предыдущего определения изоморфных структур, то она называется *вложением структуры* $\langle X, \ldots\rangle$ *в структуру* $\langle Y, \ldots\rangle$. Наличие вложения ψ означает, что структура $\langle X, \ldots\rangle$ отождествляется с частью структуры $\langle Y, \ldots\rangle$; а именно, она отождествляется с частью $R(\psi)$ носителя Y второй структуры.

Если ψ: $X \to Y$ произвольная функция, для которой выполнены свойства из определения изоморфных структур, то она называется *морфизмом структур* $\langle X, \ldots\rangle$ и $\langle Y, \ldots\rangle$. Изоморфизмы, вложения и морфизмы для разных видов (родов) структур широко встречаются в высшей математике. Они имеют соответствующие названия. Некоторые из этих названий приведены в следующей таблице.

№	Вид (род) структур $\langle X, \ldots\rangle$ и $\langle Y, \ldots\rangle$	Название изоморфизма	Название вложения	Название морфизма
1	Топологические пространства $\langle X, \mathscr{T}_X\rangle$ и $\langle Y, \mathscr{T}_Y\rangle$	Гомеоморфизм	Топологическое вложение	Непрерывная функция
2	Метрические пространства $\langle X, \rho_X(\cdot, \cdot)\rangle$ и $\langle Y, \rho_Y(\cdot, \cdot)\rangle$	Метрический изоморфизм	Метрическое вложение	Непрерывная функция
3	Кольца $\langle X, +_X, -_X, \cdot_X, 0_X, 1_X\rangle$ и $\langle Y, +_Y, -_Y, \cdot_Y, 0_Y, 1_Y\rangle$	Кольцевой изоморфизм	Кольцевой мономорфизм	Кольцевой гомоморфизм
4	Поля (тела) $\langle X, +_X, -_X, \cdot_X, {}^{-1_X}, 0_X, 1_X\rangle$ и аналогично $\langle Y, \ldots\rangle$	Изоморфизм	Мономорфизм	Гомоморфизм
5[1]	Векторные пространства, модули $\langle X, +_X, -_X, \omega_\lambda, 0_X, K\rangle$ и аналогично $\langle Y, \ldots\rangle$, где K — поле (тело, кольцо)	Линейный изоморфизм	Линейное вложение	Линейная функция
6	Линейные алгебры $\langle X, +_X, -_X, \omega_\lambda, 0_X, \cdot_X, K\rangle$ и аналогично $\langle Y, \ldots\rangle$	Алгебраический изоморфизм	Алгебраическое вложение	Алгебраический гомоморфизм
7	Линейные алгебры с делением $\langle X, +_X, -_X, \omega_\lambda, \cdot_X, {}^{-1_X}, 0_X, 1_X, K\rangle$ и аналогично $\langle Y, \ldots\rangle$	Алгебраический изоморфизм	Алгебраическое вложение	Алгебраический гомоморфизм
8	Группы $\langle X, \cdot_X, {}^{-1_X}, 1_X\rangle$ или $\langle X, +_X, -_X, 0_X\rangle$ и аналогично $\langle Y, \ldots\rangle$	Групповой изоморфизм	Групповое вложение	Групповой гомоморфизм

[1] По поводу обозначений и понятий из строк 5—7 этой таблицы см. Куликов Л. Я., с. 246, 298, 300.

Глава I

ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ. УГОЛ

ВВЕДЕНИЕ

В математическом курсе для студентов пединститутов «Числовые системы» рассматриваются числа и числовые множества. Мы также остановимся на понятиях числа и числового множества в главе V. А пока будем употреблять эти понятия в том интуитивном смысле, в каком они используются в школе. В главе I мы будем иметь дело с отображениями одних числовых множеств в другие. А именно будем изучать элементарные функции вида f: $R \to R$. Начать, конечно, следует с самых простых и в то же время самых важных функций — линейной, показательной, логарифмической, степенной, косинуса, синуса, арккосинуса и арксинуса. Эти функции мы будем называть *базисными элементарными функциями*.

Первые четыре названия относятся к семействам функций, которые зависят от одного числового параметра: $y = a \cdot x$, $y = a^x$, $y = x^a$, $y = \log_a x$, а последние четыре — к индивидуальным функциям: $y = \cos x$, $y = \sin x$, $y = \arccos x$, $y = \arcsin x$. В этой главе под понятиями «косинус» и «синус», «арккосинус» и «арксинус» удобнее понимать также семейства функций, зависящие от одного числового параметра, а именно семейства вида $y = \cos \frac{x}{a}$ и $y = \sin \frac{x}{a}$, $y = a \cdot \arccos x$ и $y = a \cdot \arcsin x$. Ясно, что функции из одного такого семейства имеют аналогичные между собой свойства.

Элементарные функции определяются на основе базисных элементарных функций путем применения к ним последовательно несколько раз трех операций над функциями: сложения, умножения и композиции функций. Например, элементарными функциями являются функции $y = \cos(x^2)$, $y = (\cos x)^2$, $y = ((\log_2(\cos(x^3))) + 5^{\sin x}) \cdot x^{3/2}$ и т. д. Они получаются из базисных функций соответственно следующими цепочками операций: $y = (\cos x) \circ (x^2)$, $y = (x^2) \circ (\cos x)$, $y_1 = (\cos x) \circ (x^3)$, $y_2 = (\log_2 x) \circ y_1$, $y_3 = (5^x) \circ (\sin x)$, $y_4 = y_2 + y_3$, $y = y_4 \circ (x^{3/2})$.

Три упомянутые операции над функциями назовем *базисными операциями*.

Для каждой из базисных элементарных функций существует много разных способов ее определения. Например, $\ln x$ можно определить как функцию, обратную к e^x, и как функцию вида

$\int_1^x \frac{dt}{t}$, $x>0$. В свою очередь e^x можно определить как сумму ряда $\sum_{n=0}^{\infty} \frac{x^n}{n!}$, и как решение дифференциального уравнения $y'=y$ с начальным условием $y(0)=1$, и как частный случай показательной функции a^x при $a=e \rightleftarrows \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$. Часто поступают в обратном порядке: a^x определяют через e^x и $\ln a$ с помощью равенства $a^x \rightleftarrows e^{x\ln a}$. Степенную функцию x^a с областью определения $\boldsymbol{R}_{>0}$ также можно определить с помощью ряда $x^a=1+a\cdot x+\frac{a(a-1)}{2!}x^2+\sum_{n=3}^{\infty}\frac{a(a-1)\ldots(a-n+1)}{n!}\cdot(x-1)^n$, но тогда приходится делать ограничение $0<x<2$, $a\neq -1, 0, 1, \ldots$ и для таких «особых» значений x и a определять функцию отдельно. Разумеется, x^a можно определить и как $e^{a\ln x}$ и т. д.

Аналогичное, если не большее, разнообразие наблюдается и в способах определения косинуса и синуса, арккосинуса и арксинуса.

Естественно было бы попытаться определить базисные элементарные функции единым, универсальным способом, что позволило бы прояснить их основное содержание и ответить, например, на вопросы: почему эти функции рассматриваются совместно, что у них общего и почему именно эти функции играют выдающуюся роль в математике и ее приложениях?

Такой единый подход возможен в рамках аксиоматического метода и основан на том фундаментальном наблюдении, что каждая из алгебраических базисных элементарных функций, т. е. линейная, показательная, логарифмическая и степенная функции, является непрерывным гомоморфизмом всего лишь двух числовых групп $\boldsymbol{R}^+$ и $\boldsymbol{R}^*$ в себя или друг в друга. Это проверяется тривиальным вычислением на основе определения непрерывного гомоморфизма. Соберем эти непрерывные гомоморфизмы соответствующих групп в таблицу.

Таблица 1

	$\boldsymbol{R}^+$	$\boldsymbol{R}^*$
$\boldsymbol{R}^+$	$a\cdot x$	a^x
$\boldsymbol{R}^*$	$\log_a x (\neq 0)$	x^a

запись в таблице, например, на месте 2,1 обозначает все функции вида $y=\log_a x$ и функцию $y\equiv 0$. Если для какой-то пары групп G_1 и G_2 из множества групп $\{\boldsymbol{R}^+, \boldsymbol{R}^*\}$ возможен непрерывный гомоморфизм вида f: $G_1\to G_2$, который отличается от указанных в таблице 1, то он имеет такое же право называться базисной элементарной функцией, как и перечисленные функции $a\cdot x$, a^x, $\log_a x$, x^a. Иными словами, если понимать термин «базисные элементарные функции» как синоним слов «непрерывные гомоморфизмы», то вопрос состоит в том, не существуют ли какие-то пока не обнаруженные базисные элементарные функции (непрерывные гомоморфизмы). Ответ таков: не существует других непрерывных гомоморфизмов между группами из множества $\{\boldsymbol{R}^+, \boldsymbol{R}^*\}$, кроме тех, которые указаны в таблице 1. Иными словами, не могут быть обнаружены (открыты) новые базисные элементарные функции, по крайней мере если не привлекать групп, отличных от групп $\boldsymbol{R}^+$ и $\boldsymbol{R}^*$.

Для функций $y=\cos x$ и $y=\sin x$ не выполняется условие гомоморфности $f(x*y)=f(x)\times f(y)$, какие бы операции из числа сложения и умножения вещественных чисел мы ни подставили вместо знаков $*$ и $\times$.

Однако по своим свойствам при решении задач и т. д. эти две функции встречаются обычно вместе. Поэтому их естественно рассматривать как две половинки одной функции. А именно рассмотрим функцию $y=f_a(x) \rightleftarrows (\cos x/a,\ \sin x/a)$: $\boldsymbol{R}\to\boldsymbol{C}$. Поскольку $\cos^2 x/a+\sin^2 x/a=1$, то f: $\boldsymbol{R}\to S$, где S — окружность единичного радиуса, $S\subseteq\boldsymbol{C}$. Во множестве S имеется одна естественная групповая операция: умножение точек из S как комплексных чисел. При этом нейтральным элементом является число 1, а симметричным элементом к числу $z=a+bi$ является число $\bar{z}=a-bi$. Получается группа! Она обозначается также S. Легко проверить, что функция f_a: $\boldsymbol{R}^+\to S$ — непрерывный гомоморфизм групп $\boldsymbol{R}^+$ и S. Поэтому функции $\cos x/a=P_1(f_a(x))$ и $\sin x/a=P_2(f_a(x))$ являются двумя половинками непрерывного гомоморфизма f_a. Опять-таки следует спросить: нельзя ли с помощью новой группы S обнаружить какие-то новые «элементарные функции»? При этом опять термин «элементарные функции» понимается как «непрерывные гомоморфизмы групп $\boldsymbol{R}^+$, $\boldsymbol{R}^*$ и S и, кроме того, все те функции, которые получаются из этих гомоморфизмов базисными операциями». В частности, мы спрашиваем: нет ли каких-то непрерывных гомоморфизмов из группы $\boldsymbol{R}^+$ в группу S, отличных от $f_a(\cdot)$, где $a\in\boldsymbol{R}^*$? Ответ: любой непрерывный гомоморфизм вида $\boldsymbol{R}^+\to S$ совпадает с функцией $f_a(x)$ или с функцией $f_0(x)\equiv 1$.

Функции вида $f_a(x)$ называются *экспонентами*. Это связано с тем, что по формуле Эйлера выполняется условие:

$$f_a(x)=e^{ix/a}=\cos\frac{x}{a}+i\sin\frac{x}{a}=f_1(x/a).$$

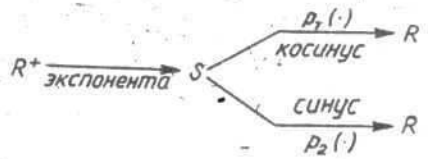


Рис. 1

Итак, если мы хотим определить функции косинус и синус, то нужно рассмотреть непрерывные гомоморфизмы из группы $\boldsymbol{R}^+$ в группу S и затем применить функции $p_1(\cdot)$ и $p_2(\cdot)$ (рис. 1).

Поэтому план изложения в каждом параграфе этой главы примерно таков: сначала аксиоматически определяются базисные элементарные функции, затем из аксиом выводятся их основные свойства; после этого доказывается существование и единственность этих аксиоматически определенных объектов.

Отметим, что такой план аксиоматического изложения материала достаточно традиционен для пединститутских курсов алгебры, геометрии и математического анализа.

§ 1. ЛИНЕЙНАЯ ФУНКЦИЯ

1. Аксиоматическое определение линейной функции.

Определение 1. **Линейной функцией называется любой непрерывный гомоморфизм группы $\boldsymbol{R}^+$ в себя, т. е. любая функция f из множества $\boldsymbol{R}$ в себя, обладающая свойствами:**

1) $f(x+y)=f(x)+f(y)$ **для всех** $x, y\in\boldsymbol{R}$;

2) f **непрерывна.**

Определение 1′. **Линейной функцией с коэффициентом** a, **где** $a\in\boldsymbol{R}$, **называется любая линейная функция, обладающая дополнительным свойством:**

3) $f(1)=a$.

Линейную функцию с коэффициентом a будем обозначать $l_a(x)$.

2. Свойства линейной функции. Всюду ниже f — произвольная линейная функция. При перечислении свойств f продолжим нумерацию из определений 1 и 1′.

4) $f(0)=0$.

▷ По свойству 1 $f(0)=f(0+0)=f(0)+f(0)$. Поэтому $f(0)=0$. ■

5) $f(-x)=-f(x)$.

▷ $f(-x)+f(x)=f(-x+x)=f(0)=0$. Поэтому $f(-x)=-f(x)$. ■

Следующее свойство f связано с умножением в $\boldsymbol{R}$. Оно довольно неожиданно, так как в самом определении линейной функции ничего не говорится об операции умножения, а используется только операция сложения:

6) $f(\lambda x)=\lambda f(x)$ **для всех** $\lambda, x\in\boldsymbol{R}$.

▷ а) Если $\lambda=n\in\boldsymbol{N}$, то равенство $f(nx)=nf(x)$ легко доказывается индукцией.

б) Если $\lambda=-n$, $n\in\boldsymbol{N}$, то из п. а и свойства 5 следует, что

$$f((-n)x)=f(-nx)=-f(nx)=(-n)f(x).$$

в) Если $\lambda=\frac{1}{n}$, $n\in\boldsymbol{N}$, то $f(x)=f\left(n\left(\frac{1}{n}x\right)\right)=nf\left(\frac{1}{n}x\right)$, и поэтому $f\left(\frac{1}{n}x\right)=\frac{1}{n}f(x)$.

г) Если

$$\lambda=r\in\boldsymbol{Q},$$

то

$$r=\frac{m}{n}\ (m\in\boldsymbol{Z},\ n\in\boldsymbol{Q})$$

и

$f(rx)=f\left(m\left(\frac{1}{n}x\right)\right)=\frac{m}{n}f(x)=rf(x)$. Наконец, пусть λ иррационально. Возьмем любую последовательность рациональных чисел $(r_n)_{n=1}^{\infty}$, сходящуюся к λ. Тогда $f(\lambda x)=f((\lim\limits_{n\to\infty} r_n)x)=f(\lim\limits_{n\to\infty}(r_n x))=$

$=\lim f(r_n x)=\lim\limits_{n\to\infty}(r_n f(x))=(\lim\limits_{n\to\infty} r_n)f(x)=\lambda f(x)$. ■

Заметим, что непрерывность f нужна только в третьем равенстве этой цепочки.

3. Теорема существования и единственности линейной функции.

▷ С существованием линейных функций никаких трудностей не возникает. Действительно, фиксируем число $a\in\boldsymbol{R}$ и определим функцию $f(\cdot)$ равенством $f(x)=a\cdot x$ для всех чисел $x\in\boldsymbol{R}$. Легко видеть, что таким образом определенная функция $f(\cdot)$ удовлетворяет определению 1′ и, следовательно, является линейной функцией.

Единственность же функции $f(\cdot)$, обладающей свойствами 1—3, следует из свойства 6.

Действительно, $f(x)\equiv f(1\cdot x)\equiv x\cdot f(1)=a\cdot x$, и если g также обладает свойствами 1—3, то $g(x)\equiv a\cdot x$, т. е. $f(x)=g(x)$ при всех $x\in\boldsymbol{R}$. Тем самым доказана следующая теорема. ■

Теорема 1. ***Пусть число*** $a\in\boldsymbol{R}$. ***Существует единственная линейная функция*** $l_a(x)$ ***с коэффициентом*** a. ***При этом*** $l_a(x)=a\cdot x$ ***для всех чисел*** $x\in\boldsymbol{R}$.

Замечание. В школьной математике линейной называют функцию $y=a\cdot x+b$. В высшей математике ее называют аффинной, линейной или, короче, аффинной функцией. Подробнее этот вопрос обсуждается в главе II. Во всяком случае функция $y=a\cdot x+b$ при $b\neq 0$ не удовлетворяет определению 1.

Упражнение 1. *Докажите, что линейная функция* $f(\cdot)$ *является изоморфизмом тогда и только тогда, когда* $f(1)\neq 0$. *Укажите в этом случае явный вид обратной к* $f(\cdot)$ *функции и покажите, что она также линейная.*

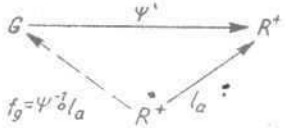


Рис. 2

Следствие 1. Если топологическая группа G изоморфна группе $\boldsymbol{R}^+$, то для всякого $g \in G$ существует единственный непрерывный гомоморфизм вида f_g: $\boldsymbol{R}^+ \to G$, для которого $f_g(1)=g$. Если g не является нейтральным элементом в G, то f_g — изоморфизм.

▷ Пусть ψ: $G \to \boldsymbol{R}^+$ некоторый изоморфизм и $\psi(g) \neq a \in \boldsymbol{R}$. Определим f_g: $\boldsymbol{R}^+ \to G$ равенством $f_g = \psi^{-1} \circ l_a$, что очень удобно изобразить как «построение» третьей стороны в следующем треугольнике, «стороны» которого являются непрерывными гомоморфизмами (рис. 2).

Тогда f_g как композиция непрерывных гомоморфизмов также является непрерывным гомоморфизмом, и если g не нейтрален, то $a=\psi(g)\neq 0$. Поэтому l_a, как и ψ, изоморфизм. Следовательно, f_g — изоморфизм. Кроме того, $f_g(1)=\psi^{-1}(l_a(1))=\psi^{-1}(a)=g$.

Если существует еще один непрерывный гомоморфизм h_g: $\boldsymbol{R} \to G$ со свойством $h_g(1)=g$, то композиция $\psi \circ h_g$: $\boldsymbol{R}^+ \to \boldsymbol{R}^+$ есть непрерывный гомоморфизм со свойствами $\psi \circ h_g(1)=a$, и по теореме 1 $\psi \circ h_g = l_a$, т. е. $h_g = \psi^{-1} \circ l_a = f_g$. ■

§ 2. ПОКАЗАТЕЛЬНАЯ ФУНКЦИЯ

1. Аксиоматическое определение показательной функции.

Определение 2. **Показательной функцией называется любой непрерывный гомоморфизм из группы $\boldsymbol{R}^+$ в группу $\boldsymbol{R}$, т. е. любая функция $f(\cdot)$ из множества $\boldsymbol{R}$ в множество $\boldsymbol{R}_{>0}$, обладающая свойствами:**

1) $f(x+y)=f(x)f(y)$ для всех $x, y \in \boldsymbol{R}$;
2) $f(\cdot)$ непрерывна.

Определение 2′. **Показательной функцией с основанием a, где $a>0$, называется показательная функция, обладающая дополнительным свойством: $f(1)=a$, т. е. любая функция f из $\boldsymbol{R}$ в $\boldsymbol{R}_{>0}$ со свойствами:**

1) $f(x+y)=f(x)f(y)$ для всех $x, y \in \boldsymbol{R}$;
2) f непрерывна;
3) $f(1)=a$.

Любую показательную функцию с основанием a обозначим a^x. Подчеркнем, что пока не выяснено, существует ли хотя бы одна функция a^x, а если таковая существует, то единственная ли она. Положительный ответ будет дан в п. 3 этого параграфа. После этого a^x становится обозначением одной определенной функции. Но сначала установим некоторые свойства аксиоматически определенного объекта — функции a^x.

Мы не определили показательную функцию с основанием $a<0$, так как такая функция была бы определена не при всех значениях аргумента x. Например, $(-2)^{1/2}$ не является вещественным числом.

2. Свойства показательной функции. Начнем со свойств 1—3 из определения 2′ и перепишем их, используя для показательной функции введенное обозначение a^x:

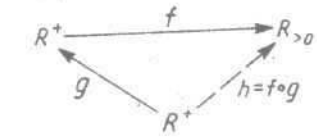


Рис. 3

1) $a^{x+y}=a^x \cdot a^y$;
2) a^x непрерывна;
3) $a^1=a$;
4) $a^0=1$.

▷ $a^0=a^{0+0}=a^0 \cdot a^0$, поэтому $a^0(a^0-1)=0$. Поскольку любое значение показательной функции по определению положительно, то $a^0 \neq 0$, и следовательно, $a^0=1$. ■

5) $a^{-x}=(a^x)^{-1}$.

▷ $a^{-x} \cdot a^x = a^{-x+x}=a^0=1$. Поэтому $a^{-x}=\dfrac{1}{a^x}=(a^x)^{-1}$. ■

6) $(a^b)^x=a^{bx}$ для всех $b, x \in R$.

▷ Рассмотрим при фиксированном b «треугольник» из непрерывных гомоморфизмов (рис. 3), где f — показательная функция с основанием a, т. е. $f(x)=a^x$, и g — линейная функция с коэффициентом b, т. е. $g(x)=b \cdot x$.

Тогда $h=f \circ g$ как композиция непрерывных гомоморфизмов является непрерывным гомоморфизмом из R^+ в $\boldsymbol{R}$, т. е. является показательной функцией с основанием $h(1)=f(g(1))=f(b)=a^b$.

Поэтому $h(x)=(a^b)^x$. С другой стороны, $h(x)=f(g(x))=f(bx)=a^{bx}$. ■

7) Функция a^x монотонно возрастает, если $a>1$, и монотонно убывает, если $0<a<1$.

▷ Разберем случай $a>1$. Сначала докажем, что $a^r>1$ при $r \in Q_{>0}$.

Если $r=\dfrac{1}{n}$, $n \in \boldsymbol{N}$, то, предположив, что $a^{\frac{1}{n}} \leqslant 1$, получим, что $a=\left(a^{\frac{1}{n}}\right)^n \leqslant 1$, так как произведение n чисел из $[0; 1]$ лежит в $[0; 1]$. Это противоречит условию $a>1$. Если $r=\dfrac{m}{n}(m, n \in \boldsymbol{N})$, то $a^r=\left(a^{\frac{1}{n}}\right)^m>1$. Докажем, что $a^x>1$ при любых $x>0$. Если $x_0>0$ и $x_0=\overline{\lim}\, r_n$, где $r_n \in \boldsymbol{Q}_{>0}$, то по непрерывности a^x получаем, что $a^{x_0}=\lim\limits_n a^{r_n} \geqslant 1$.

Если предположить, что $a^{x_0}=1$, то аналогично началу этого доказательства можно показать, что $a^{rx_0}=1$ при всех $r \in \boldsymbol{Q}$. Так как множество $\{r \cdot x_0 \mid r \in \boldsymbol{Q}\}$ всюду плотно на $\boldsymbol{R}$, то непрерывность функции a^x влечет, что она тождественно равна единице, что противоречит условию $a>1$. Следовательно, $a^{x_0}>1$, и теперь легко доказать возрастание a^x при $a>1$: если $x>y$, то $x=y+z$, где $z>0$ и $a^x=a^{y+z}=a^y \cdot a^z > a^y$, так как $a^z>1$ и $a^y>0$. ■

Упражнение 2. *Докажите, что:*

а) множество $\{r \cdot x_0 \mid r \in \boldsymbol{Q}\}$, где $x_0 \neq 0$, всюду плотно на $\boldsymbol{R}$;

б) если $a^{x_0}=1$, *где* $x_0\neq 0$, *то* $a^{rx_0}=1$ *для всех* $r\in\boldsymbol{Q}$;

в) 1^x *— функция, тождественно равная единице;*

г) свойство 7 *выполняется и при условии* $0<a<1$.

8) **Показательная функция с основаниями** $a\neq 1$ **является изоморфизмом группы** $\boldsymbol{R}^+$ **на группу** $\dot{\boldsymbol{R}}$.

▷ Пусть $a\neq 1$, $a>0$. По определению a^x — гомоморфизм из $\boldsymbol{R}^+$ в $\dot{\boldsymbol{R}}$, а по свойству 7 a^x — инъективное отображение. Поэтому достаточно проверить, что a^x сюръективно, т. е., что для любого $y_0\in\dot{\boldsymbol{R}}$ найдется $x_0\in\boldsymbol{R}$, такой, что $a^{x_0}=y_0$. Рассмотрим случай $a>1$. Тогда найдется натуральное число n, такое, что $n\cdot a>y_0$. Действительно, по свойству архимедовости вещественных чисел такое натуральное число n существует. Аналогичным образом выберем $m\in\boldsymbol{N}$ так, чтобы $a^m>\frac{1}{y_0}$. Тогда $a^{-m}<y_0$.

Наконец, рассмотрим непрерывную функцию a^x на отрезке $[-m, n]$. По построению m и n число y_0 лежит на отрезке $[a^{-m}, a^n]$, и по теореме о промежуточном значении непрерывной функции найдется число $x_0\in[-m, n]$, такое, что $a^{x_0}=y_0$. ■

Упражнение 3. *Докажите свойство* 8 *для* $0<a<1$.

Свойство 8 и упражнение 2 удобно объединить в отдельное свойство:

9) **Показательная функция с основанием** a **является изоморфизмом групп** $\boldsymbol{R}^+$ **и** $\dot{\boldsymbol{R}}$ **тогда и только тогда, когда** $a\neq 1$.

3. Теорема существования и единственности показательной функции. Доказательство существования и единственности показательной функции с основанием $a>0$ существенно сложнее доказательства аналогичной теоремы для линейной функции. Существование и единственность функции a^x мы получим как следствие теоремы.

Теорема 2. ***Топологические группы*** $\boldsymbol{R}^+$ ***и*** $\dot{\boldsymbol{R}}$ ***изоморфны.***

Если бы мы уже доказали существование хотя бы одной показательной функции, например функции e^x, то она и была бы изоморфизмом групп $\boldsymbol{R}^+$ и $\dot{\boldsymbol{R}}$. Поэтому доказательство теоремы 2 тривиально вытекает из факта существования хотя бы одной показательной функции с основанием $a\neq 1$. Однако пока мы не знаем, существует ли хотя бы одна показательная функция с основанием $a\neq 1$, и в этом заключена вся трудность доказательства теоремы 2.

Мы дадим два доказательства этой теоремы, которые будут основаны на существенно разных идеях.

Первое из них использует сформулированную ниже теорему 3. В ней утверждается, что существуют ровно две связные топологические группы, нейтральные элементы которых имеют окрестности, гомеоморфные какому-нибудь интервалу в $\boldsymbol{R}$. Очевидно, одна из таких групп — это группа $\boldsymbol{R}^+$ (которая по теореме 2 изоморфна и в этом смысле не различается с группой $\dot{\boldsymbol{R}}$).

Определим вторую из этих групп, которую будем обозначать буквой T, а операции в T — $\oplus$ и $\ominus$.

По определению обозначим через T множество, которое является фактор-множеством множества $\boldsymbol{R}$ по отношению эквивалентности

$$x\sim y\Leftrightarrow(x-y)\in\boldsymbol{Z}.$$

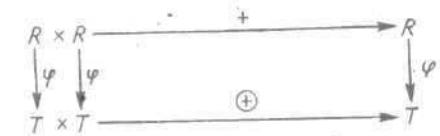


Рис. 4

Следовательно, элементами множества T являются произвольные подмножества $\boldsymbol{R}$ вида $\{..., x-1, x, x+1, x+2, ...\}$, где x — фиксированное число из $\boldsymbol{R}$.

Канонической функцией (каноническим отображением) называется функция φ вида $\boldsymbol{R}\to\boldsymbol{R}/\sim$, которая задается формулой $\varphi(x)=[x]$.

В множестве T определим операцию сложения $\oplus$ следующим образом: $[x]\oplus[y]\rightleftarrows[x+y]$.

Операция $\oplus$ в фактор-множестве T определяется с помощью операции $+$ в исходном множестве $\boldsymbol{R}$ (такой переход от операции $+$ к операции $\oplus$ называется иногда *опусканием* операции в фактор-множество). Конечно, такое опускание не всегда возможно: может случиться, что результат операции $\oplus$ изменится, если в классах эквивалентности $[x]$ и $[y]$ выбрать какие-то другие представители x' и y' вместо x и y. В данном случае это не так.

Действительно, если $x'\in[x]$ и $y'\in[y]$, то $x'+y'\in[x+y]$, так как $(x'+y')-(x+y)=(x'-x)+(y'-y)\in Z$. Поэтому $[x'+y']=[x+y]$. То же самое можно проиллюстрировать рисунком (см. рис. 4).

Для нахождения суммы двух элементов из множества T необходимо сначала поднять их в $\boldsymbol{R}$, затем сложить их прообразы в $\boldsymbol{R}$ и потом полученную сумму опустить в T. Аналогично определяется и операция вычисления симметричного элемента в T относительно операции $\oplus$, которая обозначается $\ominus$.

Определим во множестве T топологию, а именно:

(множество $\mathcal{O}$ открыто в T) $\rightleftarrows$ (множество $\varphi^{-1}(\mathcal{O})$ открыто в $\boldsymbol{R}$).

Иными словами, $\mathcal{T}_T\rightleftarrows\{\mathcal{O}\mid\varphi^{-1}(\mathcal{O})\in\mathcal{T}_R\}$. Такая топология называется *фактор-топологией*. Поскольку для отображения φ прообраз всякого открытого в T множества по определению открыт в $\boldsymbol{R}$, то каноническое отображение φ непрерывно.

Предложение 1. а) ***Каноническое отображение*** φ: $\boldsymbol{R}\to T$ ***непрерывно и открыто.***

б) ***Фактор-топология*** $\mathcal{T}_T$ ***содержит по включению всякую топологию в*** T, ***относительно которой*** φ ***непрерывно.***

в) ***Структура*** $\langle T, \ominus, \oplus, 0, \mathcal{T}_T\rangle$ ***является топологической группой, которая связна и компактна.***

▷ а) Первое утверждение доказано выше.

Напомним, что отображение f: $X\to Y$ называется открытым, если образ любого открытого в X множества открыт в Y, т. е.

$$\forall\,\mathcal{O}\in\mathcal{T}_X\ (f(\mathcal{O})\in\mathcal{T}_Y).$$

Упражнение 4. *Приведите примеры, показывающие, что*

из непрерывности функции $f(\cdot)$ не вытекает ее открытость и, наоборот, открытость функции необязательно влечет ее непрерывность.

Пусть $\mathcal{O}$ — открытое в $\boldsymbol{R}$ множество. Нужно доказать, что $f(\mathcal{O})$ открыто в T. Рассмотрим множество $\mathcal{O}+\boldsymbol{Z}=\bigcup\limits_{n\in Z}(\mathcal{O}+n)$. Множество $(\mathcal{O}+n)$, где n — фиксированное целое число, открыто в $\boldsymbol{R}$ (почему?). Поэтому и множество $\mathcal{O}_1 \rightleftarrows \mathcal{O}+\boldsymbol{Z}$ открыто в $\boldsymbol{R}$. Если $x\in\mathcal{O}_1$ и $y\sim x$, то $y\in\mathcal{O}_1$ (такое множество $\mathcal{O}_1$ называется насыщенным относительно отношения $\sim$). Выполняется $\varphi(\mathcal{O})=\varphi(\mathcal{O}_1)$: если $y=\varphi(x)$ и $x\in\mathcal{O}_1$ то $x=a+n$, где $a\in\mathcal{O}$, $\varphi(a)=[a]$, $\varphi(x)=[a+n]$, $\varphi(a)=\varphi(x)$. Выполняется $\varphi^{-1}(\varphi(\mathcal{O}_1))=\mathcal{O}_1$: если $\varphi(x)=\varphi(y)$, где $y\in\mathcal{O}_1$, то $[x]=[y]$, $x=y+n$, $x\in\mathcal{O}_1$. По определению топологии в T получаем, что $\varphi(\mathcal{O}_1)=\varphi(\mathcal{O})$ открыто в T.

б) Пусть $\mathcal{T}'_T$ — топология в T, относительно которой φ: $\boldsymbol{R}\to T$ непрерывна и $\mathcal{O}\in\mathcal{T}'_T$. Тогда $\varphi^{-1}(\mathcal{O})\in\mathcal{T}_R$ и по определению $\mathcal{T}_T\mathcal{O}\in\mathcal{T}_T$. Другими словами, $\mathcal{T}'_T\subseteq\mathcal{T}_T$.

Учитывая это утверждение, фактор-топологию $\mathcal{T}_T$ часто определяют как наибольшую среди всех топологий в T, относительно которых непрерывно каноническое отображение φ: $\boldsymbol{R}\to T$.

в) Проверка того, что алгебра $\langle T, \oplus, \ominus, 0\rangle$ является коммутативной группой, рекомендуется читателю в виде легкого упражнения. Докажем непрерывность групповых операций в T. Пусть $\mathcal{O}\in\mathcal{T}_T$. Для нахождения прообраза открытого в T множества $\mathcal{O}$ относительно операции $\oplus$ (см. рис. 4 на с. 27), можно сначала образовать множество $\varphi^{-1}(\mathcal{O})$, которое в силу непрерывности φ (или в силу определения топологии в T) принадлежит $\mathcal{T}_R$, и затем образовать относительно операции $+$ прообраз $\mathcal{O}'$ множества $\varphi^{-1}(\mathcal{O})$, который принадлежит $\mathcal{T}_{R^2}$, так как $+$ — непрерывное отображение вида $+$: $\boldsymbol{R}^2\to\boldsymbol{R}$. Затем для полученного прообраза $\mathcal{O}'$ следует найти образ при отображении $\varphi\times\varphi$, т. е. множество $(\varphi\times\varphi)(\mathcal{O}')$, где $(\varphi\times\varphi)(<x, y>)\rightleftarrows<\varphi(x), \varphi(y)>$. Это последнее множество открыто в $T\times T$.

Действительно, если $\mathcal{O}'$ открыто в $\boldsymbol{R}\times\boldsymbol{R}$, то по определению топологии произведения $\mathcal{O}'=\bigcup(U_\alpha\times V_\alpha)$, где множества U_α и V_α открыты в $\boldsymbol{R}$. Тогда $(\varphi\times\varphi)(\mathcal{O}')=\bigcup\limits_\alpha(\varphi\times\varphi)(U_\alpha\times V_\alpha)$, и нужно доказать, что все слагаемые в этом объединении открыты в $T\times T$, т. е. доказать, что $(\varphi\times\varphi)(U\times V)$ открыто в T, когда множества U и V открыты в $\boldsymbol{R}$. Выполняется $(\varphi\times\varphi)(U\times V)=\varphi(U)\times\varphi(V)$. Из этого предложения следует, что множества $\varphi(U)$ и $\varphi(V)$ открыты в T, и тогда по определению топологии произведения множество $\varphi(U)\times\varphi(V)$ открыто в $T\times T$.

Поэтому операция $\oplus$ непрерывна. Непрерывность операции $\ominus$ проверяется аналогично.

Множество T связно, так как является непрерывным образом связного множества $\boldsymbol{R}$.

Множество T компактно, так как является непрерывным образом компакта $[0, 1]$. ■

Операции $\oplus$ и $\ominus$ в T обычно записываются короче: $+$ и $-$.

Теорема 3. *Пусть G — связная топологическая группа, в которой нейтральный элемент e имеет открытую окрестность $\mathcal{O}_e$, гомеоморфную какому-нибудь интервалу в $\boldsymbol{R}$. Тогда G изоморфна $\boldsymbol{R}^+$ или T.*

Доказательство теоремы 3 содержится в приложении к этой главе, а сейчас мы используем ее для доказательства теоремы 2.

▷ Первое доказательство теоремы 2. Применим теорему 3 в случае, когда $G=\boldsymbol{R}^\cdot$. Нейтральным элементом в этой группе служит число 1, в качестве $\mathcal{O}_e$ можно взять любой интервал из положительных чисел, содержащий 1. Например, $\left]\frac{1}{2}, \frac{3}{2}\right[$ или $\left]\frac{1}{17}, \frac{16}{15}\right[$. Тогда гомеоморфизмом будет являться тождественное отображение $\mathcal{O}_e$ на себя. Так как $\boldsymbol{R}^\cdot_{>0}$ связно, то $\boldsymbol{R}^\cdot$ изоморфно $\boldsymbol{R}^+$ или T. Но $\boldsymbol{R}^\cdot$ не может быть изоморфно T, так как T компактно, а $\boldsymbol{R}^\cdot_{>0}$ нет. Следовательно, группы $\boldsymbol{R}^+$ и $\boldsymbol{R}^\cdot$ изоморфны и теорема 2 доказана. ■

▷ Второе доказательство теоремы 2. Сейчас мы не будем опираться на топологические свойства групп $\boldsymbol{R}^+$ и $\boldsymbol{R}^\cdot$. Вместо этого явно укажем вид одного определенного изоморфизма между этими группами. Определение этого изоморфизма совпадает с известным из анализа определением функции e^x с помощью ряда. Иными словами, мы определим функцию e^x, которая и будет служить этим изоморфизмом. Отметим, что такое доказательство имеет характер искусственного приема. Непонятно, например, какие свойства $\boldsymbol{R}^+$ и $\boldsymbol{R}^\cdot$ обеспечивают существование такого ряда. Кроме того, трудно представить себе, что такой ряд можно в самом деле угадать, не зная его заранее.

Итак, определим функцию ψ: $\boldsymbol{R}\to\boldsymbol{R}$ равенством $\psi(x)\rightleftarrows\sum\limits_{n=0}^{\infty}\frac{x^n}{n!}$.

По признаку Даламбера этот ряд сходится в каждой точке x из R. Радиус сходимости этого ряда равен бесконечности. Поскольку внутри интервала сходимости сумма степенного ряда непрерывна, то функция $\psi(x)$ является непрерывной на всем множестве $\boldsymbol{R}$. По теореме о произведении абсолютно сходящихся рядов получаем, что произведение $\psi(x)\cdot\psi(y)$ является суммой абсолютно сходящегося ряда с общим членом вида

$$\frac{x^n}{n!}\cdot 1+\frac{x^{n-1}}{(n-1)!}\cdot\frac{y}{1!}+\frac{x^{n-2}}{(n-2)!}\cdot\frac{y^2}{2!}+\ldots+\frac{x^{n-k}}{(n-k)!}\cdot\frac{y^k}{k!}+\ldots+$$
$$+\frac{y^n}{n!}\cdot 1=\frac{1}{n!}\sum_{k=0}^{n}C_n^k\cdot x^{n-k}\cdot y^k=\frac{(x+y)^n}{n!}.$$

Таким образом, функция ψ является непрерывным гомоморфизмом.

Докажем, что $\psi(R)=\boldsymbol{R}^\cdot_{>0}$.

Если $\psi(x)=0$ для какого-то $x\in\boldsymbol{R}$, то $\psi(1)=\psi(x+(1-x))=\psi(x)\cdot\psi(1-x)=0$, но $\psi(1)>1$. Поскольку $\psi(\boldsymbol{R})$ — связное мно-

жество и подмножество множества $\boldsymbol{R}$, не содержащее 0 и содержащее число $\psi(1)>1$, то $\psi(\boldsymbol{R})$ — связное подмножество $\boldsymbol{R}_{>0}$.

Поскольку $\psi(n)=(\psi(1))^n\to+\infty$, то $\psi(R)$ не ограничено сверху, и, следовательно, $\sup\psi(R)=+\infty$. Поскольку $\psi(-n)=(\psi(1))^{-n}\underset{n\to\infty}{\to}0$, то $\inf\psi(R)=0$. Поэтому $\psi(R)=R_{>0}$.

Инъективность функции ψ вытекает из ее монотонного возрастания, которое очевидно для положительных чисел: если $x>y\geqslant 0$, то $x^n>y^n$. Для отрицательных чисел монотонность функции ψ вытекает из равенства $\psi(-x)=(\psi(x))^{-1}$. Если $y<0<x$, то неравенство $\psi(y)=(\psi(-y))^{-1}<\psi(0)$ доказывает монотонность функции и в этом случае. ■

Теперь докажем теорему существования и единственности для аксиоматического объекта — функции a^x.

Теорема 4. ***Существует единственная показательная функция с основанием a, $a>0$, т. е. единственная функция f из R в $R_{>0}$ со свойствами:***

1) $f(x+y)=f(x)\cdot f(y)$ для всех $x, y\in R$;

2) f непрерывна;

3) $f(1)=a$.

▷ Пусть $\psi(\cdot)$: $\boldsymbol{R}^+\leftrightarrow\boldsymbol{R}$ — изоморфизм, существование которого утверждается в теореме 2. Если $\psi(1)=a$, то все доказано. Если $\psi(1)\neq a$, то положим $f(a)\rightleftarrows\psi(\psi^{-1}(1)\cdot x)$. Очевидно, $f(1)=a$ и $f(\cdot)$ — непрерывный гомоморфизм. ■

§ 3. ЛОГАРИФМИЧЕСКАЯ ФУНКЦИЯ

1. Аксиоматическое определение логарифмической функции.

Определение 3. **Логарифмической функцией называется любой непрерывный гомоморфизм группы R в группу R^+, т. е. любая функция f из множества $R_{>0}$ в множество R, обладающая свойствами:**

1) $f(x\cdot y)=f(x)+f(y)$;

2) f непрерывна.

Определение 3′. **Логарифмической функцией с основанием a, где $a>0$, называется логарифмическая функция, обладающая дополнительным свойством:**

3) $f(a)=1$.

Из свойства 1, т. е. из гомоморфности $f(\cdot)$, аналогично тому, как это делалось в § 1, 2, получим свойство:

4) $f(1)=0$.

При $a=1$ свойства 4 и 3 противоречат друг другу. Поэтому логарифмическая функция с основанием $a=1$ не существует. Часто ограничение $a\neq 1$ включают в само определение логарифмической функции по основанию a, что, как мы видим, находит свое объяснение с аксиоматической точки зрения.

Логарифмическую функцию с основанием a обозначим $\log_a x$.

Как и в предыдущем параграфе, возникают вопросы о существовании хотя бы одной такой функции $\log_a x$ и о ее единственности. Они будут решены позже. Перечислим основные свойства $\log_a x$ в предположении, что такая функция существует и единственна.

2. Свойства логарифмических функций. Теорема существования и единственности логарифмической функции. Начнем со свойств 1—3 и перепишем их, используя введенное обозначение $\log_a x$.

1) $\log_a(x\cdot y)=\log_a x+\log_a y$ для всех $x, y\in R_{>0}$;

2) функция $\log_a x$ непрерывна;

3) $\log_a a=1$.

Упражнение 5. *Докажите следующие свойства функции $\log_a x$:*

4) $\log_a(1)=0$;

5) $\log_a x^{-1}=-\log_a x$ при $x>0$;

6) $\log_a\left(\frac{x}{y}\right)=\log_a x-\log_a y$ при $x, y>0$;

7) $\log_a(a^x)=x$ для всех $x\in R$.

▷ Доказательство проводится по схеме, аналогичной той, которая использовалась в § 1 при доказательстве свойства 6. Иначе говоря, свойство 7 последовательно проверяется в случаях:

а) $x=n\in\boldsymbol{N}$; б) $x=-n$, $n\in\boldsymbol{N}$; в) $x=\frac{1}{n}$, $n\in\boldsymbol{N}$; г) $x=r\in\boldsymbol{Q}$;

д) x — любое число.

Рассмотрим, например, случай в: $n\cdot\log_a a^{1/n}=$

$$=\log_a a^{1/n}+\underset{n\text{ раз}}{\ldots}+\log_a a^{1/n}=\log_a\underbrace{a^{1/n}\cdot\ldots\cdot a^{1/n}}_{n\text{ раз}}=\log_a(a^{1/n})^n=$$

$=\log_a a=1$, что доказывает свойство 7 для таких чисел x. ■

Упражнение 6. *Докажите свойство 7 в случаях а, б, г, д, используя в случае д непрерывность $\log_a x$ и приближения любого вещественного числа x рациональными числами.*

Свойство 7 имеет важное следствие. Оно говорит о том, что композиция $(\log_a x)\circ a^x$ равна тождественной функции id: $\boldsymbol{R}\to\boldsymbol{R}$. Функция $\log_a(\cdot)$: $\boldsymbol{R}_{>0}\to\boldsymbol{R}$ сюръективна, так как для любого $x\in\boldsymbol{R}$ по свойству 7 $x=\log_a(a^x)$. Она инъективна: если y_1, $y_2\in\boldsymbol{R}_{>0}$ и $y_1\neq y_2$, то $y_1=a^{x_1}$, $y_2=a^{x_2}$ и $\log_a(a^{x_1})=x_1$, $\log_a(a^{x_2})=x_2$ и $x_1\neq x_2$, так как иначе $y_1=y_2$. Поэтому $\log_a x$ — биекция. Далее, $\log_a(a^{\log_a x})=\log_a x$ и $\log_a(x)=\log_a x$, и по инъективности функции $\log_a(\cdot)$ получаем $a^{\log_a x}=x$. Итак, мы доказали свойство:

8) $(a^x)\circ(\log_a x)=id$.

Следовательно, обе функции $\log_a x$ и a^x являются биекциями и эти биекции взаимно обратны. Таким образом, получено свойство:

9) логарифмическая функция по основанию a является биекцией, которая взаимно обратна с показательной функцией с основанием a.

10) Всякая логарифмическая функция по основанию a является изоморфизмом группы R на группу R^+.

▷ Функция $\log_a x$ обратна к функции a^x, $a \neq 1$. Функция a^x — изоморфизм, а любая функция, обратная к изоморфизму, является изоморфизмом. ■

Следствием свойства 9 и монотонности показательной функции (свойство 7, § 2) является монотонность функции $\log_a x$:

11) функция $\log_a x$ монотонно возрастает, если $a > 1$, и монотонно убывает, если $0 < a < 1$.

Наконец, приведем еще два традиционных свойства.

12) $\log_a (x^b) = b \log_a x$ при всех $b \in R$ и $x \in R > 0$.

▷ По свойству 8 $a^{\log_a x} = x$. Далее, $a^{b \cdot \log_a x} = (a^{\log_a x})^b = x^b$. Поскольку показательная функция с основанием $a \neq 1$ инъективна, то $\log_a (x^b) = b \cdot \log_a x$. ■

13) $\log_a x = \log_a c \cdot \log_c x$.

▷ По свойству 8 $a^{\log_a x} = x$.

По свойствам показательной функции с основанием a и свойству 8 получаем $a^{(\log_a c) \cdot \log_c x} = (a^{\log_a c})^{\log_c x} = c^{\log_c x} = x$.

Как и выше, доказательство заканчивается ссылкой на инъективность функции a^x при $a \neq 1$. ■

Теорема 5. ***Существует единственная логарифмическая функция с основанием $a > 0$, т. е. функция из $R_{>0}$ в R со свойствами:***

1) $f(x \cdot y) = f(x) + f(y)$;

2) f непрерывна;

3) $f(a) = 1$.

▷ Существование. Поскольку $a > 0$ и $a \neq 1$, то функция a^x будет изоморфизмом группы R^+ на группу R. Тогда изоморфизм, обратный к ней, обладает свойствами 1—3. Поэтому он является логарифмической функцией с основанием a.

Единственность. Если существуют две функции со свойствами 1—3, то каждая из них по свойству 8 будет обратной функцией к одной и той же функции, а именно к показательной функции a^x. Поэтому эти две функции совпадают. ■

§ 4. СТЕПЕННАЯ ФУНКЦИЯ

1. Аксиоматическое определение степенной функции.

Определение 4. **Степенной функцией называется любой непрерывный гомоморфизм группы R в себя, т. е. любая функция f, отображающая множество $R_{>0}$ в себя, обладающая свойствами:**

1) $f(xy) = f(x) f(y)$ для всех $x, y \in R_{>0}$;

2) f непрерывна.

Определение 4′. **Степенной функцией с показателем степени a, где число $a \in R$, называется степенная функция, обладающая дополнительным свойством:**

3) $f(2) = 2^a$.

Любую степенную функцию с показателем степени a будем обозначать через x^a.

Здесь мы несколько отступим от плана изложения, принятого в § 1—3. Сначала мы обсудим альтернативное определение степенной функции (замечание 2), затем докажем теорему существования и единственности аксиоматического объекта — степенной функции. Что же касается свойств степенной функции, то они будут перечислены в п. 3 и оставлены в качестве упражнения.

Замечания **1.** В свойстве 3 из определения 4′ число 2 не играет никакой специальной особой роли. С таким же успехом это свойство можно сформулировать в виде $f(137) = (137)^a$ или же $f\left(\frac{\pi}{2}\right) = \left(\frac{\pi}{2}\right)^a$. Единственное число, которое бессмысленно написать вместо числа 2, — это число 1. Дело в том, что для числа 1 свойство 3 выглядит так: $f(1) = 1^a = 1$, и точно такое же равенство вытекает из свойства 1.

2. Обозначение 2^a (или $\left(\frac{\pi}{2}\right)^a$) в свойстве 3 понимается в смысле значения показательной функции с основанием 2 (или $\frac{\pi}{2}$) в точке a. На этом пути можно по-новому определить функцию x^a как функцию, которая каждому положительному числу x ставит в соответствие значение показательной функции с основанием x и аргументом a.

Такое определение степенной функции содержит важный момент: оно иллюстрирует «двойственность» между функциями и их аргументами. Функции и их аргументы в определенном смысле взаимозаменяемы.

Более точно: фиксируем некоторое значение аргумента x и каждой функции f, определенной в точке x, из множества функций F поставим в соответствие значение этой функции в точке x, т. е. число $f(x)$. Тем самым фиксированное число x канонически определяет функцию, заданную на множестве функций F. Другими словами, аргумент x и функция f «поменялись местами». В этом смысле степенная функция x^a является функцией, канонически определяемой фиксированным числом a, на множестве F, состоящем из всех показательных функций.

3. Для некоторых значений a степенная функция x^a допускает продолжение на более широкую область определения, нежели $R_{>0}$. Например, если $a = n \in N$, то по четности или нечетности (в зависимости от n) x^a продолжается на $R_{<0}$. Кроме того, по непрерывности 0^n полагают равным 0. Если же $a = -n$, $n \in N$, то x^a продолжается только на $R_{<0}$. Если $a > 0$, то можно доказать, что $\lim\limits_{\alpha \to 0+} x^a = 0$. Чтобы не нарушалась непрерывность функции x^a, в этом случае полагают значение 0^a равным 0.

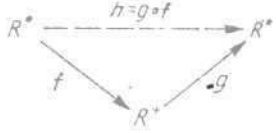


Рис. 5

При нечетном $n \in \boldsymbol{N}$ функция $x^{1/n}$ допускает естественное продолжение на всю числовую прямую, а при четном $n \in \boldsymbol{N}$ подобное продолжение невозможно. Такая ситуация связана с инъективностью или неинъективностью функции x^n, так как равенство $(x^{\frac{1}{n}})^n = x$ (§ 2, свойство 6) по существу задает $x^{\frac{1}{n}}$ как функцию, обратную к x^n. Поэтому, скажем, функцию $x^{\frac{20}{13}}$ можно считать определенной для всех $x \in \boldsymbol{R}$, а функцию $x^{\frac{13}{20}}$ — только для неотрицательных x.

Мы хотим изучать функцию x^a, не накладывая условий на значение показателя степени a. Поэтому приходится считать функцию x^a определенной на множестве $\boldsymbol{R}_{>0}$, в противном случае пришлось бы учитывать специфические особенности числа a.

2. Теорема существования и единственности степенной функции.

Теорема 6. *Существует единственная степенная функция с показателем степени a, где $a \in R$.*

▷ Существование. Рассмотрим треугольник из непрерывных гомоморфизмов (рис. 5), где $f(\cdot)$ переводит число 2 в число 1, а $g(\cdot)$ переводит число 1 в число 2^a. Такие $f(\cdot)$ и $g(\cdot)$ существуют по теоремам 5 и 4. Тогда $h(\cdot)$ как композиция непрерывных гомоморфизмов является непрерывным гомоморфизмом, переводящим число 2 в число 2^a.

В обозначениях § 2 получаем $x^a = (2^a)^{\log_2 x}$.

Единственность. Если функции h_1 и h_2 обладают свойствами 1—3 из определения 4′, а $f = \log_2 x$, то функции $g_1 \rightleftarrows h_1 \circ f^{-1}$ и $g_2 \rightleftarrows h_2 \circ f^{-1}$ являются двумя различными непрерывными гомоморфизмами из группы $\boldsymbol{R}^+$ в группу $\boldsymbol{R}$, которые оба переводят число 1 в число 2^a. Это противоречит единственности показательной функции с основанием 2^a (теорема 4). ■

3. Свойства степенных функций.

4) $1^a = 1$;

5) $\left(\frac{1}{x}\right)^a = \frac{1}{x^a}$;

6) $(x^a)^b = x^{a \cdot b}$;

7) $x^0 = 1$;

8) функция x^a монотонно возрастает при $a > 0$ и монотонно убывает при $a < 0$;

9) функция x^a является изоморфизмом группы R^+ на себя тогда и только тогда, когда $a \neq 0$.

Упражнение 7. *Докажите перечисленные выше свойства степенной функции.*

§ 5. ФУНКЦИИ КОСИНУС И СИНУС ЧИСЛОВОГО АРГУМЕНТА

Можно было бы начать этот параграф с определения, аналогичного определениям из § 1—4. Однако прежде чем сделать это, мы попытаемся объяснить, почему нам придется рассматривать непрерывные гомоморфизмы уже знакомой нам группы $\boldsymbol{R}^+$ в новую для нас группу S — комплексных чисел, по модулям равных единице.

При любом из многочисленных способов определения косинуса и синуса эти определения похожи. Аналогичны свойства этих функций, аналогичны методы их исследования и аналогично их использование в процессе решения задач.

Так как косинус и синус обычно сопутствуют друг другу, то естественно их рассматривать вместе, т. е. как функцию $f(x) \rightleftarrows \langle \cos x, \sin x \rangle$, которая отображает $\boldsymbol{R}$ в $\boldsymbol{R}^2$. Учитывая, что одним из ожидаемых свойств тригонометрических функций является тригонометрическое тождество $\cos^2 x + \sin^2 x = 1$, мы приходим к тому, что следует рассматривать непрерывную функцию из $\boldsymbol{R}$ в S.

Поскольку мы хотим, чтобы эта функция была похожа на линейную, показательную, логарифмическую и степенную функции, то нужно потребовать, чтобы она являлась не просто непрерывным отображением, а непрерывным гомоморфизмом, в данном случае из группы $\boldsymbol{R}^+$ в группу S. Любой такой непрерывный гомоморфизм назовем *экспонентой (экспоненциальной функцией)*.

Непрерывное отображение $\boldsymbol{R}$ в S для наглядности можно изобразить следующим механическим способом. Представим себе прямую $\boldsymbol{R}$ в виде бесконечной натянутой нити, которую произвольным непрерывным образом «наматывают» на обруч — окружность S. В случае произвольного непрерывного отображения $\boldsymbol{R}$ в S слово «наматывать», быть может, не слишком удачно: ведь нить при этом может не покрывать всю окружность или покрывать всю окружность, но при этом произвольным образом и в произвольных местах менять направление намотки и т. п.

Однако если, кроме непрерывности отображения $\boldsymbol{R}$ в S, потребовать его гомоморфность (относительно сложения в $\boldsymbol{R}$ и умножения в S), то это слово очень точно отражает то, что происходит на самом деле. Оказывается (теорема 7), что всякую экспоненту на интуитивном уровне можно описать так: все точки вида ..., $-a$, 0, a, $2a$, ..., где a — некоторое фиксированное положительное число, перейдут в точку $1 \in S$ и при этом отрезки ..., $[-a, 0]$, $[0, a]$, $[a, 2a]$, ... равномерно, периодически и в постоянном направлении накладываются на всю окружность S. Теорема 7 показывает, что периодичность возникает «сама собой», как следствие свойств непрерывности и гомоморфности.

Сначала мы займемся изучением произвольных непрерывных отображений множества $\boldsymbol{R}$ в множество S.

1. Экспоненциальная функция и ее периодичность.

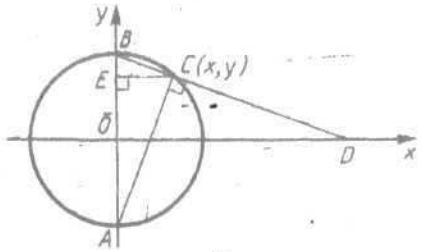


Рис. 6

Лемма 1. ***Окружность, из которой удалена какая-то одна фиксированная точка, гомеоморфна числовой прямой.***

▷ Доказательство заключается в аналитическом задании хорошо известного отображения — *стереографической проекции* $p(\cdot, \cdot)$, которая точке $C(x, y)$ сопоставляет точку D на оси Ox (рис. 6).

Из подобия треугольников BEC и OBD следует, что $|OD| = \frac{|x|}{1-y}$, так как $|BE| = 1-y$ и $|EC| = |x|$.

После этого нетрудно проверить, что отображение $p\colon S\setminus\{B\}\to\boldsymbol{R}$, задаваемое формулой $p(x, y) = \frac{x}{1-y}$, и является гомеоморфизмом. ■

В следующей лемме мы соберем необходимые в дальнейшем сведения о непрерывных функциях, определенных на отрезке.

Лемма 2. ***Пусть функция $g\colon J\to\boldsymbol{R}$ непрерывна на отрезке J. Тогда $g(J)$ — отрезок. Если $g(\cdot)$ инъективна, то $g(\cdot)$ монотонна.***

▷ Первая часть леммы является по существу объединением трех известных теорем о функциях, непрерывных на отрезке:

а) $g(J)$ ограничено, так как существует $m = \inf g(J)$ и $M = \sup g(J)$;

б) $m = g(x_0)$ и $M = g(x_1)$ для некоторых точек $x_0, x_1 \in J$;

в) если $m \leqslant c \leqslant M$, то при некотором $x \in J$ выполняется $g(x) = c$.

Докажем вторую часть леммы. Если g не монотонна на J, то найдутся такие точки $x_1 \leqslant x_2 \leqslant x_3$, что или $g(x_1) < g(x_2)$ и $g(x_3) < g(x_2)$, или $g(x_1) > g(x_2)$ и $g(x_3) > g(x_2)$.

Для определенности рассмотрим первый случай. Пусть $g(x_1) \geqslant g(x_3)$ (рис. 7). Тогда $g(x_4)$ — промежуточное значение функции g на отрезке $[x_2x_3]$, такое, что $g(x_1) = g(x_4)$ для некоторой точки $x_4 \in [x_2x_3]$, что противоречит инъективности функции g. ■

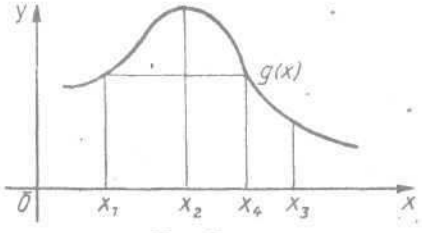


Рис. 7

Следующая лемма показывает, что прямую нельзя непрерывно и биективно отобразить на всю окружность.

Лемма 3. ***Пусть $f\colon\boldsymbol{R}\to S$ — непрерывное и сюръективное отображение. Тогда найдутся такие числа x_1 и x_2, что $f(x_1) = f(x_2)$.***

▷ Допустим, что f инъективно.

Уберем из множества S точку $f(0)$ и рассмотрим функцию вида $g = p \circ f$, где p — стереографическая проекция из точки $f(0)$. Тогда функция g инъективно и непрерывно отображает множество $\boldsymbol{R}\setminus\{0\}$ на множество $\boldsymbol{R}$. При этом прямая $\boldsymbol{R}$ разбивается на два непересекающихся множества:

$$A_+ \rightleftarrows \{x \mid \exists t > 0 (x = g(t))\} \text{ и } A_- \rightleftarrows \{x \mid \exists t < 0 (x = g(t))\}.$$

Поскольку $\boldsymbol{R}$ связно, то в одном из этих множеств найдется предельная точка другого.

Пусть $x_0 = g(t_0)$, где $t_0 > 0$, и $x_n \to x_0$, где $\{x_n\} \subseteq A_-$. Выберем такой отрезок J, состоящий из положительных чисел, что t_0 — внутренняя точка J. В силу монотонности g на J (лемма 2) $x_0 = g(t_0)$ — внутренняя точка $g(J)$. Поэтому все x_n, начиная с некоторого номера, лежат в A_+, что противоречит выбору x_n. ■

Теперь посмотрим, что дает объединение свойств непрерывности и гомоморфности в случае отображения из группы $\boldsymbol{R}^+$ в группу S.

Определение 5. **Экспоненциальной функцией (короче, экспонентой) называется любой непрерывный гомоморфизм из группы R^+ в группу S, т. е. любая функция из множества R в множество S, обладающая свойствами:**

1) $f(x+y) = f(x) \cdot f(y)$ для всех $x, y \in R$;

2) f непрерывна.

Отметим, что если в § 2—4 выражение $f(x) \cdot f(y)$ было произведением чисел из $\boldsymbol{R}_{>0}$, то теперь так обозначается и произведение комплексных чисел, равных по модулю единице. Следовательно, хотя название и обозначение операции умножения остались прежними, речь идет о двух принципиально разных операциях.

Упражнение 8. *Докажите, что: а) S — топологическая группа; б) если f — экспонента, то $f(0) = 1 \in S$, $f(-x) = (\overline{f(x)})$ и $f(x-y) = \frac{f(x)}{f(y)}$;*

в) ядро экспоненты f — замкнутая подгруппа в группе R^+;

г) если H — замкнутая подгруппа в группе $\boldsymbol{R}^+$, то возможны только три случая: $H = \boldsymbol{R}$, $H = \{0\}$ и $H = \{n \cdot a \mid n \in \boldsymbol{Z}\}$ при некотором фиксированном $a > 0$.

Приведите пример, показывающий, что утверждение г неверно, если H — незамкнутая подгруппа.

Лемма 4. ***Образ любой экспоненты — замкнутое подмножество в множестве S.***

▷ Пусть какое-то число $z_0 \in S$ и $f(x_n) = z_n \to z_0$ (см. рис. 8). Требуется доказать, что $z_0 = f(x_0)$ при некотором $x_0 \in \boldsymbol{R}$. Если $z_0 = 1$, то $f(0) = z_0$, и все доказано. Допустим, что $z_0 \neq 1$, т. е. $\mathrm{Re}(z_0) < 1$.

Как важное следствие равен-

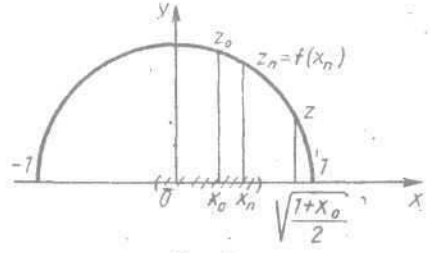


Рис. 8

ства $f(-x)=(f(x))^-$ (упражнение 8, б) получаем, что множество $f(\boldsymbol{R})$ симметрично относительности Ox. Поэтому можно считать, что число z_0 лежит в верхней полуокружности и, следовательно, имеет вид $z_0=x_0+i\sqrt{1-x_0^2}$, где $-1\leqslant x_0<1$. Эти ограничения на число x_0 влекут неравенства $x_0<\sqrt{\frac{1+x_0}{2}}<1$.

Рассмотрим комплексное число $z=\sqrt{\frac{1+x_0}{2}}+i\sqrt{\frac{1-x_0}{2}}\in S$. Легко видеть, что $z^2=z_0$. Докажем, что $z\in f(\boldsymbol{R})$, и тогда $(z=f(x))\Rightarrow(z^2=f(2\cdot x))$, что и требуется доказать.

Для этого рассмотрим непрерывную функцию вида $g=p_1\circ f:\boldsymbol{R}\to[-1,\ 1]$, где p_1 — проектирование на ось Ox, т. е. $p_1(x,\ y)\rightleftarrows x$. Отметим, что $g(0)=1$.

При достаточно больших n числа $x_n=g(t_n)=p_1(z_n)$ будут меньше числа $\sqrt{\frac{1+x_0}{2}}$, так как $x_n\to x_0$ и $\sqrt{\frac{1+x_0}{2}}>x_0$. Поэтому число $\sqrt{\frac{1+x_0}{2}}$ будет промежуточным значением функции g на любом отрезке с концами в точках O и t_n при достаточно больших n. Из непрерывности g следует, что $\sqrt{\frac{1+x_0}{2}}=g(t)$ при некотором t из отрезка $[O,\ t_n]$. Но это означает, что число $\sqrt{\frac{1+x_0}{2}}+i\sqrt{\frac{1-x_0}{2}}$ или число $\sqrt{\frac{1+x_0}{2}}-i\sqrt{\frac{1-x_0}{2}}$ равно $f(t)$. По симметричности множества $f(\boldsymbol{R})$ получаем, что оба эти числа и, в частности, число z лежат в $f(\boldsymbol{R})$. ■

Теорема 7. *Пусть $f:R^+\to S$ — произвольная экспонента, отличная от постоянной. Тогда существует такое число $a>0$, что $f(t\pm a)=f(t)$ при всех $t\in R$ и при этом функция f биективно отображает полуинтервал $[0,\ a[$ на всю окружность S.*

▷ Обозначим H ядро экспоненты f, т. е. множество тех $t\in\boldsymbol{R}^+$, для которых $f(t)=1$. Возможны три случая (упражнение 8): а) $H=\boldsymbol{R}$; б) $H=\{0\}$; в) $H=\{n\cdot a\,|\,n\in Z\}$ при некотором фиксированном $a>0$.

Случай а невозможен, так как f отлична от постоянной. Докажем, что и случай б невозможен. Допустим, $H=\{0\}$. Это означает, что функция f инъективна. Действительно, если $f(t_1)=f(t_2)$, то $f(t_1-t_2)=1$, т. е. $(t_1-t_2)\in H$ и $t_1=t_2$. Тогда по лемме 3 функция f не сюръективна. Выберем какое-то число $z_0\notin f(\boldsymbol{R})$ и обозначим через p стереографическую проекцию из точки z_0. Образуем непрерывную инъективную функцию вида $g=p\circ f$, определенную на множестве $\boldsymbol{R}$ (рис. 9). По лемме 2 эта функция монотонна. При

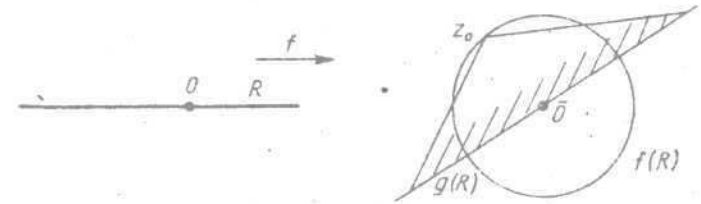


Рис. 9

этом множество $g(\boldsymbol{R})$ замкнуто (лемма 4), связно (как непрерывный образ связного множества) и ограничено (так как z_0 не принадлежит $f(\boldsymbol{R})$ вместе с некоторой ее окрестностью). Другими словами, множество $g(\boldsymbol{R})$ — отрезок, что противоречит непрерывности и монотонности функции g.

Итак, возможен только случай в, когда $H=\{n\cdot a\,|\,n\in\boldsymbol{Z}\}$, где число $a>0$. Так как $f(a)=1$, то $f(t+a)=f(t)\cdot f(a)=f(t)$ и $f(t-a)=f(t)\cdot(f(a))^{-1}=f(t)$, т. е. число a — период функции f.

Проверим, что a является главным периодом этой функции: если какое-то число b, где $0<b<a$, также является периодом функции f, то $b\in H$. Однако $b\neq n\cdot a$ ни при каком значении n из $\boldsymbol{Z}$. Получим противоречие. Итак, a — главный период экспоненты f.

Осталось доказать биективность функции f на $[0,\ a[$. Отображение f инъективно на $[0,\ a[$: если $f(t_1)=f(t_2)$ при $0\leqslant t_1<t_2<a$, то $f(t_2-t_1)=1$, т. е. $0<(t_2-t_1)<a$ и $(t_2-t_1)\in H$ — противоречие.

Докажем, что функция f сюръективно отображает отрезок $\left[0,\ \frac{a}{4}\right]$ на одну из четвертей окружности S. Действительно, $1=f(a)=f\left(4\cdot\frac{a}{4}\right)=\left(f\left(\frac{a}{4}\right)\right)^4$, т. е. число $f\left(\frac{a}{4}\right)$ — корень уравнения $z^4-1=0$. У этого уравнения имеются всего четыре корня, это числа ±1 и $\pm i$. Однако $f\left(\frac{a}{4}\right)\neq1$, так как $0<\frac{a}{4}<a$, и $f\left(\frac{a}{4}\right)\neq-1$, иначе $f\left(\frac{a}{2}\right)=\left(f\left(\frac{a}{4}\right)\right)^2=1$, что противоречит неравенствам $0<\frac{a}{2}<a$. Остались две возможности: $f\left(\frac{a}{4}\right)=i$ или $f\left(\frac{a}{4}\right)=-i$. Рассмотрим случай, когда $f\left(\frac{a}{4}\right)=i$. Образуем непрерывную функцию вида $p_1(f(t))$, определенную на отрезке $\left[0,\ \frac{a}{4}\right]$. Если при некотором значении t из этого отрезка $p_1(f(t))<0$, то $t<\frac{a}{4}$ и найдется число $t_0\in(0,\ t)$, для которого $p_1(f(t_0))=0$ (рис. 10).

Последнее означает, что $f(t_0)=\pm i$. Неравенство $f(t_0)=i$ противоречит инъективности f на отрезке $\left[0,\ \frac{a}{4}\right]\subset[0,\ a[$, а из равенства

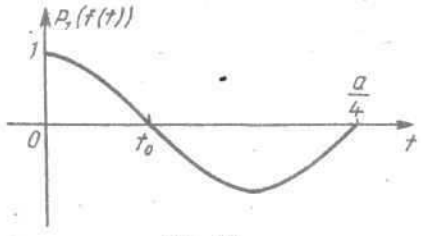


Рис. 10

$f(t_0)=-i$ следует, что $f(-t_0)=i$ и $f(a-t_0)=i$, где $\frac{3a}{4}<a-t_0<a$, что также противоречит инъективности f на $[0, a[$. Следовательно, $p_1(f(t))\geqslant 0$ на $\left[0, \frac{a}{4}\right]$. Аналогично доказывается и неотрицательность функции $p_2(f(t))$ на отрезке $\left[0, \frac{a}{4}\right]$. Поэтому значение $f(t)$ лежит в первой четверти окружности. По непрерывности значения функции $p_1(f(t))$ при $t\in\left[0, \frac{a}{4}\right]$ пробегают отрезок $[0, 1]$, и, следовательно, образ $f\left(\left[0, \frac{a}{4}\right]\right)$ совпадает с первой четвертью окружности S.

Если теперь $z_0=x_0+iy_0\in S$ и $x_0\leqslant 0$, $y_0\geqslant 0$, то $z_0=i(y_0-ix_0)=$ $=if(t)=f\left(t+\frac{a}{4}\right)$, где $0\leqslant t\leqslant\frac{a}{4}$. Следовательно, функция f сюръективно отображает отрезок $\left[\frac{a}{4}, \frac{a}{2}\right]$ на вторую четверть S.

Аналогично рассматриваются отрезки $\left[\frac{a}{2}, \frac{3a}{4}\right]$ и $\left[\frac{3a}{4}, a\right]$ и случай, когда $f\left(\frac{a}{4}\right)=-i$. ■

2. Теоремы существования и единственности экспоненциальной функции.

Определение 5′. **Экспонентой с основанием a, где число $a>0$, называется любая экспонента, у которой a — главный период и «отображение происходит в положительном направлении». Иными словами, экспонентой с основанием a называется любая функция $f(\cdot)$ из множества R в множество S со свойствами:**

1) $f(x+y)=f(x)\cdot f(y)$ для всех $x, y\in R$;
2) f непрерывна;
3) a — главный период f;
4) $f\left(\frac{a}{4}\right)=i$.

Экспонента с основанием $a>0$ обозначается $e_a(x)$. Подчеркнем, что теорема 7 позволяет из свойств 1 и 2 вывести существование такого числа $a>0$, которое является главным периодом экспоненты. Поэтому свойство 3 заключается только в фиксировании некоторого числа $a>0$ в качестве главного периода экспоненты.

Что означает фраза «отображение в положительном направлении»? Можно сказать, что она по определению означает наличие свойства 4.

Однако мы попробуем объяснить это менее формально.

Обозначим $T_z(S)$ касательную прямую к окружности S в точке $z\in S$. Отложим на прямой $T_z(S)$ какой-то вектор $p(z)$ с началом в точке z. Поступим таким образом для каждой точки z из S (рис. 11).

Получится отображение вида $p(\cdot)$: $S\to\boldsymbol{R}^2$. Его можно интерпретировать как некоторое множество $R(p)$ касательных векторов $p(z)$, параметризованных точками z окружности S. Такое отображение $p(\cdot)$ называется *касательным векторным полем на S*. Если функция $p(\cdot)$ вида p: $\boldsymbol{R}^2\to\boldsymbol{R}^2$ непрерывна, то и *поле* называют *непрерывным*.

Касательное векторное поле называют *невырожденным*, если для любой точки z выполняется $p(z)\neq 0$. Далее до конца этого пункта предполагается, что все рассматриваемые касательные векторные поля непрерывны и невырожденны.

Для любой точки $z\in S$, очевидно, существуют такие неравные точки z_1 и z_2, что один отрезок $[z_1, z]$ целиком принадлежит единичному кругу D, а второй отрезок $[z, z_2]$ целиком ему не принадлежит. Произвольно выберем такие две точки. Внешним вектором в точке z назовем любой вектор вида $\lambda\cdot\overrightarrow{z, z_2}$, где $\lambda\in\boldsymbol{R}_{>0}$; обозначим его $q(z)$. Если точка z_2 лежит на нормали к прямой $T_z(S)$, то внешний вектор $\lambda\cdot\overrightarrow{z, z_2}$ называется *внешней нормалью* и обозначается $n(z)$. *Правой* (положительной) *ориентацией* S называется касательное векторное поле $p(\cdot)$, для которого базис $\langle q(z), p(z)\rangle$ является правым для всех точек z. Последнее по определению означает, что определитель матрицы вида $(q(z)p(z))$ положителен для всех точек z (он не может равняться нулю, так как по условию $p(z)$ невырожденно).

Аналогично определяется *левая* (отрицательная) *ориентация* S. Такое определение ориентации достаточно для практических подсчетов, однако оно имеет принципиальный недостаток: существует много разных правых и левых ориентаций. Если произвольно выбрана какая-то функция $\lambda(\cdot)$ вида $\lambda(\cdot)$: $S\to\boldsymbol{R}_{>0}$ и $p(\cdot)$: $S\to\boldsymbol{R}^2$ — правая ориентация, то касательное векторное поле $\lambda(\cdot)\cdot p(\cdot)$ также правая ориентация. И кроме того, оно использует понятие внешнего вектора $q(z)$, которое также неоднозначно и основано на том, что S — определенная часть $\boldsymbol{R}^2$.

Более глубокая форма того же определения выглядит следующим образом.

Касательные векторные поля $p_1(\cdot)$ и $p_2(\cdot)$ называются одинаково направленными (эквивалентными), если для каждой точки $z\in S$ векторы $p_1(z)$ и $p_2(z)$ одинаково направлены. Иными словами, если существует такое отображение $\lambda(\cdot)$: $S\to\boldsymbol{R}_{>0}$, то $p_1(z)=\lambda(z)\cdot p_2(z)$ для всех точек $z\in S$.

Оказывается, если $p_1(\cdot)$ и $p_2(\cdot)$ — непрерывные и невырожденные касательные векторные поля, то возможны только два случая:

а) $p_1(\cdot)$ и $p_2(\cdot)$ одинаково направлены (эквивалентны);

б) $p_1(\cdot)$ и $p_2(\cdot)$ противоположно направлены, т. е. $p_1(\cdot)$ и $-p_2(\cdot)$ одинаково направлены (эквивалентны).

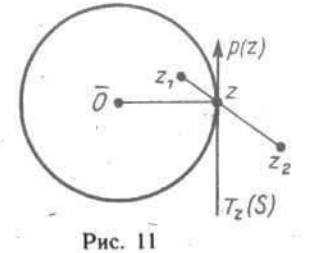


Рис. 11

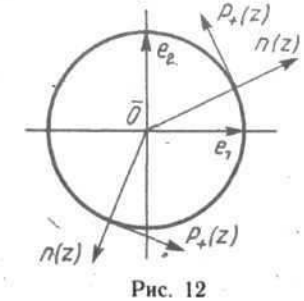


Рис. 12

Итак, *ориентацией окружности* S называют любой из двух классов эквивалентности в множестве всех непрерывных невырожденных касательных векторных полей относительно указанного выше отношения эквивалентности. *Правой* (положительной) *ориентацией* S называют тот из этих двух классов эквивалентности, в который входит правое векторное поле $p_+(\cdot)$, т. е. такое, что базис $\langle n(z),\ p_+(z)\rangle$ правый для всех точек $z\in S$. Правая ориентация как класс эквивалентности однозначно определяется любым своим представителем, т. е. правой ориентацией как отдельным касательным векторным полем, у которого все базисы $\langle q(z),\ p(z)\rangle$ правые. И наоборот, такое отдельное векторное поле однозначно определяет класс эквивалентности в множестве векторных полей, т. е. правую ориентацию. Поэтому два приведенных определения правой (и левой) ориентации по существу не отличаются друг от друга.

В приложении к этой главе с. 326—340 доказывается, что свойство 4 в определении 5′ эквивалентно тому, что отображение из множества $\boldsymbol{R}$ в множество S, осуществляемое экспонентой f, таково, что касательное векторное поле p вида $p(f(x))\rightleftarrows f'(x)$ является правой ориентацией окружности S. Здесь $f'(x)\rightleftarrows(\varphi'(x),\ \psi'(x))$, где $f(x)=\varphi(x)+i\psi(x)$, или, иными словами, $\psi(x)\rightleftarrows \mathrm{Re} f(x)$ и $\psi(x)\rightleftarrows \mathrm{Im} f(x)$. Вектор $f'(x)$ по определению откладывается от точки $f(x)=(\varphi(x),\ \psi(x))$. Следовательно, он лежит на прямой $T_{f(x)}(S)$. Поэтому отображение $p(\cdot)$ действительно является касательным векторным полем на S. Это свойство касательного векторного поля $p(\cdot)$ и имеют в виду, когда говорят: «Отображение $f(\cdot)$ происходит в положительном направлении» (рис. 12).

Теоремы существования и единственности для экспоненциальной функции сформулируем и докажем в два этапа, первый из которых — единственность такой функции.

Теорема 8. *Если $f(\cdot)$ и $g(\cdot)$ — любые экспоненты с основанием $a>0$, то они совпадают.*

▷ Как и во второй части доказательства теоремы 7, достаточно проверить совпадение функций f и g на отрезке $\left[0,\ \frac{a}{4}\right]$. Допустим, это не так. Рассмотрим новую функцию $h:\boldsymbol{R}^+\to S$, где $h(x)\rightleftarrows\frac{f(x)}{g(x)}$. Легко показать, что h — экспонента. Поскольку h не постоянная экспонента, то по теореме 7 функция $h(\cdot)$ имеет какой-то главный период $b>0$. При этом $\frac{a}{4}=n\cdot b$ при некотором $n\in\boldsymbol{N}$, так как $f\left(\frac{a}{4}\right)=g\left(\frac{a}{4}\right)=i$ и $h\left(\frac{a}{4}\right)=1$, и, следовательно, $\frac{a}{4}$ — период функции h. Опять же по теореме 7 получаем $h[0,\ b]=h[0,\ \frac{a}{4}]=S$. Число $h(x)$ при $x\in\left[0,\ \frac{a}{4}\right]$ является отношением комплексных чисел из первой четверти окружности (по доказательству теоремы 7). Поэтому число $h(x)$ не может лежать ни во второй, ни в третьей четверти, что противоречит равенству $h\left[0,\ \frac{a}{4}\right]=S$. ■

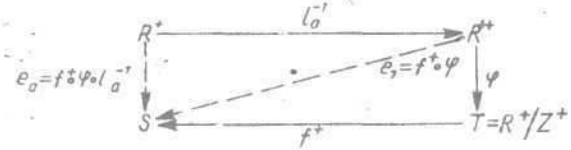


Рис. 13

Упражнение 9. *Докажите, что:*

а) если f и g — экспоненты, то и функция $h\rightleftarrows\frac{f}{g}$ — экспонента;

б) если $\frac{a+bi}{c+di}=\alpha+\beta i$; где $a,\ b,\ c,\ d>0$, то $\alpha\geqslant 0$.

Теорема 9. *Для любого числа $a>0$ существует экспонента с основанием a.*

Внимательно рассмотрим следующую схему, составленную из непрерывных гомоморфизмов (рис. 13).

Здесь l_a — линейная функция с коэффициентом a, т. е. $l_a(x)=a\cdot x$ и $l_a^{-1}(x)=\frac{1}{a}\cdot x$, φ — каноническое отображение множества $\boldsymbol{R}$ в фактор-множество T, а функция $f^+(\cdot)$ определяется следующим образом.

Поскольку S — связная топологическая группа, нейтральный элемент которой имеет окрестность, гомеоморфную открытому интервалу (докажите!), то по теореме 3 группы S и T изоморфны.

Пусть $f:T\to S$ — один из таких изоморфизмов. Рассмотрим относительно f образ элемента $\left[\frac{1}{4}\right]\in T$, т. е. образ класса эквивалентности, содержащего представитель $\frac{1}{4}$:

$$1=f([0])=f([1])=f\left(4\left[\frac{1}{4}\right]\right)=\left(f\left(\left[\frac{1}{4}\right]\right)\right)^4.$$

Поэтому число $f\left(\left[\frac{1}{4}\right]\right)$ — корень уравнения $z^4-1=0$, и, следовательно, $f\left(\left[\frac{1}{4}\right]\right)\in\{\pm 1;\ \pm i\}$. Однако случай $f\left(\left[\frac{1}{4}\right]\right)=1$ невозможен, так как $f([0])=1$ и функция $f(\cdot)$ инъективна. Случай $f\left(\left[\frac{1}{4}\right]\right)=-1$ также невозможен, так как тогда $f\left(\left[\frac{1}{2}\right]\right)=\left(f\left(\left[\frac{1}{4}\right]\right)\right)^2$, что снова противоречит инъективности функции $f(\cdot)$. Следовательно, возможны только два случая: $f\left(\left[\frac{1}{4}\right]\right)=i$ и $f\left(\left[\frac{1}{4}\right]\right)=-i$.

Если $f\left(\left[\frac{1}{4}\right]\right)=i$, то положим $f^+(\cdot)\rightleftarrows f(\cdot)$. Если $f\left(\left[\frac{1}{4}\right]\right)=-i$,

то положим $f^+(\cdot) \rightleftarrows -f(\cdot)$. Итак, функции, соответствующие на рисунке 13 сплошным стрелкам, определены. Функции, соответствующие на этом рисунке пунктирным стрелкам определяются по сплошным стрелкам! Положим $e_a \rightleftarrows f^+ \circ \varphi \circ l_a^{-1}$. Функция e_a — непрерывный гомоморфизм, так как она является композицией непрерывных гомоморфизмов. Найдем H — ядро функции $e_a(\cdot)$. При изоморфизме $f^+(\cdot)$ в число 1 перейдет элемент $[0] \in T$, в который при отображении φ перейдут все целые числа, т. е. все z. В целые числа при отображении $l_a^{-1}(\cdot)$ перейдет множество $\{n \cdot a \mid n \in \mathbf{Z}\}$. Другими словами, $H = \{n \cdot a \mid n \in \mathbf{Z}\}$, и поэтому верно свойство 3. Наконец, число $\frac{a}{4}$ относительно функции $l_a^{-1}(\cdot)$ перейдет в число $\frac{1}{4}$ и число $\frac{1}{4}$ относительно $\varphi(\cdot)$ перейдет в класс эквивалентности $\left[\frac{1}{4}\right]$, а класс $\left[\frac{1}{4}\right]$ по построению функции $f^+(\cdot)$ перейдет в число i. Поэтому $e_a\left(\frac{a}{4}\right) = i$, что доказывает свойство 4 для функции $e_a(\cdot)$. ■

Замечания. 1. Можно дать «механическое» описание $e_a(\cdot)$: при отображении $l_a^{-1}(\cdot)$ полуинтервалы ..., $[-a, 0[$; $[0, a[$; $[a, 2a[$, ... сжимаются соответственно в полуинтервалы ...; $[-1, 0[$; $[0, 1[$; $[1, 2[$; ... . При отображении $\varphi(\cdot)$ эти полуинтервалы отождествляются с одним и тем же полуинтервалом[1] $[0, 1[$, который затем при изоморфизме $f^+(\cdot)$ «наматывается» биективно в положительном направлении на окружность S.

2. Наиболее употребительны экспоненты с основаниями 1 и 2π. Первая из них обозначается $e_1(x)$, а вторая $e_{2\pi}(x)$.

Выделение экспоненты $e_1(\cdot)$ происходит потому, что при $a = 1$ функции l_a и l_a^{-1} совпадают с тождественной функцией на $\boldsymbol{R}$, и, следовательно, $e_1 = f^+ \circ \varphi$; поскольку $e_a \rightleftarrows f^+ \circ \varphi \circ l_a^{-1}$, то получаем $e_a(x) = e_1\left(\frac{x}{a}\right)$.

Выделение экспоненты $e_{2\pi}(\cdot)$ происходит по более сложным, отчасти историческим причинам — это экспонента с наиболее привычным, «тригонометрическим» главным периодом 2π.

3. Как и в случае показательной функции (см. второе доказательство теоремы в § 2), можно доказать существование функции $e_a(\cdot)$, используя вместо теоремы 3 ряды определенного вида:

$$e_a(x) \rightleftarrows \sum_{n=0}^{\infty} \frac{(2\pi i)^n}{n!}\left(\frac{x}{a}\right)^n. \qquad (*)$$

Это определение выглядит более привычно для функции $e_1(\cdot)$:

$$e_1(x) \rightleftarrows \sum_{n=0}^{\infty} \frac{(2\pi i x)^n}{n!}.$$

[1] Здесь полуинтервал $[0, 1[$ отождествляется с T по канонической биекции $\varphi: [0, 1[\leftrightarrow T$.

Можно использовать ряд только для определения функции $e_1(\cdot)$ и тогда функцию $e_a(\cdot)$ определить как $e_1\left(\frac{x}{a}\right)$.

Итак, функция $e_a(x)$ определена нами двумя способами: во-первых, аксиоматически в определении 5′ и, во-вторых, явной конструкцией (конструктивно) с помощью вышеуказанного ряда. Возникает вопрос: эквивалентны ли эти определения, т. е. задают ли они одну и ту же функцию? Ответ положителен.

Чтобы в этом убедиться, нужно только проверить, что сумма ряда $\sum_{n=0}^{\infty} \frac{(2\pi i)^n}{n!}\left(\frac{x}{a}\right)^n$ удовлетворяет всем четырем аксиомам из определения 5′. Это делается прямыми вычислениями, аналогичными тем, которые составляют второе доказательство теоремы 2. Этим четырем аксиомам удовлетворяет только одна функция (по теореме 4), которая и была нами обозначена $e_a(\cdot)$. Поэтому $e_a(\cdot)$ совпадает с суммой этого ряда. Тем самым равенство $(*)$ можно рассматривать не как определение его левой части с помощью правой части, а как утверждение о совпадении его левой части (аксиоматически введенной в определении 5′) с его правой частью (понимаемой как сумма конкретного ряда).

Сумму ряда $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ по традиции обозначают e^x. Таким образом, мы приходим к важным равенствам:

$$e_a(x) = e^{\frac{2\pi}{a} \cdot ix}, \quad e_1(x) = e^{2\pi i x}, \quad e_{2\pi}(x) = e^{ix}.$$

3. Функции косинус и синус числового аргумента: аксиоматические определения и свойства.

Определение 6. **Косинусом и синусом называются функции, определенные следующими равенствами:**

$$\cos x \rightleftarrows P_1(e_{2\pi}(x)) = P_1(e^{ix}),$$
$$\sin x \rightleftarrows P_2(e_{2\pi}(x)) = P_2(e^{ix}),$$

где P_1 и P_2 — обычные функции из C в R, определенные равенствами $P_1(x+iy) = x$, $P_2(x+iy) = y$.

Это определение использует одну конкретную экспоненту $e_{2\pi}(x)$, выбранную среди всех возможных экспонент $e_a(x)$. Иными словами, оно использует одно конкретное число 2π, выбранное среди всех чисел $a > 0$. В качестве основания экспоненты число 2π ничем не лучше любого другого числа $a > 0$. По крайней мере в элементарной математике при решении задач и в теоретических вопросах с тем же успехом вместо $\cos x$ и $\sin x$ можно рассматривать следующие функции:

$$\cos_a x \rightleftarrows P_1(e_a(x)) = P_1(e^{\frac{2\pi}{a} ix}),$$

$$\sin_a x \rightleftarrows P_2(e_a(x)) = P_2(e^{\frac{2\pi}{a} ix}).$$

Однако, следуя исторической традиции, обычно рассматривают именно функции $\cos_{2\pi} x = \cos x$ и $\sin_{2\pi} x = \sin x$.

Как определить число 2π? Первый способ определения заключается в том, что сначала дается определение длины окружности S, затем доказывается, что эта длина существует, и число 2π определяется как длина окружности S. Если начинать разговор о длине окружности, то неестественно отрывать его от общего понятия длины кривой на плоскости и, например, от понятия длины дуги окружности, которое в свою очередь требует определения понятия дуги. Этот материал вместе с определением тригонометрических функций и меры угла на основе понятия длины дуги (фактически на основе понятия интеграла) мы поместим в приложении 1 к этой главе в конце книги.

Другой способ определения числа 2π аксиоматический: числом 2π называется такое число $a > 0$, для которого выполняется соотношение

$$\lim_{x \to 0} \frac{e_a(x) - 1}{x} = i. \qquad (*)$$

Конечно, при таком аксиоматическом определении числа 2π возникает вопрос о его существовании и единственности. Если воспользоваться равенством $e_a(x) = e^{i\frac{2\pi}{a} x}$, то по правилу Лопиталя предел в равенстве $(*)$ равен

$$\frac{(e^{i\frac{2\pi}{a} x} - 1)'}{(x)'}\Big|_0 = \frac{\left(\frac{2\pi}{a} i e^{\frac{\pi}{2a} ix}\right)}{1}\Big|_0 = \frac{\frac{2\pi}{a} i}{1}.$$

Приравняв это выражение i, получим:

$$\frac{2\pi}{a} i = i, \quad a = 2\pi.$$

Однако это рассуждение опирается на замечание 2 к теореме 9, которое уже использует число 2π. Строгое доказательство существования и единственности числа, определяемого равенством $(*)$, довольно сложно.

Свойства тригонометрических функций $\cos x$ и $\sin x$ непосредственно следуют из определения экспоненты $e_{2\pi}(\cdot)$ с основанием 2π и из свойств операции произведения комплексных чисел. Например, основное тригонометрическое тождество $\cos^2 x + \sin^2 x = 1$ является просто следствием того, что значения функции $e(\cdot)$ лежат на единичной окружности S. Из гомоморфности экспоненты

получаем $e_{2\pi}\left(x + \frac{\pi}{2}\right) = e_{2\pi}(x) \cdot e_{2\pi}\left(\frac{\pi}{2}\right) = (\cos x + i \sin x) i = -\sin x + i \cos x = \cos\left(x + \frac{\pi}{2}\right) + i \sin\left(x + \frac{\pi}{2}\right)$, откуда вытекают формулы приведения для аргументов x и $x + \frac{\pi}{2}$. Гомоморфность экспоненты $e_{2\pi}(\cdot)$ в точности эквивалентна формулам сложения, т. е. свойству б из упражнения 6.

Упражнение 10. *Докажите, основываясь на аксиоматическом определении функций* $\cos x$ *и* $\sin x$, *следующие их свойства:*

а) $\cos(-x) = \cos x$, $\sin(-x) = -\sin x$;

б) $\cos(x_1 + x_2) = \cos x_1 \cdot \cos x_2 - \sin x_1 \cdot \sin x_2$;

$$\sin(x_1 + x_2) = \sin x_1 \cdot \cos x_2 + \cos x_1 \cdot \sin x_2;$$

в) функции $\cos x$ *и* $\sin x$ *непрерывны и имеют главные периоды* 2π.

Вычислите $\cos x$ *и* $\sin x$ *при* $x = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}, 2\pi$.

Вычислите $\cos x$ *и* $\sin x$ *при* $x = \pm\frac{\pi}{6}, \pm\frac{\pi}{3}, \pm\frac{\pi}{4}$.

Найдите для функций $\cos x$ *и* $\sin x$ *промежутки, на которых они имеют постоянный знак.*

Докажите формулу $\cos x_1 - \cos x_2 = -2 \sin\frac{x_1 + x_2}{2} \cdot \sin\frac{x_2 - x_1}{2}$ *и используйте ее при доказательстве строгого убывания функции* $\cos x$ *на отрезке* $[0, \pi]$. *Объясните возможность определения функции* $\arccos x$ *как функции, обратной к* $\cos x$ *на отрезке* $[0, \pi]$, *и зависимость этого определения от выбора отрезка, на котором рассматривается* $\cos x$.

Установите участки монотонности для $\sin x$ *и определите* $\arcsin x$ *как функцию, обратную к* $\sin x$.

Докажите формулы приведения для аргументов x *и* $x + \pi$, *и* $x \pm \frac{3\pi}{2}$.

Докажите, что уравнение $z^n = z_0 = \cos x + i \sin x$ *имеет ровно* n *корней, лежащих на* S, *которые получаются при подстановке чисел* $k = 0, 1, \ldots, n-1$ *в равенство*

$$z = \cos\frac{x + 2\pi k}{n} + i \sin\frac{x + 2\pi k}{n}.$$

Докажите, что $\cos\left(\frac{x}{a}\right) = P_1(e_{2\pi a}(x))$ *и* $\sin\left(\frac{x}{a}\right) = P_2(e_{2\pi a}(x))$. *Это дает интерпретацию числа* a, *так как период функций* $\cos\left(\frac{x}{a}\right)$ *и* $\sin\left(\frac{x}{a}\right)$ *равен* $2\pi a$.

§ 6. УГОЛ. ФУНКЦИИ КОСИНУС И СИНУС УГЛОВОГО АРГУМЕНТА. ИЗМЕРЕНИЕ УГЛОВ

1. Введение. В предыдущем параграфе были определены и изучены тригонометрические функции числового аргумента, причем это было сделано, с одной стороны, достаточно геометрично, а с другой стороны, без всякого использования углов на плоскости.

Однако на практике тригонометрические функции углов рассматриваются столь же часто, как и тригонометрические функции чисел.

Для построения тригонометрических функций углов необходимо:

а) определить, что такое угол;

б) уметь переходить от углов к комплексным числам — элементам S.

Второе необходимо потому, что косинус и синус угла α по традиции определяются как первая и вторая координаты той точки z из S, которая однозначно соответствует исходному углу α.

Кроме этих вопросов, в данном параграфе рассматривается соответствие между углами и вещественными числами, которые измеряют эти углы, и на этом пути формируется еще одно определение тригонометрических функций числа.

Прежде чем перейти к реализации этой программы, нужно подчеркнуть, что понятие угла предполагает уже известным понятие плоскости. Это видно хотя бы из выражения «угол α в плоскости $\boldsymbol{P}$», которое употребляют иногда вместо слов «угол α». Поэтому возникает вопрос, в какой плоскости мы собираемся определять и рассматривать углы. Интуитивное представление о плоскости формируется у ребенка очень рано, едва ли не раньше, чем интуитивное представление о числе. Однако (и это далеко не общеизвестно!) в математике изучается несколько неэквивалентных (в строго математическом смысле) определений плоскости, каждое из которых только в большей или меньшей степени соответствует интуитивному представлению о плоскости.

Перечислим основные подходы к введению понятия плоскости (которые будут подробнее рассматриваться в § 2 главы II):

1) арифметическая плоскость, т. е. декартово произведение множества вещественных чисел $\boldsymbol{R}$ на себя, такая плоскость обозначается $\boldsymbol{R}^2$;

2) векторная плоскость, т. е. произвольное двумерное векторное пространство над полем $\boldsymbol{R}$;

3) аффинная плоскость;

4) проективная плоскость;

5) аксиоматически определенная плоскость (по Евклиду — Гильберту или в другой аксиоматике).

В рамках каждого из этих подходов можно определить понятие угла в данной плоскости. В этом параграфе мы рассмотрим простейший вид углов — углы в арифметической плоскости $\boldsymbol{R}^2$. Иными словами, в этом параграфе под словом «плоскость» всюду понимается арифметическая плоскость, т. е. множество $\boldsymbol{R}^2$.

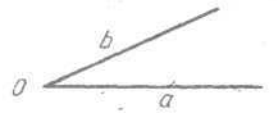


Рис. 14

Напомним, что лучом (с началом в точке $\mathbf{O}\rightleftarrows(0, 0)$ называется любое множество точек вида $\{\lambda\cdot\binom{x}{y} \mid \lambda\in\mathrm{R}_{\geqslant 0}\}$, где (x, y) — какая-то фиксированная точка в $\boldsymbol{R}^2$, не равная точке $\mathbf{O}$. Точка в $\boldsymbol{R}^2$ с числовыми координатами x и y записывается в равной мере $\binom{x}{y}$ или (x, y).

2. Определение угла в арифметической плоскости.

Определение 7. Обозначим X множество всех лучей в R^2 с началом в точке O. В множестве X^2 определим отношение эквивалентности: $\langle a, b\rangle\sim\langle c, d\rangle\rightleftarrows\exists z\in S\,(c=z\cdot a\wedge d=z\cdot b)$. Углом называется любой класс эквивалентности в множестве X^2 по этому отношению эквивалентности. Множество всех углов, т. е. фактор-множество X^2, обозначим An (от латинского *angle* — угол). Угол, содержащий пару лучей $\langle a, b\rangle$, обозначим $[\langle a, b\rangle]$.

Сравним это определение угла с более распространенным элементарным определением, когда углом называют (упорядоченную) пару лучей, выходящих из одной точки (рис. 14).

В определении 7 (первом определении угла) в качестве точки $\mathbf{O}$ для всех углов берется одна и та же точка $\mathbf{O}$. Однако это различие не существенно: в случае второго определения угол $a\mathbf{O}b$ рассматривают «с точностью до параллельного переноса», т. е. $\angle a\mathbf{O}b$ и $\angle a'\mathbf{O}'b'$ считают равными, если существует такое ξ из $\boldsymbol{R}^2$, что $\mathbf{O}=\mathbf{O}'+\xi$, $a=a'+\xi$, $b=b'+\xi$. Поэтому всегда можно перейти к «равному» углу с вершиной в точке $\mathbf{O}$. Следовательно, вершину угла можно не указывать и $\angle ab$ в смысле второго определения однозначно задает угол $[\langle ab\rangle]$ в смысле первого определения.

Почему угол в первом смысле — это класс пар лучей, а угол во втором смысле — это одна пара таких лучей? Дело в том, что при втором определении всегда добавляют: углы $\angle ab$ и $\angle cd$ по определению «равны», если существует поворот A (геометрическое преобразование A плоскости $\boldsymbol{R}^2$ определенного вида), для которого $A(c)=a$ и $A(d)=b$. Геометрические преобразования, которые называют поворотами, как раз совпадают с умножением на фиксированное комплексное число z из S (см. ниже изложение мелким шрифтом). Поэтому такое «равенство» углов на самом деле задает отношение эквивалентности в множестве пар лучей, совпадающее с тем, которое указано в определении 7. Итак, угол в смысле второго определения, рассматриваемый с точностью до этого «равенства» (отношения эквивалентности), соответствует классу всех пар лучей, эквивалентных данной паре лучей, т. е. соответствует углу в смысле первого определения. Это показывает, что определение 7 является точной формулировкой обычного понятия угла.

Прежде чем изучать определенное нами понятие угла, рассмотрим понятие поворота в арифметической плоскости и покажем, что повороты в $\boldsymbol{R}^2$ — это биекции (преобразования) вида f: $\boldsymbol{R}^2 \to \boldsymbol{R}^2$, которые определяются формулой $f(u) \rightleftarrows z \cdot u$, где $u \in \boldsymbol{R}^2$ и z — фиксированное число из S.

Напомним, что скалярным произведением в $\boldsymbol{R}^2$ называется отображение вида $(\cdot, \cdot)$: $\boldsymbol{R}^2 \times \boldsymbol{R}^2 \to \boldsymbol{R}$, определяемое равенством $\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) \rightleftarrows x_1 \cdot x_2 + y_1 \cdot y_2$. Длиной вектора (точки) из $\boldsymbol{R}^2$ с координатами x и y называется число $\sqrt{(\xi, \xi)} = \sqrt{x^2+y^2}$. Множество всех матриц размера 2×2 из вещественных чисел обозначим M_2. Каждая матрица $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ определяет функцию $\xi \mapsto A \cdot \xi$, где $A \cdot \xi \rightleftarrows \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix}$. Множество M_2 относительно обычных операций сложения и умножения матриц является кольцом, но не телом и не полем. Обозначим GL_2 множество всех матриц с определителем, не равным нулю, т. е. $GL_2 \rightleftarrows \{A \in M_2 \mid \det A \neq 0\}$. Это множество относительно обычной операции умножения матриц является группой, в которой симметричным элементом для A является элемент A^{-1} и нейтральным элементом является матрица $E \rightleftarrows \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Матрица A называется ортогональной, если выполняется условие $A^t = A^{-1}$, где A^t — транспонированная матрица.

Упражнение 11. *Докажите, что:*

а) условие $A^t = A^{-1}$ эквивалентно сохранению скалярного произведения, т. е. условию $\forall \xi, \eta \in \boldsymbol{R}^2 ((\xi, \eta) = (A\xi, A\eta))$, и эквивалентно условию: столбцы матриц A образуют ортонормированную систему векторов. Последнее означает:

$$\left(\begin{pmatrix} a_{1i} \\ a_{2i} \end{pmatrix} \cdot \begin{pmatrix} a_{1j} \\ a_{2j} \end{pmatrix}\right) = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j; \end{cases}$$

б) определитель всякой ортогональной матрицы равен числам $+1$ или -1.

Обозначим O_2 множество всех ортогональных матриц. Относительно обычной операции умножения матриц множество O_2 является группой.

Нам будут нужны ортогональные матрицы с положительным определителем. Множество таких матриц обозначим SO_2, т. е. $SO_2 \rightleftarrows \{A \in O_2 \mid \det A > 0\} = \{A \in O_2 \mid \det A = +1\}$.

Лемма 5. *Матрица $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in SO_2$ **тогда и только тогда, когда для некоторого числа $t \in \boldsymbol{R}$ выполняются условия $a_{11} = a_{22} = \cos t$ и $a_{12} = -a_{21} = -\sin t$.***

▷ Обозначим $e_1 \rightleftarrows \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ и $e_2 \rightleftarrows \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Так как матрица A ортогональна, то $a_{11}^2 + a_{21}^2 = (Ae_1, Ae_1) = (e_1, e_1) = 1$. Поэтому точка (a_{11}, a_{21}) лежит на единичной окружности S. Так как экспонента $e_{2\pi}(x) = e^{ix}$ является отображением $\boldsymbol{R}^+$ на S (теорема 7), то для некоторого числа $t \in \boldsymbol{R}$ выполняется $a_{11} = \cos t$ и $a_{21} = \sin t$.

Из равенств $(Ae_2, Ae_2) = a_{12}^2 + a_{22}^2 = 1$ и $(Ae_1, Ae_2) = a_{11}a_{12} + a_{21}a_{22} = 0$ получаем

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Поэтому произведение A на транспортированную матрицу A^t равно единичной матрице. При умножении перемножаются и их определители $\det(A^t) = \det A$. По условию $\det A > 0$, отсюда

$$\det A = a_{11} \cdot a_{22} - a_{21} \cdot a_{12} = 1.$$

Поэтому для неизвестных a_{12} и a_{22} имеем два линейных уравнения:

$$a_{22} \cos t - a_{12} \sin t = 1,$$
$$a_{22} \sin t + a_{12} \cos t = 0.$$

Решая эту систему, получаем, что $a_{22} = \cos t = a_{11}$, $a_{12} = -\sin t = -a_{21}$. Тот факт, что любая матрица вида $\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ принадлежит SO_2, проверяется прямым вычислением. ■

Следствие. Матрица A принадлежит SO_2 тогда и только тогда, когда ее столбцы образуют ортонормированную систему векторов и ее определитель равен 1.

Упражнение 12. *а) Проверьте, что GL_2 и O_2 — группы и O_2 — подгруппа GL_2.*

б) Докажите, что SO_2 — коммутативная группа и отображение $f: SO_2 \to S$, определяемое равенством $f\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \rightleftarrows \cos t + i \sin t$, является изоморфизмом этих групп.

Топология в S, конечно, индуцируется топологией в $\boldsymbol{R}^2$, т. е.

$$\mathcal{O}_1 \in \mathcal{T}_S \rightleftarrows \exists \mathcal{O}_2 \in \mathcal{T}_{R^2} (\mathcal{O}_1 = \mathcal{O}_2 \cap S).$$

Топологию в SO_2 определим с помощью биекции из упражнения 12, б, а именно: $\mathcal{O} \in \mathcal{T}_{SO_2} \rightleftarrows f(\mathcal{O}) \in \mathcal{T}_S$. При таком определении топологии в SO_2 биекция $f(\cdot)$, очевидно, является гомеоморфизмом. И следовательно, топологические группы SO_2 и S изоморфны.

Может возникнуть вопрос, нет ли другого определения топологии в SO_2, которое не было бы связано с окружностью S. Действительно, множество M_2 канонически биективно с $\boldsymbol{R}^4$ с помощью биекции $g(\cdot)$: $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \leftrightarrow \langle a_{11}, a_{12}, a_{21}, a_{22} \rangle$. Поэтому топологию в M_2 можно индуцировать и из $\boldsymbol{R}^4$, а именно:

$$\mathcal{O} \in \mathcal{T}_{M_2} \rightleftarrows g(\mathcal{O}) \in \mathcal{T}_{R^4}.$$

Поскольку в M_2 имеется топология, то ее можно индуцировать и в SO_2 как часть M_2, а именно: $\mathcal{T}_{SO_2} \rightleftarrows \{\mathcal{O} \cap SO_2 \mid \mathcal{O} \in \mathcal{T}_{M_2}\}$. Итак, в SO_2 определено две топологии: одна — индуцированная топологией в S, другая — индуцированная топологией в $\boldsymbol{R}^4$.

Упражнение 13. *Докажите, что эти две топологии в SO_2 совпадают.*

Теперь определим операцию сложения углов.

Определение 8. **Пусть даны два угла $[\langle a, b \rangle]$ и $[\langle c, d \rangle]$. Ясно, что луч c (как и любой луч) пересекает окружность S ровно в одной какой-то точке n** (рис. 15) **и луч b пересекает S в какой-то точке z. Число $\frac{z}{n}$ принадлежит S, и поэтому пара лучей $\left\langle \frac{z}{n} \cdot c, \frac{z}{n} \cdot d \right\rangle$ эквивалентна паре $\langle c, d \rangle$, т. е. $\left[\left\langle \frac{z}{n} \cdot c, \frac{z}{n} \cdot d \right\rangle\right] = [\langle c, d \rangle]$. При этом**

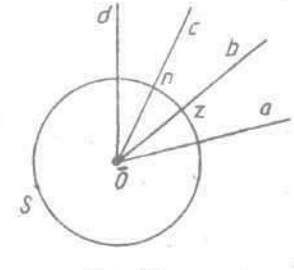


Рис. 15

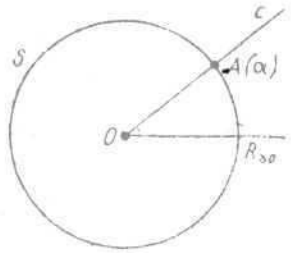


Рис. 16

$\frac{z}{n}\cdot c=b$. Положим $[\langle a,\ b\rangle]+[\langle c,\ d\rangle]\rightleftarrows$
$\rightleftarrows\left(\left\langle a,\ \frac{z}{n}d\right\rangle\right)$.

Геометрический смысл этого определения очевиден: для сложения углов нужно первую сторону второго угла повернуть до совпадения со второй стороной первого угла и затем образовать объединение полученных углов.

Упражнение 14. *Докажите, что:*

а) это определение не изменится, если вместо преобразования луча c в луч b будет преобразован луч b в луч c;

б) множество всех углов An относительно операции $+$ является коммутативной группой, в которой нейтральным элементом является угол $[\langle a,\ a\rangle]$ и симметричным элементом для угла $[\langle a,\ b\rangle]$ — угол $[\langle b,\ a\rangle]$.

в) Сформулируйте эквивалентное определение операции сложения с помощью поворотов, т. е. используя матрицы из SO_2 вместо комплексных чисел из S.

Определим теперь отображение $A(\cdot)$ вида A: $An\to S$. Для этого отметим, что в каждом угле $\alpha=[\langle a,\ b\rangle]$ содержится единственная пара лучей вида $\langle \boldsymbol{R}_{\geqslant 0},\ c\rangle$. Назовем ее каноническим представителем угла α. Вторая сторона c канонического представителя $\langle \boldsymbol{R}_{\geqslant 0},\ c\rangle$ угла пересекает S в единственной точке, которая по определению и является значением отображения $A(\cdot)$ на аргументе α (см. рис. 16).

Ясно, что отображение $A(\cdot):An\to S$ — биекция.

В множестве An определим топологию, индуцированную топологией в S. А именно: $\mathcal{O}\in\mathcal{T}_{An}\rightleftarrows(A(\mathcal{O})\in\mathcal{T}_S)$. Иными словами, множество $\mathcal{O}\subseteq An$ называется открытым, если подмножество $A(\mathcal{O})$ множества S открыто в S.

Таким образом, группы SO_2 и S можно не различать. Элементы группы SO_2 называются *поворотом плоскости*. Итак, поворот A как функция вида $\xi\mapsto A\cdot\xi$ совпадает с умножением на комплексное число z из S, соответствующее матрице A, т. е. с функцией $\xi\mapsto z\cdot\xi$.

Упражнение 15. *Докажите, что:*

а) алгебра $\langle An,\ +,\ -,\ O_{An},\ \mathcal{T}_{An}\rangle$ является топологической группой;

б) биекция $A(\cdot)$ является гомеоморфизмом.

Предложение 2.

а) Функция $A:An\to S$ является изоморфизмом групп An и S.

б) Функция $f:C^*\to An$, определяемая равенством

$$f(z)\rightleftarrows[\langle \boldsymbol{R}_{\geqslant 0};\ \lambda\cdot z\rangle],$$

является непрерывным гомоморфизмом.

Здесь запись $\lambda\cdot z$ обозначает луч, выходящий из точки $\bar{0}$ и проходящий через точку z.

▷ а) Биективность функции $A(\cdot)$ отмечалась выше. Гомеоморфность этой функции в сущности очевидна, так как топология в An определялась выше как раз так, чтобы эта функция оказалась гомеоморфизмом; впрочем, ее гомеоморфность отмечалась в упражнении 15, б.

Остается проверить гомоморфность функции A, т. е. равенство

$$A(\alpha+\beta)=A(\alpha)\cdot A(\beta).$$

По определению операции $\alpha+\beta$ получаем:

$$\alpha+\beta=[\langle \boldsymbol{R}_{\geqslant 0},\ \lambda\cdot A(\alpha)\rangle]+[\langle \boldsymbol{R}_{\geqslant 0},\ \lambda\cdot A(\beta)\rangle]=[\langle \boldsymbol{R}_{\geqslant 0},\ \lambda(A(\alpha)\cdot A(\beta))\rangle].$$

Отсюда по определению функции $A(\cdot)$ имеем $A(\alpha+\beta)=A(\alpha)\cdot A(\beta)$.

б) По доказанному в п. а функция $A^{-1}(\cdot)$ существует и непрерывна. Она имеет вид $A^{-1}(\cdot):S\to An$ (и как обратная к изоморфизму сама является изоморфизмом). Функция f является композицией функции A^{-1} и функции $z\mapsto\frac{1}{|z|}$. Вторая из этих функций, очевидно, есть непрерывный гомоморфизм. Композиция двух непрерывных гомоморфизмов является непрерывным гомоморфизмом. ■

3. Конструктивные определения функций косинус и синус углового аргумента. Свойства этих функций.

Определение 9. Косинусом и синусом угла α называются функции, определяемые соответственно равенствами

$$\cos\alpha\rightleftarrows P_1(A(\alpha)),\ \sin\alpha\rightleftarrows P_2(A(\alpha)),$$

где $\alpha\in An$.

Свойства этих функций углового аргумента легко вытекают из их определения и из свойств операции умножения комплексных чисел:

1) $\cos 0=1$, $\sin 0=0$ просто потому, что при гомоморфизме $A(\cdot)$, как и при любом гомоморфизме, нейтральный элемент 0 в группе An переходит в неитральный элемент 1 в группе S, координаты которого равны числам 1 и 0.

2) $\cos d=0$, $\sin d=1$, где d — прямой угол, т. е. угол, определяемый равенством $d\rightleftarrows A^{-1}(i)$.

3) $\cos 2d=-1$, $\sin 2d=0$, где $2d\rightleftarrows d+d$ и $+$ — определенная выше операция сложения в группе An. Действительно, $2d=A^{-1}(i)+A^{-1}(i)=A^{-1}(i\cdot i)=A^{-1}(-1)$, $A(2d)=-1$. Угол $2d$ называется развернутым углом.

4) $\cos(-\alpha)=\cos\alpha$, $\sin(-\alpha)=-\sin\alpha$ (рис. 17).

▷ По гомоморфности функции $A(\cdot)$ получаем $1=A(\alpha)\cdot A(-\alpha)$, $A(-\alpha)=(A(\alpha))^{-1}$ и п. 4 следует из определения сопряженного числа. ■

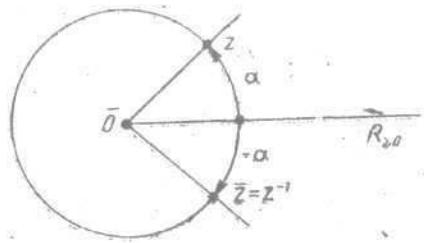


Рис. 17

5) $\cos(\alpha_1+\alpha_2)=P_1(A^{-1}(\alpha_1+\alpha_2))=P_1(A^{-1}(\alpha_1)\cdot A^{-1}(\alpha_2))=$ $P_1((\cos\alpha_1+i\sin\alpha_1)(\cos\alpha_2+i\sin\alpha_2))=P_1(\cos\alpha_1\cos\alpha_2-\sin\alpha_1\sin\alpha_2,$ $\sin\alpha_1\cos\alpha_2+\cos\alpha_1\sin\alpha_2)=\cos\alpha_1\cos\alpha_2-\sin\alpha_1\sin\alpha_2.$

Одновременно мы доказали другую хорошо известную формулу:

$$\sin(\alpha_1+\alpha_2)=\sin\alpha_1\cos\alpha_2+\cos\alpha_1\sin\alpha_2.$$

6) $\cos^2\alpha+\sin^2\alpha=1,\ \forall\alpha\in An.$

7) $\frac{(z,\ u)}{\|z\|\cdot\|u\|}=\cos|\langle\lambda z,\ \lambda u\rangle|$, где $z,\ u$ — любые точки из R^2, отличные от точки 0. Как всегда, $\|\xi\|$ обозначает длину вектора ξ. Отсюда следует известное неравенство Коши — Буняковского:

$$(z,\ u)^2\leqslant\|z\|^2\cdot\|u\|^2.$$

▷ Докажем этот пункт, используя для тренировки повороты, а не комплексные числа.

Рассмотрим точку $\frac{z}{\|z\|}\in S$ и матрицу $W\in SO_2$, которая соответствует комплексному числу $\frac{\|z\|}{z}$. Для этой матрицы выполняется соотношение $W\left(\frac{z}{\|z\|}\right)=1\in S$. Так как W сохраняет скалярное произведение, то

$$\left(\frac{z}{\|z\|},\ \frac{u}{\|u\|}\right)=\left(1,\ W\left(\frac{u}{\|u\|}\right)\right).$$

Последнее скалярное произведение по определению (учитывая $1=(1,\ 0)$) равно

$$P_1\left(W\left(\frac{u}{\|u\|}\right)\right)=\cos\left[\left\langle R_{\geqslant 0},\ \lambda\cdot W\left(\frac{u}{\|u\|}\right)\right\rangle\right]=$$
$$\cos\left[\left\langle\lambda\frac{z}{\|z\|},\ \lambda\frac{u}{\|u\|}\right\rangle\right].$$

В последнем равенстве используется определение угла как класса эквивалентности и равенство $W\left(\lambda\cdot\frac{z}{\|z\|}\right)=R_{\geqslant 0}$.

Осталось заметить, что луч $\lambda\cdot z$ совпадает с лучом $\frac{\lambda}{\|z\|}\cdot z$ и аналогично луч $\lambda\cdot u$ совпадает с лучом $\frac{\lambda}{\|u\|}\cdot u$. ■

8) Теорема Пифагора: **для любых точек A, B, C, если угол ABC прямой, то $|\overline{CA}|^2+|\overline{CB}|^2=|\overline{AB}|^2$** (см. рис. 18).

▷ Для упрощения записи обозначим $z=\overline{CA}$ и $u=\overline{CB}$. Тогда $\overline{AB}=u-z$. По свойствам скалярного произведения получаем: $\|u-z\|^2=(u-z,\ u-z)=(u,\ u)-2(u,\ z)+(z,\ z)=\|u\|^2+\|z\|^2-$ $-2\|u\|\cdot\|z\|\cos[\langle\lambda\cdot z,\ \lambda\cdot u\rangle]=\|z\|^2+\|u\|^2$. ■

4. Измерение углов. Что такое градусная и радианная мера угла? Что такое 1° и 1 радиан? Чем градусная мера отличается от радианной и что у них общего? Грамотный школьник скажет, что градусной мерой угла является некоторое число. Отсюда кажется, что градусная мера угла является какой-то функцией вида $An\to R$. Однако это не совсем так.

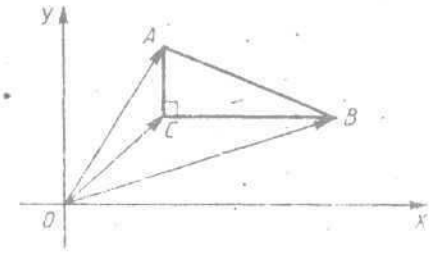


Рис. 18

С прямым углом в качестве градусной меры можно связать не только число 90, но и числа 270, 450, 810 и т. д. Получается, что одному углу соответствует несколько чисел в качестве его градусной меры. Правда, в школе иногда выходят из положения, говоря, что эти числа относятся к разным углам. Так возникают углы «против часовой» и «по часовой стрелке», «в два оборота с лишним против часовой стрелки» и т. п.

В то же время мы хотели бы остаться на той простой и ясной точке зрения, что угол — это пара лучей, рассматриваемая с точностью до поворотов. Поэтому приходится признать, то все числа 270, 450, 810 и т. д. являются градусной мерой прямого угла d. Другими словами, одному и тому же углу (элементу An) соответствует бесконечно много различных чисел (элементов $\boldsymbol{R}$) в качестве градусной меры этого угла. Тем самым измерение углов является «многозначной функцией» из множества An в множество R. И это в принципе правильная точка зрения. «Многозначности» можно избежать, если поступать следующим образом: сопоставлять не углам числа, а, наоборот, числам углы. Возвращаясь к конкретному примеру, любому из чисел 270, 90, 450 и вообще всем числам вида $90+360\cdot n$, где $n\in\boldsymbol{Z}$, поставим в соответствие (уже однозначно!) прямой угол d. Так как $\boldsymbol{R}^+$ и An являются группами, то в соответствии с идеями всей этой главы следует рассматривать не произвольные функции из множества $\boldsymbol{R}$ в множество An, а только такие функции, которые непрерывны и сохраняют групповые операции, т. е. непрерывные гомоморфизмы вида $\boldsymbol{R}^+\to An$.

Переход к непрерывным гомоморфизмам можно объяснить и более конкретным образом: при любом способе измерения углов естественно ожидать, что число нуль будет мерой нулевого угла, что симметричным углам α и $-\alpha$ будут соответствовать симметричные числовые меры x и $-x$, что при небольшом изменении числовой меры соответствующие углы мало изменятся. Эти интуитивно желательные свойства меры углов выполняются тогда, когда рассматриваются непрерывные гомоморфизмы из группы $\boldsymbol{R}^+$ в группу An.

Определение 10. **Мерой углов называется любой непрерыв-**

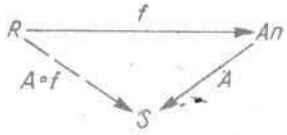


Рис. 19

ный гомоморфизм вида $f(\cdot): R^+ \to An$, т. е. любая функция f из множества R в множество An, обладающая свойствами:

1) $f(x+y) = f(x) + f(y)$ для всех $x, y \in R$;

2) $f(\cdot)$ непрерывна.

Подчеркнем, что в свойстве 1 слева знак $+$ является групповой операцией в $\boldsymbol{R}$, а справа знак $+$ является групповой операцией в An. В свойстве 2 непрерывность f понимается относительно топологии $\mathcal{T}_{An}$ в группе An и топологии $\mathcal{T}_R$ в группе $\boldsymbol{R}$.

Поскольку группы An и S изоморфны (предложение 2), то непрерывные гомоморфизмы из $\boldsymbol{R}^+$ в An тесно связаны с непрерывными гомоморфизмами из $\boldsymbol{R}^+$ в S, т. е. с экспонентами.

Теорема 10. *Всякая мера углов $f(\cdot): R^+ \to An$ имеет вид $f(x) = 0$, или $f(x) = A^{-1} \circ e_a(x)$, или $f(x) = A^{-1} \circ \bar{e}_a(x)$, где $e_a(x)$ — экспонента с каким-то основанием $a > 0$. Здесь $\bar{z}$ — комплексное число, сопряженное комплексному числу z.*

▷ Рассмотрим треугольник, состоящий из непрерывных гомоморфизмов (рис. 19).

Функция $A \circ f$ является некоторой экспонентой. По теореме 7 она либо тождественно равна числу 1 из S (тогда $f(\cdot)$ тождественно равно 0 из An), либо имеет некоторый главный период $a > 0$. Если при этом $A \circ f\left(\frac{a}{4}\right) = i$, то $A \circ f$ удовлетворяет всем условиям определения 5′, и по соответствующей теореме существования и единственности получаем $A \circ f = e_a(\cdot)$ и $f(\cdot) = A^{-1} \circ e_a(\cdot)$. Если же $A \circ f\left(\frac{a}{4}\right) = -i$, то функция $(A \circ f)^- : x \mapsto (A \circ f(x))^-$ удовлетворяет всем условиям определения 5′.

Поэтому $(A \circ f)^- = e_a(\cdot)$ и $f = A^{-1} \circ \bar{e}_a$.

Случай, когда $A \circ f\left(\frac{a}{4}\right) = -i$ или, что то же самое, когда $f(\cdot) = A^{-1} \circ \bar{e}_a(\cdot)$, соответствует прямой «намотке» $\boldsymbol{R}$ на окружность S под действием функции $\bar{e}_a$, т. е. намотке в отрицательном направлении. Действительно, точки $e_a(x)$ и $\bar{e}_a(x)$ симметричны относительно оси Ox, и если под действием функции e_a прямая $\boldsymbol{R}$ наматывается на окружность S в положительном направлении, то под действием функции $\bar{e}_a$ намотка будет происходить в обратном, отрицательном направлении. ■

Замечание. Так как $e_a(x) = e^{\frac{2\pi}{a} ix}$ и $\bar{e}_a(x) = e^{-\frac{2\pi}{a} ix}$, то случай меры углов вида $A^{-1} \circ \bar{e}_a$ соответствует параметру $(-a)$; таким образом, если только отображение $f(\cdot)$ не тождественный нуль, то ему однозначно соответствует некоторое число a из $\boldsymbol{R}^*$. Если тождественно нулевому отображению f сопоставить параметр 0, то каждой мере углов $f: \boldsymbol{R}^+ \to An$ соответствует ровно один числовой параметр.

Условимся впредь не рассматривать тривиальную меру углов $f(x) = 0$.

Число a, однозначно соответствующее по теореме 10 и предыдущему замечанию любой мере углов $f(\cdot)$, будем называть основанием этой меры $f(\cdot)$ и саму меру обозначать $f_a(\cdot)$. Основанием меры может служить любое число из $\boldsymbol{R}^*$. Как и в случае экспонент, будем рассматривать только меры углов с основанием $a > 0$, т. е. только меры углов, отображающие $\boldsymbol{R}^+$ в S в положительном направлении. Другими словами, будем всегда считать, что $(A \circ f)\left(\frac{a}{4}\right) = i$, где число $a > 0$ — главный период функции $(A \circ f)$. Экспоненты и меры углов с основанием $a < 0$ легко сводятся к экспонентам и мерам углов с основанием $a > 0$.

Меру углов с основанием $a > 0$ можно определить и чисто аксиоматически.

Определение 11. **Мерой углов с основанием $a > 0$ называется любая мера углов $f(\cdot)$, обладающая дополнительными свойствами:**

3) **a — главный период функции $f(\cdot)$;**

4) **$f\left(\frac{a}{4}\right) = d$, где d — прямой угол.**

Теорема о существовании и единственности меры углов f_a с фиксированным основанием $a > 0$ очевидным образом вытекает из теорем о существовании и единственности экспоненты (теоремы 8 и 9) и теоремы 10, так как $f_a(\cdot) = A^{-1} \circ e_a(\cdot)$.

Определение 12. **Будем говорить, что:**

а) **угол α имеет меру x по основанию $a > 0$ (числовая мера угла α равна x, мера угла α равна x), если $\alpha = f_a(x)$;**

б) **угол α имеет главную меру x по основанию $a > 0$ (главная мера угла α равна x), если**

$$\alpha = f_a(x) \text{ и } x \in [0;\ a[;$$

в) **радианная мера угла α равна x, если $\alpha = f_{2\pi}(x)$;**

г) **градусная мера угла α равна x, если $\alpha = f_{360}(x)$.**

Отметим, что главная мера угла является биекцией между множествами An и $[0,\ a)$. Поэтому можно образовать функцию, обратную к этой биекции, которая будет иметь вид $An \to [0,\ a)$. Допуская вольность речи, эту функцию будем также называть *главной мерой угла*.

Рассмотрим конкретный угол $\alpha = -d$. Тогда:

а) мера угла $(-d)$ по основанию $a > 0$ равна любому числу из множества $\left\{\ldots, \frac{-5a}{4}, -\frac{a}{4}, \frac{3a}{4}, \ldots\right\}$. Действительно, значение экспоненты $e_a(x)$ для любого числа x из этого множества равно $-i$ и $A^{-1}(-i) = -d$, т. е. $-d = A^{-1} \circ e_a(x) = f_a(x)$;

б) главная мера угла $(-d)$ по основанию $a>0$ равна числу $\frac{3a}{4}$; в частности, при $a=2\pi$ получаем число $\frac{3\pi}{2}$;

в) главная радианная мера угла $(-d)$ равна числу $\frac{3\pi}{2}$;

г) главная градусная мера угла $(-d)$ равна числу 270.

Эти подсчеты можно провести, пользуясь формулой $e_a(x)=e^{\frac{2\pi}{a}ix}$ (см. § 5). Например, если $x=\frac{3a}{4}$, то $e^{\frac{2\pi}{a}ix}=e^{i\frac{3\pi}{2}}=\cos\frac{3\pi}{2}+i\sin\frac{3\pi}{2}$, и поэтому $-d=A^{-1}(-i)=A^{-1}(e^{\frac{2\pi}{a}i\cdot\frac{3a}{4}})=f_a\left(\frac{3a}{4}\right)$, т. е. число $\frac{3a}{4}$ — главная мера угла $(-d)$ по основанию $a>0$. Или другой пример, когда $a=2\pi$ и $x=\pi$: $e_a(x)=e^{\frac{2\pi}{2\pi}i\pi}=e^{i\pi}=-1$. Учитывая $A^{-1}(-1)=A^{-1}(i)+A^{-1}(i)=d+d=2d$, получаем, что число π — главная радианная мера угла $2d$.

Упражнение 16. *а) В соответствии с определением 12 найдите различные меры углов d, $2d$, $3d$, ..., nd, где nd по определению угол, равный $\underbrace{d+\ldots+d}_{n}$, $n\in\mathbf{N}$.*

б) Определим углы $\frac{d}{2}$ и $\frac{d}{3}$ как решения уравнений $2x=d$ и $3x=d$. Докажите, что $\frac{d}{2}=A^{-1}\left(\frac{\sqrt{2}}{2}\pm i\frac{\sqrt{2}}{2}\right)$ и $\frac{d}{3}=A^{-1}\left(\frac{\sqrt{3}}{2}\pm\pm i\frac{1}{2}\right)$. (Используйте изоморфизм $A(\cdot)$ и свойства умножения комплексных чисел.)

в) Найдите меры углов $\pm\frac{d}{2}$; $\pm\frac{d}{3}$; $\pm4\frac{d}{3}$.

Предложение 3. ***Если x — мера угла α по основанию 2π, то $\cos x=\cos\alpha$ и $\sin x=\sin\alpha$.***

▷ Действительно,

$$\alpha=f_{2\pi}(x)\Leftrightarrow\alpha=A^{-1}\circ e_{2\pi}(x)\Leftrightarrow A(\alpha)=e_{2\pi}(x)\Leftrightarrow$$
$$\Leftrightarrow(P_1A(\alpha)=P_1(e_{2\pi}(x))\wedge(P_2A(\alpha)=P_2(e_{2\pi}(x))\Leftrightarrow$$
$$\Leftrightarrow(\cos\alpha=\cos x)\wedge(\sin\alpha=\sin x).\ \blacksquare$$

Это важное наблюдение позволяет определить тригонометрические функции числового аргумента на основе тригонометрических функций углового аргумента. При этом тригонометрические функции угла естественно определить по любому положительному основанию, а не только по основанию 2π.

Определение 13. **Косинусом (синусом) числа x по основанию $a>0$ называется косинус (синус) такого угла α, мера которого по основанию a равна x:**

$$(\cos_a x=\cos\alpha)\wedge(\sin_a x=\sin\alpha)\rightleftarrows\alpha=f_a(x).$$

Предложение 4.

$$\cos_a x=\cos\left(\frac{2\pi x}{a}\right)\ \text{и}\ \sin_a x=\sin\left(\frac{2\pi x}{a}\right),$$

где $\cos(\cdot)$ и $\sin(\cdot)$ — тригонометрические функции числа, введенные в § 5.

▷ $(\cos_a x=\cos\alpha)\wedge(\sin_a x=\sin\alpha)\Leftrightarrow M_a(x)=\alpha\Leftrightarrow\alpha=$
$=A^{-1}\circ e_a(x)\Leftrightarrow A(\alpha)=e_a(x)=e_1\left(\frac{x}{a}\right)=e_{2\pi}\left(\frac{2\pi x}{a}\right)=e\left(\frac{2\pi x}{a}\right)$. Следовательно, у точек единичной окружности $A(\alpha)$ и $e\left(\frac{2\pi x}{a}\right)$ совпадают первая и вторая координаты, т. е. $\cos\alpha=\cos\frac{2\pi x}{a}$ и $\sin\alpha=\sin\frac{2\pi x}{a}$. ■

Предложение 4 позволяет получать свойства функций $\cos_a x$ и $\sin_a x$ на основе свойств функций $\cos x$ и $\sin x$.

Упражнение 17. *Докажите, что:*

а) функции $\cos_a x$ и $\sin_a x$ непрерывны и имеют главный период a;

б) $e_a(x)=\cos_a x+i\sin_a x$;

в) функция $\cos_a(\cdot)$ четна и функция $\sin_a(\cdot)$ нечетна;

г) $e_a(-x)=\cos_a x-i\sin_a x$ и выразите функции $\cos_a x$ и $\sin_a x$ через функцию $e_a(\cdot)$ и, в частности, через функцию $e_1(\cdot)$;

д) $\cos_b(x)=\cos_a\left(\frac{ax}{b}\right)$ (формула перехода к новому основанию).

Обычно рассматриваются меры углов с основанием a, равным 2π, или 360, или 400.

Важно сравнить следующие определения:

радианом называется угол, главная мера которого по основанию 2π равна числу 1, т. е. 1 рад $\rightleftarrows f_{2\pi}(1)=A^{-1}\circ e_{2\pi}(1)=A^{-1}\circ e_1\left(\frac{1}{2\pi}\right)$;

градусом называется угол, главная мера которого по основанию 360 равна числу 1, т. е. $1^\circ\rightleftarrows f_{360}(1)=A^{-1}\circ e_{360}(1)=A^{-1}\circ e_1\left(\frac{1}{360}\right)$;

десятичным градусом называется угол, главная мера которого по основанию 400 равна числу 1, т. е.

1 десятичный градус $\rightleftarrows f_{400}(1)=A^{-1}\circ e_{400}(1)=A^{-1}\circ e_1\left(\frac{1}{400}\right)$;

градусом по основанию $a>0$ называется угол, главная мера которого по основанию a равна числу 1, т. е.

$$1_a\rightleftarrows f_a(1)=A^{-1}\circ e_a(1)=A^{-1}\circ e_1\left(\frac{1}{a}\right).$$

Уравнение $z^n=1$ имеет ровно n корней, и все они лежат на

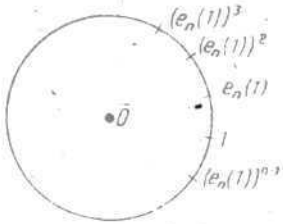


Рис. 20

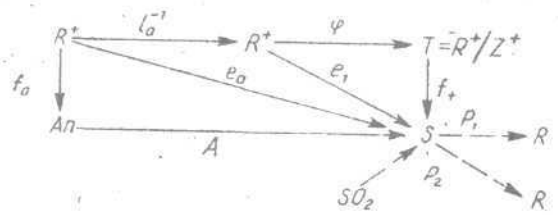


Рис. 21

единичной окружности, так как условие $z^n=1$ влечет $|z|^n=1$, $|z|\in R_{>0}$, $|z|=1$. Среди них по многим причинам выделяется такой корень z_0, для которого все остальные корни являются его натуральными степенями. Такой корень называют первообразным. Комплексное число $e_n(1)$ является как раз первообразным корнем уравнения $z^n=1$. Действительно, $(e_n(1))^n=e_n(n\cdot 1)=1$, и все числа вида $(e_n(1))^j$, где $1\leqslant j\leqslant n$, $j\in Z$, являются корнями этого уравнения: $((e_n(1))^j)^n=((e_n(1))^n)^j=1^j=1$. В этом смысле число $e_n(1)=e^{\frac{2\pi}{n}i}$ является n-й частью окружности S (рис. 20). Как мы увидим, длина дуги $\overset{\frown}{1,\ e_n(1)}$ равна числу $\frac{2\pi}{n}$, и в этом смысле число $e_n(1)$, и дугу $\overset{\frown}{1,\ e_n(1)}$, и угол, «опирающийся» на эту дугу, можно называть n-й частью окружности. Поэтому говорят, что угол $1^\circ=A^{-1}\circ e_{360}(1)$ является «одной трехсотшестидесятой частью окружности» или «одной девяностой частью прямого угла», так как

$$90\cdot(A^{-1}\circ e_{360}(1))=A^{-1}(e_{360}(1)^{90})=A^{-1}(e_{360}(90))=A^{-1}(i)=d.$$

Аналогично для одного десятичного градуса:

$$100\cdot(A^{-1}\circ e_{400}(1))=A^{-1}(e_{400}(1)^{100})=$$
$$=A^{-1}(e_{400}(100))=A^{-1}\left(e_{400}\left(\frac{400}{4}\right)\right)=A^{-1}(i)=d.$$

Поэтому угол 1 десятичный градус можно назвать «одной четырехсотой частью окружности» или «одной сотой частью прямого угла». Для 1 радиана подобные выкладки не проходят (почему?).

5. Обсуждение полученных результатов. Рассмотрим диаграмму, составленную из непрерывных гомоморфизмов (рис. 21), по которой вспомним оба определения тригонометрических функций числового аргумента, отношение между этими определениями и определением тригонометрических функций углового аргумента.

На этой диаграмме жирными стрелками указаны основные непрерывные гомоморфизмы, а обычными стрелками — производные (вспомогательные) непрерывные гомоморфизмы. Пунктиром обозначены функции, которые не являются гомоморфизмами, а также изоморфизм между группами SO_2 и S, использование которого носит чисто технический характер.

Например, для определения тригонометрических функций угла нужно пройти по стрелке $A(\cdot)$ и затем по стрелке $p_1(\cdot)$ и

$p_2(\cdot)$. Первое определение тригонометрических функций числа заключается в прохождении стрелки $e_a(\cdot)$ (или l_a^{-1} и затем e_1, или l_a^{-1} и затем φ и f_+) и снова стрелок $p_1(\cdot)$ и $p_2(\cdot)$.

При втором определении тригонометрических функций числа нужно пройти стрелку $f_a(\cdot)$, затем $A(\cdot)$ и $p_1(\cdot)$ и $p_2(\cdot)$. Для установления связи между этими подходами следует вспомнить, что пройти по $f_a(\cdot)$ — это то же самое, что пройти по $e_a(\cdot)$ и затем по $A(\cdot)$ в обратном направлении.

Наконец, следует подчеркнуть, что наиболее нетривиальным моментом во всей этой схеме является существование стрелки (изоморфизма) $f_+(\cdot)$ (теорема 3): без нее нельзя было бы включить в схему группу T.

Вернемся теперь к вопросу, который мы обсуждали во введении к этой главе (полезно еще раз прочесть его). А именно к вопросу о том, не существует ли функция, которую можно отнести к базисным элементарным функциям с тем же основанием, как и изученные выше функции от линейной до тригонометрической. В точной постановке этот вопрос состоит в том, существует ли непрерывный гомоморфизм основных числовых групп (к которым во всяком случае относятся группы $\boldsymbol{R}^+$, $\boldsymbol{R}^\cdot$ и S), который не выражался бы с помощью базисных операций и уже рассмотренных непрерывных гомоморфизмов. Как мы увидим, ответ на этот вопрос отрицательный; иными словами, нет оснований отнести еще какую-то функцию к числу базисных элементарных функций. Для решения этого вопроса построим таблицу всех непрерывных гомоморфизмов между любыми двумя парами групп из множества числовых групп $\{\boldsymbol{R}^+, \boldsymbol{R}^\cdot, S\}$.

Пусть f — непрерывный гомоморфизм вида $f:\boldsymbol{R}^\cdot\to S$. Выберем какое-нибудь основание показательной функции, например, $a=2$ и рассмотрим непрерывный гомоморфизм $2^x:\boldsymbol{R}^+\to\boldsymbol{R}^\cdot$. Эта функция является биекцией $2^x:\boldsymbol{R}\leftrightarrow\boldsymbol{R}_{>0}$ и изоморфизмом групп $\boldsymbol{R}^+$ и $\boldsymbol{R}^\cdot$. Образуем композицию непрерывных гомоморфизмов $f\circ 2^x$. Она является непрерывным гомоморфизмом вида $\boldsymbol{R}^+\to S$, и следовательно, $f\circ 2^x\equiv e^{ix/a}$ или $f\circ 2^x\equiv 1$. Подставляя вместо аргумента x функцию, обратную к функции 2^x, т. е. функцию $\log_2 x$, получим $f(x)=e^{i/a\cdot\log_2 x}$, или $f(x)\equiv 1$. Поэтому $p_1(f)=\cos\frac{1}{a}\cdot\log_2 x$ и $p_2(f)=\sin\frac{1}{a}\cdot\log_2 x$ или $f(x)\equiv 1$. Все эти функции, очевидно, элементарные. Итак, обе половинки любого непрерывного гомоморфизма вида $f:\boldsymbol{R}^\cdot\to S$ являются элементарными функциями.

Пусть $f:S\to\boldsymbol{R}^+$ — какой-то непрерывный гомоморфизм. Рассмотрим композицию непрерывных гомоморфизмов вида $f\circ e^{it}:\boldsymbol{R}^+\to\boldsymbol{R}^+$, где $t\in\boldsymbol{R}$. Получился непрерывный гомоморфизм, и по таблице 1 имеем $f\circ e^{it}\equiv a\cdot t$, где a — какое-то фиксированное число. Итак, функция f вида $f:e^{it}\mapsto a\cdot t$. Как превратить функцию вида $f:S\to\boldsymbol{R}$ в функцию вида $g:\boldsymbol{R}\to\boldsymbol{R}$? Естественный ответ состоит в том, что нужно в f подставить функции, «обратные» к проекциям $p_1(\cdot)$

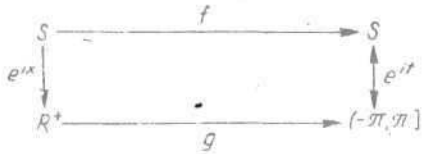


Рис. 22

и $p_2(\cdot)$, т. е. точке x сопоставить точку z из S_+, где S_+ — верхняя полуокружность, или точке y сопоставить точку z из правой полуокружности. Здесь $-1 \leqslant x \leqslant +1$ и $-1 \leqslant y \leqslant +1$. Аналогичным образом можно считать, что z из нижней или левой полуокружности, но ничего нового при этом не получится (проверьте). Точка z по x вычисляется в виде $z = e^{i \arccos x}$ и по y в виде $z = e^{i \arcsin x}$.

Вспомним, что $R(\arccos x) = [0, \pi]$ и $R(\arcsin x) = \left[-\frac{\pi}{2}, +\frac{\pi}{2}\right]$. Поэтому из функции $f(\cdot)$ получаем $g(x) = f(z) = f(e^{i \arccos x}) = a \cdot \arccos x$ и аналогично $g(x) = f(z) = f(e^{i \arcsin x}) = a \cdot \arcsin x$. Итак, $\arccos x$ и $\arcsin x$ — две половинки непрерывного гомоморфизма вида $f: e^{it} \mapsto a \cdot t$. Таким образом, все гомоморфизмы вида $f: S \to \boldsymbol{R}^+$ сводятся к функциям $a \cdot \arccos x$ и $a \cdot \arcsin x$, которые являются элементарными функциями.

Пусть $f: S \to \boldsymbol{R}$ — какой-то непрерывный гомоморфизм. Выбирая, например, изоморфизм $y = \log_2 x: \boldsymbol{R}^* \to \boldsymbol{R}^+$, получаем непрерывный гомоморфизм вида $((\log_2 x) \circ f): S \to \boldsymbol{R}^+$. Как было только что проверено, он имеет вид $e^{it} \mapsto at$, поэтому $\log_2(f(e^{it})) = a \cdot t$, $f(e^{it}) = 2^{a \cdot t}$, где a — какое-то фиксированное вещественное число. Половинки функции f являются элементарными функциями вида $2^{a \cdot \arccos x}$ и $2^{a \cdot \arcsin x}$. Таким образом, половинки всех непрерывных гомоморфизмов вида $S \to \boldsymbol{R}$ суть элементарные функции.

Наконец, пусть f — какой-то непрерывный гомоморфизм вида $f: S \to S$. Рассмотрим диаграмму (рис. 22), где функция g определяется как композиция $e^{it} \circ f \circ (e^{ix})^{-1}$. Это определение корректно, так как функция e^{it} на отрезке $(-\pi, +\pi]$ биективна, и поэтому существует функция $(e^{it})^{-1}$. Функция g непрерывна и является «локальным гомоморфизмом», т. е. для достаточно малых чисел x, y выполняются соотношения $g(x+y) = g(x) + g(y)$ и $g(-x) = -g(x)$. Требование малости x и y существенно, так как иначе может оказаться, что числа $g(x) + g(y)$ или $-g(x)$ не принадлежат отрезку $(-\pi, \pi]$. Используя функцию g, нетрудно показать, что функция f вида $f: e^{it} \mapsto e^{iat}$, где a — фиксированное вещественное число. Аналогично предыдущему одна из половинок этой функции вида $x \mapsto e^{i \arccos x} \mapsto e^{ia \cdot \arccos x} \mapsto \cos(a \cdot \arccos x)$, т. е. вида $g(x) = \cos(a \cdot \arccos x)$. Другая аналогичного вида. Таким образом, и в этом случае половинки всех непрерывных гомоморфизмов вида $S \to S$ — элементарные функции.

Итак, соберем все непрерывные гомоморфизмы рассмотренных числовых групп в одну таблицу, продолжающую таблицу 1 из введения к этой главе:

Таблица 2

	R^+	R^*	S
R^+	$a \cdot x$	$a^{\log_2 x} (a \neq 1)$	$e^{2\pi i a t}$
R^*	$\log_a x (a \neq 0)$	x^a	$e^{a \log 2^{\pi i} t}$
S	$e^{it} \mapsto a \cdot t$	$e^{it} \mapsto 2^{ait}$	$e^{it} \mapsto e^{ait}$

Остается последний вопрос: почему рассматриваются именно такие числовые группы, как $\boldsymbol{R}^+$, $\boldsymbol{R}^*$ и S? Все они являются подгруппами полей $\boldsymbol{R}$ и $\boldsymbol{C}$. И поэтому можно спросить, почему рассматриваются такие числовые поля, как $\boldsymbol{R}$ и $\boldsymbol{C}$. Выбор именно этих полей в качестве базисных (основных) полей в элементарной математике обсуждается в главе V. Впрочем, и без детальной аргументации ясно, что поля $\boldsymbol{R}$ и $\boldsymbol{C}$ суть основные поля по крайней мере в элементарной математике и ее приложениях.

Подгруппа какого-то поля K — это по определению подгруппа его аддитивной группы K^+ или его мультипликативной группы K^*. Итак, кандидатами на рассмотрение являются подгруппы следующих групп: $\boldsymbol{R}^+$, $\boldsymbol{R}^*$, $\boldsymbol{C}^+$ и $\boldsymbol{C}^*$. Разумно рассматривать нетривиальные подгруппы, носители которых связны. По крайней мере это простейшие подгруппы, кроме тривиальных. Тривиальные группы приводят к постоянным функциям и поэтому не представляют интереса. Если G — связная подгруппа $\boldsymbol{R}^+$ и $G \neq \{0\}$, то G содержит какое-то $\varepsilon \neq 0$ и интервал $(0, \varepsilon)$. Если $\varepsilon < 0$, то его можно заменить на $(-\varepsilon)$. Следовательно, для любого числа $\lambda \in \boldsymbol{R}_{\geqslant 0}$ существует такое $x \in (0, \varepsilon)$ и $n \in \boldsymbol{N}$, что $\lambda \leqslant n\varepsilon = \underbrace{\varepsilon + \ldots + \varepsilon}_{n}$. Тогда $n \cdot \varepsilon \in G$, $(0, n \cdot \varepsilon) \subseteq G$, $\lambda \in G$. Поэтому $\boldsymbol{R} \subseteq G$ и $G \subseteq \boldsymbol{R}$, т. е. $\boldsymbol{R} = G$. Если G — подгруппа $\boldsymbol{R}^*$ и $G \neq \{1\}$, то существует $\varepsilon \neq 1$, $\varepsilon \in G$ и $(1, \varepsilon) \subseteq G$, и аналогично получаем, что $\boldsymbol{R}_{>1} \subseteq G$, $\boldsymbol{R}_{>0} = G$. Итак, $\boldsymbol{R}^+$ и $\boldsymbol{R}^*$ — единственные связные подгруппы поля $\boldsymbol{R}$.

Для поля $\boldsymbol{C}$ дело обстоит чуть сложнее. Для краткости изложения заменим свойство связности группы G на также очень естественное свойство компактности группы G. Если G — компактная подгруппа поля $\boldsymbol{C}$, то G — ограниченное множество в обычной метрике пространства $\boldsymbol{C}$. Легко доказать, что S — единственная ограниченная и, в частности, компактная подгруппа поля $\boldsymbol{C}$.

Глава II

ВЕКТОР. ПЛОСКОСТЬ. ПЛАНИМЕТРИЯ

ВВЕДЕНИЕ

Целью этой главы является сравнение различных подходов к определению и пониманию основных геометрических объектов: вектора, плоскости, планиметрии.

Для упрощения вычислений мы ограничимся двумерным случаем. Аналогичное сравнение в более высоких размерностях отличается лишь удлинением технических выкладок.

Напомним, что векторным пространством X над полем K называется алгебраическая система вида $\langle X, +, -, 0, \cdot, K\rangle$, где X — некоторое множество с определенной в нем двуместной операцией $+$ вида $X\times X\to X$, относительно которой X является коммутативной группой, а также двуместной операцией вида: $\cdot : K\times X\to X$, где K — носитель поля K. При этом операция умножения $\cdot$ обладает обычными свойствами:

$$\lambda\cdot(x+y)=\lambda x+\lambda y,$$
$$(\lambda+\mu)\cdot x=\lambda x+\mu x,$$
$$\lambda\cdot(\mu\cdot x)=(\lambda\cdot\mu)\cdot x,$$
$$1_K\cdot x=x,$$

где 1_K — единичный элемент поля и λ, μ, $1_K\in K$, x, $y\in X$. (Одну двуместную операцию $\cdot : K\times X\to\mathscr{X}$ часто трактуют как бесконечное семейство одноместных операций $w_\lambda:\mathscr{X}\to\mathscr{X}$, где $w_\lambda(x)\rightleftarrows\lambda\cdot x$ и λ пробегает K.)

Можно рассматривать векторные пространства над телом K и даже над кольцом K (в последнем случае вместо «векторное пространство» говорят «модуль над K»).

Если между векторными пространствами X_1 и X_2 над одним и тем же телом K можно установить линейную биекцию $f:X_1\leftrightarrow X_2$, т. е. такую биекцию, при которой выполняется $f(\lambda x+\mu y)=\lambda f(x)+\mu f(y)$, то $\mathscr{X}_1$ и $\mathscr{X}_2$ называются изоморфными пространствами, а функция f называется линейной биекцией или линейным изоморфизмом. Изоморфные пространства часто отождествляются. Наконец, размерностью векторного пространства $\mathscr{X}$ над телом K называется натуральное число $\dim_K\mathscr{X}$, равное максимальному числу линейно независимых элементов в $\mathscr{X}$ (если, конечно, такое натуральное число существует). Если в векторном пространстве X существует хотя бы одна бесконечная система векторов, всякая конечная подсистема которой линейно независима, то говорят, что X бесконечномерно. Напомним, что любые векторные пространства над K одной и той же (конечной) размерности изоморфны между собой. В дальнейшем принимается, что $\forall n\in\boldsymbol{N}(\underbrace{1+\ldots+1}_{n}\neq 0)$.

Векторное пространство X над K размерности 2 назовем векторной плоскостью. При таком аксиоматическом определении векторной плоскости как всегда возникает вопрос о существовании хотя бы одного примера векторной плоскости. Иными словами, возникает вопрос о существовании хотя бы одной модели соответствующей аксиоматики.

Типичным примером векторной плоскости (моделью аксиоматики векторной плоскости) является множество K^2, где K — носитель поля (или тела) K, с покоординатно определенными операциями сложения и умножения на скаляры. Эту векторную плоскость обозначают K^2. В качестве поля K часто рассматривают поля $\boldsymbol{R}$ или $\boldsymbol{C}$. Можно построить много примеров векторных плоскостей над $\boldsymbol{R}$, которые хотя все и изоморфны векторной плоскости $\boldsymbol{R}^2$, тем не менее по их интуитивному пониманию и по сфере их приложений совсем не похожи друг на друга и на плоскость $\boldsymbol{R}^2$. Эти примеры отражают различные подходы к понятию вектора, обычно определяемого как элемент произвольной векторной плоскости над $\boldsymbol{R}$. В них сначала определяются сами векторы, т. е. элементы будущего множества $\mathscr{X}$, затем — операции над векторами и только после этого возникает векторная плоскость $\mathscr{X}$ над $\boldsymbol{R}$. Каждый из этих подходов (примеров аксиоматически определенной векторной плоскости) в той или иной мере использует множество $\boldsymbol{R}^2$, поэтому их нельзя считать вполне независимыми друг от друга. Тем не менее, знакомство с каждым из них кажется полезным для осознания всей широты концепции вектора. Эти подходы к понятию вектора когда-то мотивировали аксиоматическое определение векторного пространства.

Такие два пути: от аксиоматического определения к примеру (модели) и от примера (модели) к аксиомам — типичны для всей математики.

Два определения понятия вектора считаются эквивалентными, если соответствующие этим определениям векторные пространства изоморфны. Параграф 1 содержит примеры векторных плоскостей.

Параграфы 2 и 3 посвящены ретроспективному обсуждению интуитивного понятия плоскости. При этом рассматриваются два подхода: один из них использует понятие векторной плоскости, как уже известное, заранее заданное понятие (подход Вейля). Второй из них является чисто аксиоматическим подходом в том смысле, что он не связан ни с какими предварительными представлениями о векторах, числах и т. п. объектах; этот подход основан на аксиоматике инцидентности или аксиоматике Евклида — Гильберта (здесь же приводятся аксиоматики Лобачевского и Римана).

В связи с аксиоматическим подходом очень кратко рассматривается та структура, в рамках которой определяются все обычные

модели евклидовой и неевклидовой планиметрий; а именно — двумерное риманово многообразие.

В § 4—5 обсуждается роль теории групп в геометрии. В § 4 описываются основные из таких групп, а в § 5 изучаются такие объекты и величины, которые остаются неизменными при преобразованиях плоскости под действием фиксированной группы G. Все такие объекты и величины по определению образуют планиметрию группы G. Этот подход, наряду с аксиоматическим, позволяет определить понятие планиметрии (клейновский подход). В качестве G рассматриваются подгруппы аффинной группы.

§ 1. СРАВНЕНИЕ РАЗЛИЧНЫХ ПОДХОДОВ К ПОНЯТИЮ ВЕКТОРА

1. Вектор как пара чисел. Свободный вектор. Вектор как параллельный перенос. Мы приведем шесть подходов к понятию вектора. В первом из них вектором называется пара вещественных чисел. В последующих определениях прямо или косвенно используются числа или пары чисел и, следовательно, эти подходы взаимозависимы. Тем не менее каждый из этих подходов имеет самостоятельное значение и в соответствующем круге вопросов и приложений математики принимается за исходное определение вектора. Если мы примем какое-то одно определение вектора за исходное, то все другие можно рассматривать как свойства определяемого понятия.

Итак, мы приводим ряд определений вектора, нумеруя их 1а, 1б, ..., и после каждого из них определяем функцию ψ, которая является изоморфизмом между множеством всех векторов в смысле определения 1а и множеством всех векторов в смысле текущего определения. Поэтому после этих определений (в теореме 1) можно сделать вывод, что все они эквивалентны.

а) Определение 1а. **Вектором называется произвольная упорядоченная пара вещественных чисел.**

В этом случае покоординатно определяются операции сложения векторов и умножения вектора на скаляр — вещественное число. Другими словами, вектором называется произвольный элемент из множества $\boldsymbol{R}^2$, и в этом множестве рассматриваются обычные операции $+$ и $\cdot$.

Очевидно, это определение сохраняет смысл, если в нем заменить поле $\boldsymbol{R}$ на произвольное тело K.

б) Определение 1б. **Вектором называется произвольный класс эквивалентности в множестве $(\boldsymbol{R}^2\times\boldsymbol{R}^2)$ относительно отношения эквивалентности:**

$$\langle a, b\rangle \sim \langle c, d\rangle \rightleftarrows (c-a=d-b),$$

где $a, b, c, d\in\boldsymbol{R}^2$.

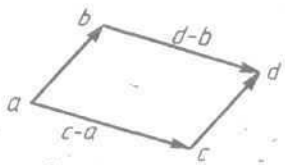


Рис. 23

Смысл такого отношения эквивалентности виден из рисунка 23. Класс эквивалентности с представителем $\langle a, b\rangle$, как всегда, обозначается $[\langle a, b\rangle]$. Сложение векторов и их умножение на числа определяются естественным образом:

$$[\langle a, b\rangle]+[\langle a', b'\rangle]\rightleftarrows[\langle a+a', b+b'\rangle] \text{ и}$$
$$\lambda\cdot[\langle a, b\rangle]\rightleftarrows[\langle \lambda\cdot a, \lambda\cdot b\rangle].$$

Итак, вектор в этом смысле — это как раз то, что называют свободным вектором. Множество всех векторов в смысле определения 1б совпадает с фактор-множеством $(\boldsymbol{R}^2\times\boldsymbol{R}^2)/\sim$, или, что то же самое, с $(\boldsymbol{R}^4)/\sim$.

Чтобы связать первое определение со вторым, положим:

$$\psi(x)\rightleftarrows[\langle 0, x\rangle],\ \psi:\boldsymbol{R}^2\to((\boldsymbol{R}^2\times\boldsymbol{R}^2)/\sim).$$

Функция $\psi(\cdot)$ — изоморфизм. Действительно, функция $\psi(\cdot)$ инъективна: если $x\neq y$ и $\psi(x)=\psi(y)$, то $\langle 0, x\rangle\sim\langle 0, y\rangle$, $0=y-x$, $x=y$, — противоречие. Она сюръективна: в любом классе эквивалентности $[\langle a, b\rangle]$ из $(\boldsymbol{R}^2\times\boldsymbol{R}^2)/\sim$ содержится представитель вида $\langle 0, x\rangle$, где $x=b-a$, так как $a-0=b-(b-a)$. Ясно, что $\psi(x)=[\langle 0, x\rangle]$. Поэтому ψ биекция. Она сохраняет операции, т. е. $\psi(x+y)=\psi(x)+\psi(y)$ и $\psi(\lambda\cdot x)=\lambda\cdot\psi(x)$. Проверьте это самостоятельно. Итак, ψ биекция и гомоморфизм и, следовательно, изоморфизм.

в) Следующее определение вектора связано с его трактовкой как параллельного переноса плоскости, т. е. некоторой биекции (преобразования) плоскости в себя.

Параллельный перенос в плоскости $\boldsymbol{R}^2$ определяется как отображение $f_\xi:\boldsymbol{R}^2\to\boldsymbol{R}^2$, для которого $f_\xi(x)\rightleftarrows x+\xi$, где ξ — фиксированный элемент $\boldsymbol{R}^2$. Ясно, что f_ξ действительно биекция $\boldsymbol{R}^2$ на себя. Операции над такими биекциями определяются как $(f_\xi+f_\eta)\rightleftarrows f_\xi\circ f_\eta$ и $\lambda\cdot f_\xi\rightleftarrows f_{\lambda\cdot\xi}=H_\lambda\circ f_\xi\circ H_{1/\lambda}$, где H_λ — гомотетия с коэффициентом λ.

Часто дают другое определение параллельного переноса, которое использует только метрику и понятие прямой, а именно: параллельным переносом называется такая биекция $f:\boldsymbol{R}^2\to\boldsymbol{R}^2$, что для любых двух различных точек $a, b\in\boldsymbol{R}^2$ выполняются условия: $\rho(a, f(a))=\rho(b, f(b))$ и $\rho(a, b)=\rho(f(a), f(b))$, причем прямая $(f(a), f(b))$ либо совпадает с прямой (a, b), либо не пересекается с ней. Здесь ρ обычная метрика в $\boldsymbol{R}^2$, т. е.

$$\rho(a, b)=\sqrt{(b_1-a_1)^2+(b_2-a_2)^2},$$

где $a=\begin{pmatrix}a_1\\a_2\end{pmatrix}$, $b=\begin{pmatrix}b_1\\b_2\end{pmatrix}$ (см. рис. 24). Сложение параллельных переносов опять-таки определяется как их композиция, а $\lambda\cdot f$, где например $\lambda>0$ определяется условиями (если $f\neq id_{\boldsymbol{R}^2}$):

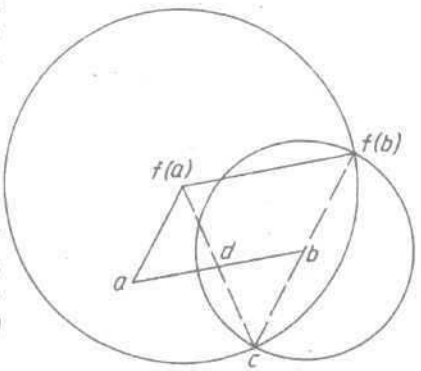


Рис. 24

для любой точки $a \in \boldsymbol{R}^2$ точка $(\lambda \cdot f)(a)$ лежит на луче $\overrightarrow{a, f(a)}$ и $\rho(a, (\lambda \cdot f)(a)) = \lambda \cdot \rho(a, f(a))$, если $f = id_{\boldsymbol{R}^2}$, то $\lambda \cdot f = f$.

Теперь можно определить вектор как параллельный перенос. Это и есть наше третье определение вектора.

Можно определить параллельный перенос, не используя метрику в $\boldsymbol{R}^2$. Такое определение предпочтительнее, так как с геометрической точки зрения в плоскости нет какой-то одной, выделенной метрики. Например, метрики в $\boldsymbol{R}^2$

$$\rho(a, b) \rightleftarrows |a_1 - b_1| + |a_2 - b_2| \text{ и}$$
$$\rho(a, b) \rightleftarrows \max\{|a_1 - b_1|, |a_2 - b_2|\},$$

как и любые другие обычные метрики ничем не хуже метрики

$$\rho(a, b) \rightleftarrows \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}\,.$$

Иными словами, при аксиоматическом определении плоскости метрика обычно не входит в число исходных понятий; если она и может быть определена как производное понятие (например, в аксиоматике Евклида — Гильберта), то только в результате сложных построений.

Итак, займемся определением параллельного переноса, которое вместо метрики использует аддитивную группу $(\boldsymbol{R}^2)^+$.

Множество всех биекций любого фиксированного множества X на себя обозначим Is (от лат. isomorfism).

В множестве $\mathrm{Is}(X)$ бинарная операция определяется как композиция двух биекций $g \circ f \rightleftarrows g(f(\cdot))$.

Упражнение 1. а) *Укажите нейтральный элемент в множестве* $\mathrm{Is}(X)$ *относительно операции композиции.*

б) *Укажите в* $\mathrm{Is}(X)$ *элемент, обратный к элементу* f *относительно операции композиции.*

в) *Докажите, что* $\mathrm{Is}(X)$ *— группа относительно операции композиции.*

г) *Является ли она коммутативной группой?*

Каждому элементу $\xi \in \boldsymbol{R}^2$ сопоставим определенную биекцию f_ξ из $\mathrm{Is}(\boldsymbol{R}^2)$, определяемую равенством $f_\xi(x) = x + \xi$. Возникает отображение $F : \xi \mapsto f_\xi$ из множества $\boldsymbol{R}^2$ в множество $\mathrm{Is}(\boldsymbol{R}^2)$. Напомним, что бинарной операцией в группе $(\boldsymbol{R}^2)^+$ является операция покоординатного сложения элементов множества $\boldsymbol{R}^2$.

Упражнение 2. *Докажите, что:*

а) *Отображение* $F : \xi \mapsto f_\xi$ *является гомоморфизмом группы* $(\boldsymbol{R}^2)^+$ *в группу* $\mathrm{Is}(\boldsymbol{R}^2)$;

б) *Ядро* $\mathrm{Ker}(F)$ *гомоморфизма* F *состоит только из нуля группы* $(\boldsymbol{R}^2)^+$, *что эквивалентно утверждению: функция* F *инъективна;*

в) *Для любых двух точек* a *и* b *из множества* $\boldsymbol{R}^2$ *найдется такой элемент* ξ *из группы* $(\boldsymbol{R}^2)^+$, *что* $f_\xi(a) = b$.

Обобщим рассмотренное выше отображение $F : \xi \mapsto f_\xi$. Для этого группу $(\boldsymbol{R}^2)^+$ заменим произвольной группой $\langle G, \cdot, -1, e\rangle$ и множество $\boldsymbol{R}^2$ — произвольным множеством X. Всякий гомоморфизм F группы G в группу $\mathrm{Is}(X)$ называется действием группы G

на множестве X. Отметим, что здесь не только F является отображением (функцией), но и каждое значение отображения F, т. е. каждое $F(g)$, также является функцией вида $X \to X$, которую будем обозначать f_g; не нужно путать $F : G \to \mathrm{Is}(X)$ и $f_g : X \to X$.

Действие F группы G на множестве X называется эффективным, если отображение F инъективно, или, что то же самое, если ядро F состоит только из нейтрального элемента e группы G.

Действие F группы G на множестве X называется транзитивным, если для любых элементов x_1 и x_2 из X найдется такой элемент g из G, что $f_g(x_1) = x_2$.

По упражнению 2 построенное выше конкретное отображение $F : \xi \mapsto f_\xi$ группы $(\boldsymbol{R}^2)^+$ в $\mathrm{Is}(\boldsymbol{R}^2)$ является эффективным и транзитивным действием группы $(\boldsymbol{R}^2)^+$ на множестве $\boldsymbol{R}^2$.

Определение 1в. **Пусть F — произвольное эффективное и транзитивное действие группы $(\boldsymbol{R}^2)^+$ на произвольном множестве X. Параллельным переносом (вектором) называется любое значение отображения F, т. е. любая биекция из множества $\mathrm{Is}(X)$ вида $F(\xi)$, где $\xi \in D(F)$. Как и в общем случае, это значение обозначим f_ξ.**

Множество всех векторов в этом смысле совпадает с областью значений отображения F, т. е. с $R(F)$. Сумму векторов и умножение вектора на число определим с помощью действия F и одноименных операций в $\boldsymbol{R}^2$, учитывая что отображение F является биекцией между $\boldsymbol{R}^2$ и своей областью значений $R(F)$, а именно положим

$$f_\xi + f_\eta \rightleftarrows f_{\xi+\eta} \quad \text{и} \quad \lambda \cdot (f_\xi) \rightleftarrows f_{\lambda \cdot \xi}\,, \qquad (*)$$

где $\lambda \in \boldsymbol{R}$, $\xi, \eta \in \boldsymbol{R}^2$. Интересно, что сложение в этом смысле векторов можно определить без использования $\boldsymbol{R}^2$, как композицию

$$f_\xi + f_\eta \rightleftarrows f_\xi \circ f_\eta = f_{\xi+\eta}.$$

Это соответствует тому, что $(\boldsymbol{R}^2)^+$ как группа (и даже как векторное пространство над $\boldsymbol{Q}$) изоморфна любой группе $(\boldsymbol{R}^n)^+$, где $n \geqslant 1$. Но чтобы определить умножение векторов на скаляры по формуле $(*)$, необходима именно группа $(\boldsymbol{R}^2)^+$: если $n = 1$, то чему равно $\lambda \cdot \xi$ в формуле $(*)$? Если $n = 3$, то в $R(F)$, где F — действие $(\boldsymbol{R}^3)^+$ в X, найдутся три линейно независимых вектора в смысле определения 1в.

Чтобы сравнить это определение вектора с его первым определением, положим $\psi(\xi) \rightleftarrows F(\xi) : \boldsymbol{R}^2 \to R(F)$. Биективность функции ψ сразу вытекает из условия эффективности действия F и того, что $R(F)$ — область значений F. Гомоморфность ψ сразу вытекает из условия, что F — действие.

До сих пор мы не использовали условие транзитивности действия F. Какова его роль? Оно обеспечивает то обстоятельство, что множество X обладает свойствами «обычной плоскости». Точнее говоря, оно обеспечивает выполнение в X двух аксиом Вейля:

1) **Для любой точки a из X функция вида $\boldsymbol{R}^2 \to X$, определяемая соотношением $\xi \mapsto f_\xi(a)$, где ξ — переменная и a — фиксировано, является биекцией** (отсюда сразу вытекает, например, что X равномощно с $\boldsymbol{R}^2$).

По аксиоме 1) для любых точек a и b из X существует равно одно $\xi \in \boldsymbol{R}^2$, для которого выполняется $f_\xi(a) = b$. Такое ξ обозначим $g(a, b)$. Ясно, что $g : X^2 \to \boldsymbol{R}^2$. Интуитивно $g(a, b)$ — тот элемент из $(\boldsymbol{R}^2)^+$, который переводит точку a в точку b.

2) $\boldsymbol{\forall a, b, c \in X (g(a, b) + g(b, c) + g(c, a) = 0)}$.

Из этих аксиом легко вывести, что

$$\forall a, b \in X (g(a, b) = -g(b, a)).$$

Поэтому вторая аксиома выражает собой обычное правило параллелограмма: $g(a, b) + g(b, c) = g(a, c)$.

Итак, благодаря транзитивности действия F любое f_ξ не просто «вектор в произвольном множестве X», а «вектор в множестве X подобном обычной плоскости».

(Интересующиеся геометрией, конечно, заметили, что с помощью транзитивного и эффективного действия F группы $(\boldsymbol{R}^2)^+$ мы определили аффинную плоскость X, которая подробнее рассматривается в § 2.1.)

Последнее определение параллельного переноса сохраняет смысл, если в нем заменить группу $(\boldsymbol{R}^2)^+$ произвольной группой G.

2. Вектор как дифференцирование. Вектор как класс касающихся кривых. Пусть ξ — фиксированный элемент из $\boldsymbol{R}^2$. Каждой дифференцируемой в точке $\overline{O} \rightleftarrows \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ функции вида $f : \boldsymbol{R}^2 \to \boldsymbol{R}$ поставим в соответствие число, которое будем обозначать $d_\xi(f)$. Возникает отображение d_ξ из множества всех таких функций f в $\boldsymbol{R}$. Оказывается, что отображение d_ξ можно определить аксиоматически, не используя самого вектора ξ. Поэтому, отождествляя ξ и d_ξ, вектором можно назвать само отображение d_ξ, определенное аксиоматически, без использования ξ. Такова вкратце идея следующего способа определения вектора. Перейдем к точным формулировкам.

Обозначим A множество всех непрерывно дифференцируемых в точке $\overline{O}$ функций вида $f : \mathcal{O}_f \to \boldsymbol{R}$, где $\mathcal{O}_f$ некоторая зависящая от f окрестность точки $\overline{O}$. Условие непрерывной дифференцируемости означает, что первые частные производные функции f существуют во всех точках окрестности $\mathcal{O}_f$, а в самой точке $\overline{O}$ эти частные производные непрерывны. Первые частные производные обозначают $f'_x(\cdot, \cdot)$ и $f'_y(\cdot, \cdot)$ или $f'_1(\cdot, \cdot)$ и $f'_2(\cdot, \cdot)$. Элементы множества A можно поточечно складывать и умножать между собой, а также умножать на вещественные числа:

$$(f_1 + f_2)(z) \rightleftarrows f_1(z) + f_2(z), \quad z \in \mathcal{O}_{f_1} \cap \mathcal{O}_{f_2},$$
$$(f_1 \cdot f_2)(z) \rightleftarrows f_1(z) \cdot f_2(z), \quad z \in \mathcal{O}_{f_1} \cap \mathcal{O}_{f_2},$$
$$(\lambda \cdot f)(z) \rightleftarrows \lambda \cdot (f(z)), \quad \lambda \in \boldsymbol{R}.$$

При этом выполняются обычные свойства операций сложения, умножения и умножения на вещественные числа. Поэтому множество A вместе с этими операциями в нем является не только векторным пространством над полем $\boldsymbol{R}$, но и алгеброй над полем $\boldsymbol{R}$. Можно доказать, что векторное пространство A бесконечномерно над полем $\boldsymbol{R}$.

Фиксируем элемент ξ из $\boldsymbol{R}^2$. Каждой функции $f \in A$ поставим в соответствие число (называемое производной функции f в точке $\overline{O}$ по направлению вектора ξ), равное

$$d_\xi(f) \rightleftarrows \lim_{t \to 0} \frac{f(t \cdot \xi) - f(0)}{t}.$$

Возникает отображение вида $d_\xi(\cdot) : A \to \boldsymbol{R}$.

Легко доказать, что

$$d_\xi(f) = f'_x(0, 0) \cdot \xi_1 + f'_y(0, 0) \cdot \xi_2,$$

где $\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$. Иными словами, $d_\xi(f)$ является скалярным произведением вида $(f'(\overline{O}), \xi)$, где $f'(\overline{O}) \rightleftarrows \begin{pmatrix} f'_x(0, 0) \\ f'_y(0, 0) \end{pmatrix}$. Проверьте эти утверждения.

Можно по-другому объяснить, что такое $d_\xi(\cdot)$, а именно: дифференциалом функции $f(x, y) : \boldsymbol{R}^2 \to \boldsymbol{R}$ в точке a называется выражение вида $f'_x(a) \cdot dx + f'_y(a) \cdot dy$ $(*)$.

В этой записи за таинственными обозначениями dx и dy скрываются аргументы отображения. Запишем их в виде $\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$, где $\xi_1 = dx$ и $\xi_2 = dy$, т. е. $\xi = \begin{pmatrix} dx \\ dy \end{pmatrix}$.

Поэтому аргументами дифференциала являются f, a и ξ, и дифференциал, как отображение, определяется формулой $(*)$, т. е.

$$d(f, a, \xi) \rightleftarrows (f'(a), \xi),$$

где $f'(a) \rightleftarrows \begin{pmatrix} f'_x(a) \\ f'_y(a) \end{pmatrix}$. Следовательно, производную функцию $f(\cdot, \cdot)$ в точке $\bar{O}$ по направлению ξ можно определить формулой

$$d\xi(f) \rightleftarrows d(f, \bar{O}, \xi).$$

Отображение $d(f, \bar{O}, \xi)$ имеет два аргумента: f и ξ. Из них фиксируем аргумент ξ и будем считать переменным аргумент f. Получится отображение вида $A \to \boldsymbol{R}$, которое и было определено выше как $d_\xi(\cdot) : A \to \boldsymbol{R}$.

Итак, определено отображение $d_\xi(\cdot) : A \to \boldsymbol{R}$.

Упражнение 3. *Докажите, что для любого фиксированного $\xi \in \boldsymbol{R}^2$ выполняются соотношения:*

а) $d_\xi(f + g) = d_\xi(f) + d_\xi(g)$, $\forall f, g \in A$;

б) $d_\xi(\lambda \cdot f) = \lambda \cdot d_\xi(f)$, $\forall \lambda \in \boldsymbol{R}$, $\forall f \in A$;

в) $d_\xi(f \cdot g) = d_\xi(f) \cdot g(\overline{O}) + f(\overline{O}) \cdot d_\xi(g)$, $\forall f, g \in A$.

Функции $P_1(\cdot)$ и $P_2(\cdot)$, как всегда, суть проекции точек из $\boldsymbol{R}^2$ соответственно на первую и вторую координатные оси, т. е. $P_1(\xi)=\xi_1$, $P_2(\xi)=\xi_2$, где $\xi=\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}$. Тогда $P_1,\ P_2\in A$. Если $\xi=\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}$, то $d_\xi(P_1)=\lim\limits_{t\to 0}\frac{P_1(t\cdot\xi)-P_1(0)}{t}=\lim\limits_{t\to 0}\frac{t\cdot\xi_1}{t}=\xi_1$, и аналогично $d_\xi(P_2)=\xi_2$. Отсюда вытекает, что разным векторам ξ и η соответствуют разные отображения $d_\xi(\cdot)$ и $d_\eta(\cdot)$. Действительно, если $d_\xi(\cdot)=d_\eta(\cdot)$, то $\xi_1=d_\xi(P_1)=d_\eta(P_1)=\eta_1$ и $\xi_2=d_\xi(P_2)=d_\eta(P_2)=\eta_2$, т. е. $\xi=\eta$. Поэтому соответствие $\xi\to d_\xi$ инъективно.

В дальнейшем удобно рассматривать отображение $d_\xi(\cdot)$ как определенное не на всей алгебре A, а на ее части: подалгебре B, по определению состоящей из всех функций вида $f: O_f\to\boldsymbol{R}$, которые имеют во всех точках окрестности O_f частные производные сколь угодно высокого порядка. Оказывается, что свойства отображения $d_\xi(\cdot)$, приведенные в упражнении 3, однозначно характеризуют его без использования вектора ξ.

Определение 1г. **Вектором называется любое отображение вида $d: B\to\boldsymbol{R}$ со свойствами:**

1) $d(f+g)=d(f)+d(g)$, $\forall f,\ g\in B$;

2) $d(\lambda\cdot f)=\lambda\cdot d(f)$, $\forall\lambda\in\boldsymbol{R}$, $\forall f\in B$;

3) $d(f\cdot g)=d(f)\cdot g(\overline{O})+f(\overline{O})\cdot d(g)$, $\forall f,\ g\in B$.

Чаще такое отображение называют *дифференцированием*. Условие 3 из определения 1г называют правилом Лейбница, так как оно похоже на равенство $(f\cdot g)'=f'\cdot g+f\cdot g'$.

Векторы в смысле этого определения, конечно, существуют. Например, по упражнению 3 любое отображение $d_\xi(\cdot)$ является вектором. Однако неясно, существует ли хотя бы один вектор в смысле определения 1г, отличный от всех отображений вида d_ξ.

Упражнение 4. *Докажите, что:*

а) A и B — алгебры и B — подалгебра A;

б) множество всех векторов в смысле определения 1г *является векторным пространством над $\boldsymbol{R}$ относительно поточечно определенных операций:*

$$(d_1+d_2)(f)\rightleftarrows d_1(f)+d_2(f)\ \text{и}\ (\lambda\cdot d)(f)\rightleftarrows\lambda\cdot d(f);$$

в) если функция h тождественно равна единице, то для любого вектора $d(\cdot)$ выполняется $d(h)=0$;

г) если h — постоянная функция, то для любого вектора $d(\cdot)$ выполняется $d(h)=0$;

д) если функции f и g совпадают в некоторой окрестности $\mathcal{O}$ точки $\overline{O}$, то для любого вектора $d(\cdot)$ выполняется $d(f)=d(g)$ (указание: используйте существование такой функции $h\in B$, что h равна нулю вне $\mathcal{O}$ и тождественно равна единице в некоторой окрестности $\mathcal{O}_1$ точки $\overline{O}$, замыкание которой содержится в $\mathcal{O}$);

е) пусть отображения $d_x, d_y: B\to\boldsymbol{R}$ определены равенствами

$$d_x(f)\rightleftarrows\lim_{x\to 0}\frac{f(x,0)-f(0,0)}{x}\quad\text{и}\quad d_y(f)\rightleftarrows\lim_{y\to 0}\frac{f(0,y)-f(0,0)}{y}.$$

Они являются векторами. Векторы $d_x(\cdot)$ и $d_y(\cdot)$ часто обозначают $\left.\frac{\partial(\cdot)}{\partial x}\right|_0$ *и* $\left.\frac{\partial(\cdot)}{\partial y}\right|_0$;

ж) выполняется $d_x(P_1)=1$, $d_x(P_2)=0$, $d_y(P_1)=0$, $d_y(P_2)=1$.

Подход к определению вектора, изложенный в этом пункте, широко используется в математике, например в дифференциальной геометрии и топологии при определении касательного вектора к абстрактному многообразию (поверхности, необязательно расположенной в некотором R^n).

Отметим, что из упражнения 4д следует, что дифференцирование на самом деле задано на фактор-множестве множества B, определяемом отношением эквивалентности $f\sim g\rightleftarrows$ (f и g совпадают в некоторой окрестности точки $\overline{O}$). Фактор-множество $B/\sim$ является и фактор-алгеброй. Ее называют алгеброй ростков бесконечно дифференцируемых функций в нуле. Она обозначается $O_2^\infty(\overline{O})$. Аналогично определяются понятия алгебр ростков дифференцируемых и непрерывных функций в нуле. Итак, любой вектор $d(\cdot)$ опускается на $O_2^\infty(\overline{O})$ и является по существу отображением вида $d(\cdot): (O_2^\infty(\overline{O}))\to\boldsymbol{R}$ со свойствами 1—3 из определения 1г.

Сумма двух векторов $d_1(\cdot)$ и $d_2(\cdot)$ определяется как поточечное сложение:

$$(d_1+d_2)(f)\rightleftarrows d_1(f)+d_2(f).$$

Аналогично $(\lambda\cdot d)(f)\rightleftarrows(\lambda\cdot d(f))$.

В случае определения 1г вектора сложнее установить его связь с первым определением вектора. Для этого нужна следующая

Лемма 1. ***Для любого вектора $d(\cdot): B\to\boldsymbol{R}$ выполняется $\forall f\in B\,(d(f)=d(P_1)\cdot d_x(f)+d(P_2)\cdot d_y(f))$, где отображения $d_x(\cdot)$ и $d_y(\cdot)$ определены в упражнении 4е и функции $P_1(\cdot)$, $P_2(\cdot)$ определены после упражнения 3; $P_1(\cdot)$, $P_2(\cdot)\in B$.***

Любой точке $\xi=\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}$ из $\boldsymbol{R}^2$ сопоставим отображение $g_\xi(\cdot)$ вида $B\to\boldsymbol{R}$, определяемое формулой $g_\xi(f)\rightleftarrows\xi_1\cdot d_x(f)+\xi_2\cdot d_y(f)$, где $f\in B$. Легко проверить, что для любого отображения $g_\xi(\cdot)$ выполняются свойства 1—3 из определения 1г, т. е. для любого ξ из $\boldsymbol{R}^2$ отображение g_ξ — вектор в смысле определения 1г. Рассмотрим функцию $\psi(\xi)\rightleftarrows g_\xi$. Она инъективна: если $\xi\neq\eta$, $\xi=\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}$, $\eta=\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}$ и $\psi(\xi)=\psi(\eta)$, то $\psi(\xi)(P_1)=\xi_1\cdot 1+\xi_2\cdot 0=\xi_1$, $\psi(\eta)(P_1)=\eta_1\cdot 1+\eta_2\cdot 0=\eta_1$, т. е. $\xi_1=\eta_1$ и, аналогично, $\xi_2=\eta_2$. Поэтому $\xi=\eta$ — противоречие.

Ценность леммы 1 в том, что она позволяет проверить сюръективность функции ψ. Действительно, для любого вектора d выберем в качестве ξ пару чисел $\begin{pmatrix}d(P_1)\\ d(P_2)\end{pmatrix}$. По лемме 1 получим $\psi(\xi)=d$. Итак, ψ — биекция. Гомоморфность функции ψ очевидна. Доказательство леммы 1 приводится ниже, после теоремы 1.

Следующий способ определения вектора связан с таким наблюдением. Рассмотрим множество всех кривых в $\boldsymbol{R}^2$, проходящих че-

рез начала координат $\bar{O}=\begin{pmatrix}0\\0\end{pmatrix}$ и имеющих в этой точке одну и ту же скорость (один и тот же касательный вектор). Имеют ли две кривые одинаковую скорость в точке $\bar{O}$, можно определить, не используя понятия касательного вектора. Следовательно, это множество кривых можно отождествить с касательным вектором, тем самым интуитивно понимая вектор как множество кривых. Перейдем к точным формулировкам.

Напомним, что кривой (гладкой кривой), проходящей через точку $\bar{O}$, называется любое отображение $\varphi: \boldsymbol{R}\to\boldsymbol{R}^2$, $\varphi(t)=\begin{pmatrix}x(t)\\y(t)\end{pmatrix}\in R^2$, $t\in\boldsymbol{R}$, для которого функции $x(t)$ и $y(t)$ имеют непрерывные первые производные во всех точках их области определения, и, кроме того, $\bar{O}=\varphi(t_0)$ для некоторого $t_0\in\boldsymbol{R}$. (В качестве ооласти определения кривой можно выбрать любой промежуток в $\boldsymbol{R}$, но можно пользоваться и приведенным определением.) Для определенности будем считать, что в качестве точки t_0 всегда выбирается число 0. Множество всех кривых, проходящих через $\bar{O}$, обозначим S.

Определение 1д. **Вектором называется любой класс эквивалентности в множестве S относительно отношения эквивалентности:**

$$\varphi_1\sim\varphi_2 \rightleftarrows \lim_{t\to 0}\frac{\|\varphi_1(t)-\varphi_2(t)\|}{t}=0.$$

Другими словами, φ_1 эквивалентно φ_2, если разность $\varphi_1(t)-\varphi_2(t)$ при $t\to 0$ является бесконечно малой более высокого порядка, чем t.

Суммой $[\varphi_1]+[\varphi_2]$ двух векторов (двух классов эквивалентности) называется вектор (класс эквивалентности) вида $[\varphi_1+\varphi_2]$. Аналогично $\lambda\cdot[\varphi]\rightleftarrows[\lambda\cdot\varphi]$.

Пусть $\varphi\in S$ и $\varphi(t)=\begin{pmatrix}x(t)\\y(t)\end{pmatrix}\in\boldsymbol{R}^2$, где $x(t)$, $y(t): \boldsymbol{R}\to\boldsymbol{R}$. Функцию ψ определим следующим образом: $\psi([\varphi])\rightleftarrows(x'(0), y'(0))$. Прежде всего проверим, что функция ψ определена корректно, т. е. не зависит от выбора представителя в классе $[\varphi]$. Если $\alpha\in[\varphi]$, то $\alpha\sim\varphi$, и

$$\alpha(t)=\begin{pmatrix}u(t)\\v(t)\end{pmatrix}\in\boldsymbol{R}^2, \text{ т. е.}$$

$$\lim_{t\to 0}\frac{u(t)-x(t)}{t}=0.$$

Учитывая, что $u(0)=x(0)=0$, получим $u'(0)=x'(0)$. Аналогично $v'(0)=y'(0)$, т. е. функция ψ корректно определена.

Докажем инъективность функции ψ. Так как ψ гомоморфизм (проверьте), то для этого достаточно убедиться в том, что $\operatorname{Ker}\psi=[[\bar{0}]]$. Если $\psi([\varphi])=(0, 0)$, то по теореме Лагранжа $\frac{x(t)-x(0)}{t}=x'(\lambda_t)$ и $y(t)-y(0)=y'(\mu_t)\cdot t$, где λ_t и μ_t какие-то фиксированные

точки из отрезка $[0, t]$. Функция $\frac{|t|}{t}$ ограничена. Точки λ_t, μ_t стремятся к точке 0 при $t\to 0$. Функции $x'(\lambda_t)$ и $y'(\mu_t)$ стремятся соответственно к числам $x'(0)$ и $y'(0)$, при $t\to 0$. Наконец, $x'(0)=y'(0)=0$. Поэтому получим

$$\frac{\|\varphi(t)\|}{t}=\frac{\sqrt{x(t)^2+y(t)^2}}{t}=\frac{|t|}{t}\cdot(x'(\lambda_t)^2+\ y'(\mu_t)^2)^{\frac{1}{2}}\to 0$$

при $t\to 0$. Следовательно, φ эквивалентна тождественно равной нулю кривой, т. е. $\varphi\in[\bar{0}]$. Итак, функция ψ инъективна.

Докажем ее сюръективность: если $(x, y)\in\boldsymbol{R}^2$ и $\varphi(t)\rightleftarrows(t\cdot x, t\cdot y)$, то $\psi([\varphi])=(x, y)$. Итак, функция ψ — изоморфизм.

Отметим, что в определениях 1г и 1д используется только метрика (расстояние между точками) в $\boldsymbol{R}^2$, т. е. в них $\boldsymbol{R}^2$ рассматривается как метрическое пространство независимо от его векторной структуры. Метрика используется в понятиях предела и производной.

3. Вектор как тензор. Последний подход к пониманию вектора основан на фундаментальной копцепции тензора, которая связана с преобразованием координат фиксированного вектора при переходе от одного базиса в $\boldsymbol{R}^2$ к другому. Мы ограничимся изложением концепции тензора только для интересующего нас случая вектора.

Сначала напомним, как преобразуются координаты фиксированного вектора $\xi=x\cdot e_1+y\cdot e_2$ при переходе от одного базиса $e=\langle e_1, e_2\rangle$ к другому базису $e'=\langle e_1', e_2'\rangle$. Пусть $A_{e'e}=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$ — матрица перехода от базиса e к базису e', т. е. числа a_{ij} удовлетворяют условиям:

$$\begin{cases}e_1'=a_{11}\cdot e_1+a_{12}\cdot e_2,\\ e_2'=a_{21}\cdot e_1+a_{22}\cdot e_2.\end{cases}$$

Из этих условий выразим векторы e_1 и e_2:

$$\begin{cases}e_1=\frac{a_{22}}{\det A_{e',e}}\ e_1'-\frac{a_{12}}{\det A_{e',e}}\cdot e_2',\\ e_2=-\frac{a_{21}}{\det A_{e',e}}\ e_1'+\frac{a_{11}}{\det A_{e',e}}\cdot e_2'.\end{cases}$$

Поэтому матрица перехода от базиса e' к базису e, т. е матрица $A_{e,e'}$, равна $((A_{e',e})^{-1})^t$.

Если $\xi=x'\cdot e_1'+y'\cdot e_2'=x\cdot e_1+y\cdot e_2$, то $x'\cdot(a_{11}e_1+a_{12}e_2)+y'\cdot(a_{21}e_1+a_{22}e_2)=xe_1+ye_2$, $\begin{cases}a_{11}x'+a_{21}y'=x,\\ a_{12}x'+a_{22}y'=y,\end{cases}$ $(A_{e',e})^t\cdot\begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}x\\y\end{pmatrix}$ и $\begin{pmatrix}x'\\y'\end{pmatrix}=(A^t_{e',e})^{-1}\cdot\begin{pmatrix}x\\y\end{pmatrix}$. Отметим, что $(A^t_{e',e})^{-1}=(A^{-1}_{e',e})^t$. Для краткости вместо $A_{e',e}$ иногда будем писать A, Итак,

$$e'=A_{e',e}\cdot e \text{ и } \begin{pmatrix}x'\\y'\end{pmatrix}=(A^{-1}_{e',e})^t\cdot\begin{pmatrix}x\\y\end{pmatrix}.$$

Сравнение этих равенств показывает, что если базисы преобразуются друг в друга матрицей A, то соответствующие координаты одного и того же фиксированного вектора ξ преобразуются матрицей $(A^{-1})^t$.

Определение 1e. **Обозначим E совокупность всех базисов в R^2. Вектором называется любое отображение $f: E \to R^2$, обладающее свойством $\forall e',\ e \in E\,(f(e') = (A_{e',e}^{-1})^t \cdot f(e))$, где $A_{e',e}$ — матрица перехода от базиса e к базису e'.**

Отметим одну особенность этого определения: если известно хотя бы одно значение функции $f(\cdot)$, то любое другое ее значение однозначно вычисляется по нему. Следовательно, если $f(\cdot)$ и $g(\cdot)$ — два вектора и $\exists e \in E(f(e) = g(e))$, то $\forall e' \in E(f(e') = g(e'))$. Кроме линейной, ни одна из привычных функций не обладает таким свойством. Операции над векторами в этом смысле определим поточечно:

$$(f_1 + f_2)(e) \rightleftarrows f_1(e) + f_2(e),\ (\lambda \cdot f)(e) \rightleftarrows (\lambda \cdot f(e)).$$

Упражнение 5. *Докажите, что:*

а) определения операций сложения векторов и умножения вектора на число, сформулированные в 1д и 1е, корректны;

б) множества векторов, определенных в 1д и 1е, образуют векторные плоскости;

в) если $A_{e',e}$ — матрица перехода от базиса e к базису e', то $\det(A_{e',e}) \neq 0$.

г) Используя явный вид матрицы A^{-1}, проверьте, что

$$\det(A^{-1}) = (\det A)^{-1} \text{ и } \det(A^t) = \det A.$$

Определим функцию $\psi(\cdot)$ равенством $\psi(f) \rightleftarrows f(e_0)$, где e_0 — какой-то один фиксированный базис в R^2. Легко видеть, что $\psi(\cdot)$ — гомоморфизм.

Чтобы проверить инъективность $\psi(\cdot)$, убедимся, что ядро $\psi(\cdot)$ тривиально. Пусть $\psi(f) = (0, 0) \in R^2$. Докажем, что тогда $f: E \to R^2$ — нулевой вектор, т. е. функция из E в R^2, тождественно равная нулю. Действительно, по определению f имеем $f(e) = (A_{e,e_0}^{-1})^t f(e_0)$. Так как $f(e_0) = (0, 0)$, то для любого базиса $e \in E$ выполняется $f(e) = (0, 0)$. Итак, функция $\psi(\cdot)$ инъективна. Докажем ее сюръективность. Пусть (x_0, y_0) — любая наперед заданная точка из R^2. Определим отображение $f: E \to R^2$ равенством $f(e) \rightleftarrows (A_{e,e_0}^{-1})^t \cdot (x_0; y_0)$, где e — любой элемент из E. Надо доказать, что определенная таким образом функция $f: E \to R^2$ является вектором, т. е.

$$f(e') = (A_{e',e}^{-1})^t \cdot f(e).$$

Так как матрица перехода A_{e',e_0} от e_0 к e' равна произведению $(A_{e',e}) \cdot (A_{e,e_0})$ матриц перехода от e_0 к e и от e к e', то, учитывая равенство $((BC)^{-1})^t = (C^{-1}B^{-1})^t = (B^{-1})^t (C^{-1})^t$, получаем:

$$f(e') = (A_{e',e_0}^{-1})^t f(e_0) = ((A_{e',e} \cdot A_{e,e_0})^{-1})^t f(e_0) =$$

$$= (A_{e',e}^{-1})^t \cdot (A_{e,e_0}^{-1})^t f(e_0) = (A_{e',e}^{-1})^t f(e).$$

Итак, функция ψ-биекция и гомоморфизм и, следовательно, изоморфизм.

Итак, выше доказана следующая теорема.

Теорема 1. *Определения вектора 1а — 1е эквивалентны.*

▷ Доказательство леммы 1. Нужно проверить, что для любой функции $f \in B$ выполняется $d(f) = d(P_1) \cdot d_x(f) + d(P_2) \cdot d_y(f)$. Фиксируем какую-то точку $\xi = (\xi_1, \xi_2)$ и рассмотрим функцию $g(t) \rightleftarrows f(t \cdot \xi) = f(t \cdot \xi_1, t \cdot \xi_2)$, зависящую только от числового аргумента t, т. е. $g(\cdot): R \to R$. Применим к ней как к функции от t формулу Ньютона-Лейбница на отрезке $[0, 1]$:

$$g(1) - g(0) = f(\xi) - f(\bar{0}) = \int_0^1 g'(t) \cdot dt.$$

Поскольку функция $g(\cdot)$ является композицией функций $t \mapsto t \cdot \xi \mapsto f(t \cdot \xi)$, то по правилу дифференцирования композиции получаем:

$$g'(t) = f'_{\xi_1}(t \cdot \xi) \cdot \xi_1 + f'_{\xi_2}(t \cdot \xi) \cdot \xi_2.$$

Поэтому $f(\xi) - f(\bar{0}) = \xi_1 \cdot \int_0^1 f'_{\xi_1}(t \cdot \xi) \cdot dt + \xi_2 \cdot \int_0^1 f'_{\xi_2}(t \cdot \xi) \cdot dt = P_1(\xi) \cdot f_1(\xi) + P_2(\xi) \cdot f_2(\xi)$, где $f_1(\xi) \rightleftarrows \int_0^1 f'_{\xi_1}(t \cdot \xi) \cdot dt$ и $f_2(\xi) \rightleftarrows \int_0^1 f'_{\xi_2}(t \cdot \xi) \cdot dt$.

В полученное равенство подставим сначала точку $\xi = (\xi_1, 0)$ и затем точку $\xi = (0, \xi_2)$. В первом случае получим $\frac{f(\xi_1, 0) - f(0, 0)}{\xi_1} = f_1(\xi_1, 0)$. Это равенство выполняется при любых ξ_1. Поэтому, переходя в нем к пределу при $\xi_1 \to 0$, получим $f_1(\bar{0}) = f'_{\xi_1}(0, 0)$. Аналогично $f_2(\bar{0}) = f'_{\xi_2}(0, 0)$. Используя упражнение 4г, получим:

$$d(f) = d(f(0) + P_1 \cdot f_1 + P_2 \cdot f_2) = d(f(0)) + d(P_1) \cdot f_1(0) + P_1(0) \cdot d(f_1) + d(P_2) \cdot f_2(0) + P_2(0) \cdot d(f_2) = d(P_1) \cdot d_1(f) + d(P_2) \cdot d_2(f). \blacksquare$$

В этом рассуждении существенно использовалось, что все рассматриваемые функции бесконечно дифференцируемы, так как иначе функции f_1 и f_2 могут не принадлежать области определения отображения d.

§ 2. ПОНЯТИЕ ПЛОСКОСТИ.

1. Аффинная плоскость. Когда в школе формируют интуитивное представление о плоскости, то говорят о плоскости стола или плоскости доски, неограниченно продолженной во все стороны. В высшей математике плоскость часто определяют как R^2. Однако R^2 мало соответствует такому интуитивному представлению о плоскости: в R^2 имеется абсолютно выделенная точка $\bar{0} \rightleftarrows \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, абсолютно выделенные оси x и y, точки с рациональными и иррациональными, алгебраическими и трансцендентными координатами и т. д. Ничего подобного нет в неограниченно продолженной плоскости классной доски. Там все точки и прямые одинаково нераз-

личимы. Чтобы лучше выразить интуитивное представление о плоскости (геометрической плоскости), вместо $\boldsymbol{R}^2$ рассматривают векторную плоскость $\mathscr{X}$. В $\mathscr{X}$ уже нет точек со специфическими координатами, нет выделенных осей. Но в $\mathscr{X}$ еще имеется абсолютно выделенное начало — точка $\overline{O}$. Поэтому делают следующий шаг и переходят к аффинной плоскости.

Напомним, что двумерным аффинным пространством над телом K (аффинной плоскостью над K) называется четверка объектов $A=\langle A, \mathscr{X}_K, g, K\rangle$, где A — произвольное множество, $\mathscr{X}_K$ — двумерное векторное пространство над телом K и g — функция, согласующая множества A и $\mathscr{X}$, а именно функция $g: A^2\to\mathscr{X}$ и она по определению удовлетворяет условиям Вейля (в которых $g(a, \cdot)$ обозначение функции, получающейся из функции $g(\cdot, \cdot)$ фиксированием первого агрумента значением a, где $a\in$A):

1) для любой точки a из A функция $g(a, \cdot): A\to\mathscr{X}$ является биекцией множеств A и $\mathscr{X}$;

2) для любых $a, b, c\in A$ выполняется соотношение

$$g(a, b)+g(b, c)+g(c, a)=0.$$

Удобно представить себе вектор $g(a, b)$ как вектор с началом в точке a и концом в точке b. Чтобы подчеркнуть такое интуитивное представление, часто используют обозначение

$$g(a, b)\rightleftarrows\overrightarrow{a, b}.$$

Элементы множества A называются точками; они соответствуют интуитивному представлению о точках: все они совершенно одинаковы и ни одна из них ничем не выделяется перед другими. Элементы множества $\mathscr{X}$ называются векторами. Таким образом, в понятии аффинной плоскости происходит разделение точек и векторов, чего нет в арифметической и векторной плоскостях. Если фиксирована аффинная плоскость $\langle A, \mathscr{X}_K, g, K\rangle$, то для краткости будем называть аффинной плоскостью и само множество A. В обозначении $\overrightarrow{a, b}$ часто не пишут запятую и опускают стрелку у верхней черты. Аффинную плоскость обозначим A_2.

Упражнение 6. *Докажите, что:*

а) $\overline{a, a}=0$;

б) $\overline{a, b}=-\overline{b, a}$;

в) $\overline{a, b}+\overline{b, c}=\overline{a, c}$;

г) $(\overline{a, b}=0)\Leftrightarrow(a=b)$.

Подчеркнем, что в множестве A нет никаких операций, т. е. A ни группа, ни кольцо, ни векторное пространство. Однако наличие отображения $g: A\times A\to\mathscr{X}$ со свойствами 1—2 позволяет определить действие на множестве A аддитивной группы $\mathscr{X}^+$; в этом смысле всякий вектор из $\mathscr{X}^+$ можно рассматривать как некоторое преобразование множества A. Более точно, фиксируем вектор $\xi\in\mathscr{X}^+$ и каждому элементу $a\in A$ поставим в соответствие такой элемент $b\in A$, что $g(a, b)=\xi$. Этот элемент обозначим $a+\xi$. По

аксиоме 1 для любых a и ξ такой элемент b существует и единствен.

Тем самым определена функция f_ξ вида

$$f_\xi: A\to A,\ f_\xi: a\mapsto a+\xi;$$

она называется параллельным переносом в аффинной плоскости A на вектор ξ. Если ξ пробегает все множество $\mathscr{X}$, то каждому ξ соответствует своя конкретная функция $f_\xi: A\to A$. Получается отображение $F: \mathscr{X}\to\mathrm{Is}(A)$, $F(\xi)\rightleftarrows f_\xi$.

Предложение 1. ***Отображение $F: \xi\mapsto f_\xi$ является эффективным и транзитивным действием аддитивной группы $\mathscr{X}^+$ векторной плоскости $\mathscr{X}$ на множестве A, где A — носитель аффинной плоскости $\langle A, \mathscr{X}_K, g, K\rangle$.***

▷ Докажем, что для фиксированных векторов ξ и η из $\mathscr{X}$ и для всех точек $a\in A$ выполняется $(a+\xi)+\eta=a+(\xi+\eta)$.

Пусть $b=a+\xi$ и $c=(a+\xi)+\eta$. Тогда $\overline{a, b}=\xi$, $\overline{b, c}=\eta$, и поэтому $\overline{a, c}=\overline{a, b}+\overline{b, c}=\xi+\eta$, т. е. $c=a+(\xi+\eta)$.

Перепишем равенство $(a+\xi)+\eta=a+(\xi+\eta)$ в терминах отображения $F: \xi\mapsto f_\xi$ и получим, что $f_\eta\circ f_\xi=f_{\xi+\eta}$. Это означает гомоморфность отображения F, а из нее вытекает равенство $f_{-\xi}\circ f_\xi=\mathrm{id}_A$, где id_A — тождественное отображение множества A на себя.

Если $f_\xi(a_1)=f_\xi(a_2)$, то $a_1=f_{-\xi}\circ f_\xi(a_1)=f_{-\xi}\circ f_\xi(a_2)=a_2$, т. е. функция f_ξ инъективна. Сюръективность функции f_ξ очевидна: $f_\xi(a)=b\Leftrightarrow\xi=\overline{a, b}$, и достаточно положить $a\rightleftarrows b+(-\xi)$. Итак, доказано, что отображение $F: \xi\mapsto f_\xi$ является гомоморфизмом группы $\mathscr{X}^+$ в группу $\mathrm{Is}(A)$. Следовательно, F — действие $\mathscr{X}^+$ на множестве A. Пусть $\xi\in\mathrm{Ker}\,F$, т. е. $f_\xi=\mathrm{id}_A$ или же $f_\xi(a)=a$ для всех $a\in A$. Выберем какую-нибудь одну точку a и по определению f_ξ получим $\xi=\overline{a, a}=0$. Следовательно, F инъективно. Иными словами, действие F эффективно.

Наконец, докажем транзитивность действия F.

Пусть a и b — элементы A. Положим $\xi\rightleftarrows\overline{a, b}$. Тогда $f_\xi(a)=b$. ■

Доказанное предложение представляет интерес в связи с тем, что подводит нас к новому пониманию аффинной плоскости. Оказывается, что задание на произвольном множестве A эффективного и транзитивного действия аддитивной группы $\mathscr{X}^+$ произвольной векторной плоскости $\mathscr{X}$ позволяет рассматривать A как аффинную плоскость над $\mathscr{X}$. (Такое понимание плоскости непосредственно связано с определением 1в вектора.) Говоря точнее, верно утверждение, обратное к предложению 1.

Теорема 2. ***Пусть $F: \xi\mapsto f_\xi$ — произвольное эффективное и транзитивное действие аддитивной группы $\mathscr{X}^+$ какой-то векторной плоскости $\mathscr{X}_K$ на произвольном множестве A. Определим $g: A\times A\to\mathscr{X}$ условием $(g(a, b)=\xi)\rightleftarrows(f_\xi(a)=b)$. Тогда $\langle A, \mathscr{X}_K, g, K\rangle$ — аффинная плоскость.***

▷ Пусть для некоторого $c\in A$ выполняется $f_\xi(c)=c$. Докажем, что тогда $f_\xi(a)=a$ для всех $a\in A$. Действительно, в силу транзитив-

ности действия F для любого $a \in A$ найдется такой вектор $\eta \in \mathscr{X}$, что $f_\eta(c) = a$. Тогда $f_\xi(a) = f_{(\eta+\xi)-\eta}(a) = f_\eta \circ f_\xi \circ f_\eta^{-1}(a) = f_\eta \circ f_\xi(c) = f_\eta(c) = a$.

Поэтому $f_\xi = \mathrm{id}_A$, и в силу эффективности действия F получаем равенство $\xi = \overline{0}$.

Докажем, что $g(a, b)$ определено однозначно. Если $f_{\xi_1}(a) = f_{\xi_2}(a) = b$, то $f_{\xi_1 - \xi_2}(b) = f_{\xi_1} \circ f_{\xi_2}^{-1}(b) = b$, и по доказанному выше $f_{\xi_1 - \xi_2} = \mathrm{id}_A$, т. е. $\xi_1 = \xi_2$. Следовательно, такое равенство $g(a, b) = \xi$ верно только при одном ξ. Существование такого ξ следует из транзитивности действия F. Проверим аксиомы Вейля.

1) Фиксируем $a \in A$. Тогда надо показать, что $g(a, \cdot) : A \to \mathscr{X}$ — биекция.

▷ Если $g(a, b_1) = g(a, b_2) = \xi$, то $f_\xi(a) = b_1$ и $f_\xi(a) = b_2$, и так как f_ξ — однозначное отображение, то $b_1 = b_2$. Сюръективность $g(a, \cdot)$ очевидна: если положить $b = f_\xi(a)$, то $g(a, b) = \xi$.

2) Достаточно показать, что $g(a, b) + g(b, c) = g(a, c)$. Если $\xi = g(a, b)$ и $\eta = g(b, c)$, то $f_\xi(a) = b$, $f_\eta(b) = c$, и поэтому $f_{\xi+\eta}(a) = f_\eta \circ f_\xi(a) = c$, т. е.

$$g(a, c) = \xi + \eta. \blacksquare$$

Например, R^2 можно рассматривать как аффинную плоскость, а именно как $\langle R^2, R^2, g, R \rangle$, где первое R^2 — просто множество, второе R^2 — векторная плоскость над R и $g(a, b) \rightleftarrows b - a$.

Аналогичным образом любую векторную плоскость $\mathscr{X}$ над K (например, $\boldsymbol{C}^2$ над $\boldsymbol{C}$) обычно рассматривают как аффинную плоскость вида $\langle \mathscr{X}, \mathscr{X}, g, K \rangle$, где первое $\mathscr{X}$ — просто множество, второе $\mathscr{X}$ — векторная плоскость над K и $g(\xi, \eta) \rightleftarrows \eta - \xi$.

2. Школьные геометрические понятия в аффинной плоскости. Прямой l в аффинной плоскости $\langle A, \mathscr{X}_K, g, K \rangle$ называется множество точек аффинной плоскости вида $\{a + \lambda \cdot \overrightarrow{a, b} \mid \lambda \in K\}$, где a и b — любые две, не равные между собой точки из A. Вектор $\overrightarrow{a, b}$ называется направляющим вектором прямой l.

Если поле (или тело) K упорядочено (в этом случае порядок всегда линейный), то в аффинной плоскости $\langle A, \mathscr{X}_K, g, K \rangle$ естественным образом определяется понятие луча с началом в точке a, проходящего через точку $b \in A$, $(b \neq a)$, а также и понятие отрезка с концами в точках a и b. А именно: лучом называется множество точек аффинной плоскости вида $\{a + \lambda \cdot \overrightarrow{a, b} \mid \lambda \in K_{\geqslant 0}\}$, где a и b — какие-то фиксированные точки из A, $a \neq b$. Отрезком с концами в точках a и b называется множество точек аффинной плоскости вида $\{a + \lambda \cdot \overrightarrow{a, b} \mid 0_K \leqslant \lambda \leqslant 1_K\}$, $a \neq b$; оно обозначается $[a, b]$.

Отметим, что для отрезка, определяемого точками a и b, не фиксируется, какая из этих точек первая и какая — вторая. Действительно, легко видеть, что множества $\{a + \lambda \cdot \overrightarrow{a, b} \mid 0_K \leqslant \lambda \leqslant 1_K\}$ и $\{b + \alpha \cdot \overrightarrow{b, a} \mid 0_K \leqslant \alpha \leqslant 1_K\}$ совпадают.

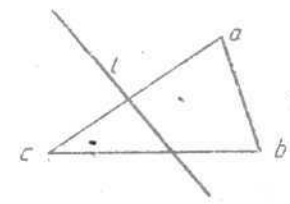


Рис. 25

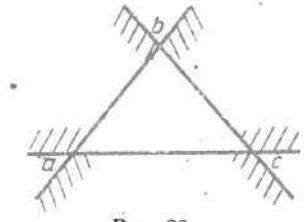


Рис. 26

Далее будем рассматривать случай, когда K совпадает с полем $\boldsymbol{R}$.

Чтобы показать, каким образом переносятся многие геометрические понятия из $\boldsymbol{R}^2$ в аффинную плоскость, в оставшейся части этого пункта остановимся на понятиях многоугольника, площади многоугольника и топологии в аффинной плоскости. Этот материал предполагает известным содержание § 3 из главы III. Поэтому при первом чтении его можно пропустить и перейти к следующему п. 3.

Пусть l — некоторая фиксированная прямая в аффинной плоскости. Рассмотрим в множестве $A \setminus l$ отношение эквивалентности: $(a \sim b) \rightleftarrows ([a, b] \cap l = \varnothing)$, т. е. точки a и b эквивалентны, если отрезок $[a, b]$ содержится в $(A \setminus l)$.

По теореме Паша для любых трех попарно различных точек a, b, c из множества $A \setminus l$, если прямая l (см. рис. 25) пересекает один из отрезков $[a, b]$, $[b, c]$ или $[c, a]$, то она пересекает еще один из этих отрезков, а третий из них не пересекает. Поэтому среди точек a, b, c обязательно найдутся две эквивалентные точки и, следовательно, в рассматриваемом фактор-множестве $(A \setminus l)/\sim$ существует не более двух различных классов эквивалентности. С другой стороны, не трудно проверить, что в $(A \setminus l)$ существуют две не эквивалентные точки. Поэтому в фактор-множестве $(A \setminus l)/\sim$ имеется ровно два класса эквивалентности. Каждый из них по определению называется открытой полуплоскостью. Открытая полуплоскость в объединении с прямой l называется замкнутой полуплоскостью или кратко — полуплоскостью.

Пусть a, b, c — точки аффинной плоскости A, не лежащие на одной прямой. Треугольником с вершинами в точках a, b, c называется пересечение всех полуплоскостей, содержащих эти три точки (рис. 26). Другое эквивалентное определение состоит в том, что треугольником называется пересечение всех выпуклых множеств, содержащих эти три точки. При этом множество X, где $X \subseteq A$, называется выпуклым, если для любых точек x, y из X в X содержится отрезок $[x, y]$.

Чтобы определить понятие многоугольника в аффинной плоскости аналогично определению многоугольника в арифметической плоскости $\boldsymbol{R}^2$, нужно сначала определить понятие внутренности треугольника.

Границей треугольника Δ называется объединение трех отрезков, соединяющих его вершины. Внутренностью треугольника Δ называется множество $(\Delta \setminus \partial\Delta)$, где $\partial\Delta$ — граница треугольника. После этого понятие многоугольника определяется точно так же, как и в случае арифметической плоскости $\boldsymbol{R}^2$.

Многоугольником в аффинной плоскости A называется любое множество, представимое в виде объединения конечного числа треугольников, внутренности которых попарно не пересекаются.

Напомним, что линейной биекцией (изоморфизмом) $f : \mathscr{X} \to \mathscr{Y}$ двух векторных плоскостей $\mathscr{X}$ и $\mathscr{Y}$ называется биекция, обладающая свойством линейности:

$$f(\lambda_1 \cdot \xi + \lambda_2 \cdot \eta) = \lambda_1 \cdot f(\xi) + \lambda_2 \cdot f(\eta), \ \forall \lambda_1, \lambda_2 \in \boldsymbol{R}, \ \forall \xi, \eta \in \mathscr{X}.$$

Теперь мы хотим непосредственно связать понятия многоугольника в арифметической и аффинной плоскостях; это делается с помощью понятия репера. Пусть $e = (e_1, e_2)$ — произвольный базис в векторной плоскости $\mathscr{X}$. Обозначим W_e отображение из множества $\boldsymbol{R}^2$ в $\mathscr{X}$, определяемое равенством

$$W_e(\langle x, y \rangle) \rightleftarrows x \cdot e_1 + y \cdot e_2.$$

Так как e базис в $\mathscr{X}$, то W_e — линейная биекция из $\boldsymbol{R}^2$ на $\mathscr{X}$. Верно и обратное: для любой линейной биекции $W : \boldsymbol{R}^2 \to \mathscr{X}$ найдется такой базис e в $\mathscr{X}$, что $W = W_e$. Действительно, достаточно положить $e_1 \rightleftarrows W(\langle 1, 0 \rangle)$ и $e_2 \rightleftarrows W(\langle 0, 1 \rangle)$. Тогда $W(\langle x, y \rangle) = W(x \cdot \langle 1, 0 \rangle + y \cdot \langle 0, 1 \rangle) = x \cdot e_1 + y \cdot e_2 = W_e(\langle x, y \rangle)$. Поэтому рассмотрение линейных биекций между $\boldsymbol{R}^2$ и $\mathscr{X}$ эквивалентно рассмотрению произвольных базисов в $\mathscr{X}$.

В соответствии с первой аксиомой Вейля для любой фиксированной точки a из A обозначим g_a биекцию $g(a, \cdot) : A \leftrightarrow \mathscr{X}$. Соединяя некоторую биекцию g_a с какой-то линейной биекцией $W : \boldsymbol{R}^2 \leftrightarrow \mathscr{X}$, получаем биекцию $g_a^{-1} \circ W : \boldsymbol{R}^2 \leftrightarrow A$ и обратную биекцию $W^{-1} \circ g_a : A \leftrightarrow \boldsymbol{R}^2$.

Пара $\langle a, W \rangle$ или, что то же самое, пара $\langle a, e \rangle$, где a — точка в A, W — линейная биекция и e — базис в $\mathscr{X}$, называется репером в аффинной плоскости A.

Итак, выбор и фиксирование репера в плоскости A позволяет отождествить плоскости A и $\boldsymbol{R}^2$ с помощью биекций $W^{-1} \circ g_a$ или $W_e^{-1} \circ g_a$. Репер $\langle a, e \rangle$ можно задать, например, тремя не коллинеарными точками в A: первая из них — сама точка a, вторая и третья — такие точки b и c, что $a, b = e_1$ и $a, c = e_2$.

Нетрудно показать, что множество u является многоугольником в A тогда и только тогда, когда для любого репера $\langle a, W \rangle$ множество $W^{-1} \circ g_a(u)$ является многоугольником в $\boldsymbol{R}^2$. Или эквивалентным образом, когда существует репер $\langle a, W \rangle$, для которого множество $W^{-1} \circ g_a(u)$ является многоугольником в $\boldsymbol{R}^2$.

Упражнение 7. *Докажите два последних утверждения.*

Множество всех многоугольников аффинной плоскости A обозначим U_A. В частности таким же образом определяется множество $U_{\mathscr{X}}$ всех многоугольников в векторной плоскости $\mathscr{X}$.

Функцию площади (многоугольника) в аффинной плоскости A можно определить как любую функцию вида $s \circ W_e^{-1} \circ g_a$, где $\langle a, e \rangle$ — какой-то репер в A и s — какая-то площадь в $\boldsymbol{R}^2$. То же самое можно сформулировать чуть иначе.

Определение 2. **Функцией площади в аффинной плоскости $\langle A, \mathscr{X}, g \rangle$ называется любая функция $s_A : U_A \to R_{>0}$, которая для любого репера $\langle a, W \rangle$ порождает функцию площади $s_A \circ g_a^{-1} \circ W$ в $\boldsymbol{R}^2$.**

Удобно показать участвующие в этом определении отображения:

$$\boldsymbol{R}_{>0} \xleftarrow{s} U_{\boldsymbol{R}^2} \xrightarrow{W} U_{\mathscr{X}} \xrightarrow{g_a} U_A \xrightarrow{s_A} \boldsymbol{R}_{>0}.$$

Теперь рассмотрим понятие открытого множества в векторной плоскости $\mathscr{X}$ и в аффинной плоскости A.

Определение 3. **Топология в A (и в $\mathscr{X}$) определяется базой, состоящей из внутренностей всех треугольников.**

Как эта топология связана с топологией в арифметической плоскости $\boldsymbol{R}^2$?

Предложение 2. ***Для любой линейной биекции $W : \boldsymbol{R}^2 \to \mathscr{X}$ и любой точки $a \in A$ функции W и g_a — гомеоморфизмы.***

▷ Так как W биекция, то нужно только проверить, что W и W^{-1} — непрерывные функции. Это следует из того, что внутренности треугольников образуют предбазы топологий в $\boldsymbol{R}^2$ и в X, а для любого треугольника $\triangle$, $W(\mathring{\triangle})$ и $W^{-1}(\mathring{\triangle})$ — внутренности треугольников.

Упражнение 8. *Проверьте, что в векторной плоскости $\mathscr{X}$ треугольник $\triangle$ с вершинами a, b, c и его внутренность $\mathring{\triangle}$ являются множествами соответственно вида*

$$\{\lambda_1 a + \lambda_2 b + \lambda_3 c \mid \lambda_1, \lambda_2, \lambda_3 \geqslant 0, \ \lambda_1 + \lambda_2 + \lambda_3 = 1\}$$

и

$$\{\lambda_1 a + \lambda_2 b + \lambda_3 c \mid \lambda_1, \lambda_2, \lambda_3 > 0, \ \lambda_1 + \lambda_2 + \lambda_3 = 1\}.$$

Аналогично, так как g_d — биекция, то нужно только проверить непрерывность функций g_d и g_d^{-1}, где $d \in A$. Действительно, если w — выпуклое множество, содержащее точки a, b, c, то $g_d(w)$ — выпуклое множество, содержащее точки $g_d(a)$, $g_d(b)$, $g_d(c)$, так как $\{g(d, x + \lambda x, y) | 0 \leqslant \lambda \leqslant 1\}$ совпадает с отрезком в $\mathscr{X}$, соединяющим точки $g(d, x)$ и $g(d, y)$ в $\mathscr{X}$. ■

Поэтому $W^{-1} \circ g_a : A \to \boldsymbol{R}^2$ — гомеоморфизм и, следовательно, топологию в A можно эквивалентным образом определить с помощью репера $\langle a, W \rangle$ и топологии в $\boldsymbol{R}^2$. А именно, множество Y, $Y \subseteq A$ назовем открытым, если существует репер $\langle a, W \rangle$, для которого множество $W^{-1} \circ g_a(Y)$ открыто в $\boldsymbol{R}^2$.

Упражнение 9. *Проверьте, что:*

а) *семейство $\{Y \subseteq A \mid \exists \langle a, W \rangle$ — репер $(W^{-1} \circ g_a(Y)$ — открыто$)\}$ образует топологию в A;*

б) *два определения топологии в аффинной плоскости эквивалентны.*

Предложение 3. ***Пусть*** $\langle a_1, W_1\rangle$ ***и*** $\langle a_2, W_2\rangle$ ***— любые два репера в аффинной плоскости*** A ***и*** $Y\subseteq A$. ***Если множество*** $W_1^{-1}\circ g_{a_1}(Y)$ ***открыто в*** $\boldsymbol{R}^2$, ***то и множество*** $W_2^{-1}\circ g_{a_1}(Y)$ ***открыто в*** $\boldsymbol{R}^2$. ***Поэтому определение топологии в*** A ***с помощью репера не зависит от выбора репера.***

▷ Второе из этих множеств (обозначим его X_2) выражается через первое (обозначим его X_1) в виде

$$X_2=W_2^{-1}\circ g_{a_2}\circ g_{a_1}^{-1}\circ W_1(X_1),$$

где композиция $W_2^{-1}\circ g_{a_2}\circ g_{a_1}^{-1}\circ W_1$ — гомеоморфизм. Образ открытого множества при гомеоморфизме — открытое множество. ■

Так как векторная плоскость $\mathscr{X}$ является частным случаем аффинной плоскости A, то мы определили топологию и в $\mathscr{X}$. Сформулируйте два аналогичных определения топологии специально для случая векторной плоскости.

Упражнение 10. *Докажите, что:*

а) *топологии в векторной и аффинной плоскости отделимы и локально компактны;*

б) *в топологии векторной плоскости операции сложения векторов и умножения векторов на числа непрерывны;*

в) *в топологии аффинной плоскости любой параллельный перенос* $f_\xi: A\to A$, *где* ξ — фиксированный элемент из $\mathscr{X}$, *является гомеоморфизмом.*

Отметим, что можно доказать совпадение между собой и с построенной выше топологией в $\mathscr{X}$ любых двух отделимых топологий в $\mathscr{X}$, для которых непрерывны главные операции векторной плоскости: сложение векторов и умножение векторов на числа. Поэтому указанную топологию в $\mathscr{X}$ можно определять и аксиоматически.

Упражнение 11. *Проверьте, что топологии в векторной плоскости* X *и аффинной плоскости* A *можно эквивалентным образом определить как наименьшие топологии, относительно которых следующие семейства* $\{W: X\to\boldsymbol{R}^2 \mid W$ *— линейная биекция*$\}$ *и* $\{g_a: X\to a\in A\}$ *состоят из непрерывных функций.*

3. Плоскость с формой. Плоскостью с формой над $\boldsymbol{R}$ называется аффинная плоскость $\langle A, \mathscr{X}, f, \boldsymbol{R}\rangle$, для которой в векторной плоскости $\mathscr{X}$ фиксирована некоторая билинейная функция g, т. е. функция $g: \mathscr{X}\times\mathscr{X}\to\boldsymbol{R}$, обладающая следующими свойствами:

1) $g(\lambda\cdot\xi+\mu\cdot\eta, \zeta)=\lambda\cdot g(\xi, \zeta)+\mu\cdot g(\eta, \zeta)$;

2) $g(\xi, \lambda\cdot\eta+\mu\cdot\zeta)=\lambda\cdot g(\xi, \eta)+\mu\cdot g(\xi, \zeta)$, $\forall\lambda, \mu\in\boldsymbol{R}$, $\forall\xi, \eta, \zeta\in\mathscr{X}$.

Плоскостью с формой над $\boldsymbol{C}$ называется аффинная плоскость $\langle A, \mathscr{X}, f, \boldsymbol{C}\rangle$, для которой в векторной плоскости $\mathscr{X}$ фиксирована некоторая билинейная или полуторалинейная функция $g: \mathscr{X}\times\mathscr{X}\to\boldsymbol{C}$; понятие билинейной функции определяется в этом случае как и выше, а полуторалинейная функция по определению обладает свойствами:

1) $g(\lambda\cdot\xi+\mu\cdot\eta, \zeta)=\lambda\cdot g(\xi, \zeta)+\mu\cdot g(\eta, \zeta)$,

2) $g(\xi, \lambda\cdot\eta+\mu\cdot\zeta)=\bar\lambda\cdot g(\xi, \eta)+\bar\mu\cdot g(\xi, \zeta)$, $\forall\lambda, \mu\in\boldsymbol{C}$, $\forall\xi, \eta, \zeta\in\mathscr{X}$.

Здесь $\bar\lambda$ и $\bar\mu$ — комплексные числа, сопряженные к комплексным числам λ и μ, т. е. $(\lambda=x+iy)\Rightarrow(\bar\lambda=x-iy)$.

Билинейную и полуторалинейную функции часто называют формами.

Любую векторную плоскость можно рассматривать как аффинную плоскость; основным примером плоскости с формой является векторная плоскость $\mathscr{X}$, в которой задана билинейная или полуторалинейная форма g.

Примеры.

1. В $\mathscr{X}$ определим билинейную форму $g(\cdot, \cdot)\equiv 0$.

2. В $\boldsymbol{R}^2$ определим билинейную форму $g(x, y)=x_1\cdot y_1+x_2\cdot y_2$, где $x=\begin{pmatrix}x_1\\x_2\end{pmatrix}$ и $y=\begin{pmatrix}y_1\\y_2\end{pmatrix}$. Эта форма — обычное скалярное произведение; оно обозначается (x, y).

3. В $\boldsymbol{C}^2$ определим полуторалинейную форму

$$g(x, y)=x_1\cdot\bar y_1+x_2\cdot\bar y_2$$

или билинейную форму

$$g_1(x, y)=x_1\cdot y_1+x_2\cdot y_2.$$

4. В $\boldsymbol{R}^2$ определим билинейную форму

$$g(x, y)=\det\begin{pmatrix}x_1 y_1\\x_2 y_2\end{pmatrix}.$$

5. В $\boldsymbol{R}^2$ определим билинейную форму

$$g(x, y)=x_1\cdot y_1-x_2\cdot y_2.$$

Эта форма обычно называется псевдоскалярным произведением и обозначается $(x, y)_1$.

Форма $g(\cdot, \cdot)$ называется:

а) *симметричной*, если $g(x, y)=g(y, x)$, $\forall x, y\in\mathscr{X}$;

б) *кососимметричной*, если $g(x, y)=-g(y, x)$, $\forall x, y\in\mathscr{X}$;

в) *эрмитовой*, если $g(x, y)=\overline{g(y, x)}$, $\forall x, y\in\mathscr{X}$;

г) *положительно определенной*, если $g(x, x)>0$, $\forall x\in\mathscr{X}$, $x\neq 0$;

д) *невырожденной*, если условие $\forall x\in\mathscr{X}(g(x, y)=0)$ влечет $y=0$ и условие $\forall y\in\mathscr{X}(g(x, y)=0)$ влечет $x=0$.

Упражнение 12. *а) Про каждую из форм 1—5 определите, какими свойствами из числа а — д она обладает.*

б) Докажите: из положительной определенности формы вытекает ее невырожденность, обратное утверждение неверно.

в) Докажите, что билинейная форма $g(\cdot, \cdot)$ *на* $\boldsymbol{R}^2$ *невырожденна тогда и только тогда, когда* $\det\begin{pmatrix}g_{11} g_{12}\\g_{21} g_{22}\end{pmatrix}\neq 0$, *где* $g_{ij}\rightleftarrows g(e_i, e_j)$, $i, j=1,2$ и $\langle e_1, e_2\rangle$ — стандартный базис в $\boldsymbol{R}^2$.

г) Докажите, что любая билинейная форма g на $\boldsymbol{R}^2$ представима в виде $g(x, y)=g_{11}\cdot x_1y_1+g_{12}\cdot x_1y_2+g_{21}\cdot x_2y_1+g_{22}\cdot x_2y_2$, *где* $x=(x_1, x_2)$, $y=(y_1, y_2)$ *и коэффициенты* g_{ij} *определены в п. в.*

Теорема 3. ***Для того чтобы функция*** $g(\cdot, \cdot): \mathscr{X}^2\to R$ ***была билинейной, симметричной и положительно определенной, необ-***

ходимо и достаточно, чтобы существовала такая линейная биекция $W: R^2 \to \mathscr{X}$, что $\forall x, y \in R^2 ((x, y) = g(W(x), W(y))$, где $(\cdot, \cdot)$ — обычное скалярное произведение в R^2.

▷ Достаточность. Пусть $W: \boldsymbol{R}^2 \to X$ — такой изоморфизм, что $\forall x, y \in \boldsymbol{R}^2 ((x, y) = g(W(x), W(y)))$. Тогда для любых векторов $\xi, \eta \in \mathscr{X}$ выполняется $g(\xi, \eta) = (W^{-1} \cdot \xi, W^{-1} \cdot \eta) = (W^{-1} \cdot \eta, W^{-1} \cdot \xi) = g(\eta, \xi)$, т. е. функция $g(\cdot, \cdot)$ симметрична. Если вектор $\xi \in \mathscr{X}$ и $\xi \neq 0$, то и $W^{-1}\xi \neq 0$, так как W — изоморфизм. Поэтому $g(\xi, \xi) = (W^{-1}\xi, W^{-1}\xi) > 0$, т. е. функция $g(\cdot, \cdot)$ положительно определена. Далее, для любых векторов $\xi, \eta, \zeta \in \mathscr{X}$ и скаляров $\lambda, \mu \in \boldsymbol{R}$ получаем:

$$g(\lambda \cdot \xi + \mu \cdot \eta, \zeta) = (W^{-1}(\lambda\xi + \mu\eta), W^{-1}\xi) = (\lambda W^{-1}\xi + \mu W^{-1}\eta, W^{-1}\zeta) = \lambda(W^{-1}\xi, W^{-1}\zeta) + \mu(W^{-1}\eta, W^{-1}\zeta) = \lambda g(\xi, \zeta) + \mu g(\eta, \zeta).$$

Аналогично проверяется второе условие из определения билинейности функции $g(\cdot, \cdot)$.

Необходимость. Матрицей формы $g(\cdot, \cdot)$ в каком-то базисе $\{e_1, e_2\}$ назовем следующую матрицу:

$$\begin{pmatrix} g(e_1, e_1) & g(e_1, e_2) \\ g(e_2, e_1) & g(e_2, e_2) \end{pmatrix} \rightleftarrows \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}.$$

Фиксируем в $\mathscr{X}$ некоторый базис $\{e_1, e_2\}$ и по матрице формы $g(\cdot, \cdot)$ в этом базисе построим такой новый базис $\{e_1', e_2'\}$, в котором матрица формы $g(\cdot, \cdot)$ имеет особенно простой вид: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Построение такого базиса выполняется следующим образом.

Если $g_{11} = g_{22} = 0$, то положим $e_1' = e_1 + e_2$ и $e_2' = e_2$. Тогда $g_{11}' \rightleftarrows g(e_1', e_1') = g(e_1 + e_2, e_1 + e_2) = 2g_{12} \neq 0$, так как все числа g_{11}, g_{12}, и g_{22} не могут быть равны нулю в силу невырожденности формы $g(\cdot, \cdot)$. Поэтому с самого начала можно считать, что число g_{11} или число g_{22} отлично от нуля. Поменяв, если надо, местами векторы e_1 и e_2, получим, что $g_{11} \neq 0$.

Положим $e_1' = e_1$ и $e_2' = -\frac{g_{12}}{g_{11}} \cdot e_1 + e_2$.

Тогда $g(e_1', e_2') = g(e_1, -\frac{g_{12}}{g_{11}} \cdot e_1 + e_2) = -\frac{g_{12}}{g_{11}} \cdot g(e_1, e_1) + g(e_1, e_2) = 0 = g(e_2', e_1')$.

При этом $\{e_1', e_2'\}$ — базис, т. е. векторы e_1' и e_2' линейно независимы, так как определитель матрицы перехода $A_{e', e}$ от базиса $\{e_1, e_2\}$ к паре векторов $\{e_1', e_2'\}$ отличен от нуля:

$$\det \begin{pmatrix} 1 & -\frac{g_{12}}{g_{11}} \\ 0 & 1 \end{pmatrix} = 1.$$

Если бы существовало $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \in \boldsymbol{R}^2$, для которого $\alpha_1 e_1' + \alpha_2 e_2' = 0$, то $0 = \alpha_1(a_{11}e_1 + a_{12}e_2) + \alpha_2(a_{21}e_1 + a_{22}e_2) = (a_{11}\alpha_1 + a_{21}\alpha_2)e_1 + (a_{12}\alpha_1 + a_{22}\alpha_2)e_2$.

По условию $e = \{e_1, e_2\}$ — базис, поэтому $A_{e', e}^t \cdot \alpha = 0$, $\det A_{e', e}^t = \det A_{e', e} \neq 0$, $\alpha = (A_{e', e}^t)^{-1} \cdot 0 = 0$ — противоречие.

Итак, в базисе $\{e_1', e_2'\}$ матрица формы $g(\cdot, \cdot)$ имеет вид $\begin{pmatrix} g_{11}' & 0 \\ 0 & g_{22}' \end{pmatrix}$.

При этом $g_{11}' > 0$ и $g_{22}' > 0$, так как форма $g(\cdot, \cdot)$ положительно определена и $g_{11}' = g(e_1', e_1')$, $g_{22}' = g(e_2', e_2')$. Положим $e_1'' \rightleftarrows \frac{1}{\sqrt{g_{11}'}} e_1'$ и $e_2'' \rightleftarrows \frac{1}{\sqrt{g_{22}'}} e_2'$. Тогда в базисе $e'' \rightleftarrows \{e_1'', e_2''\}$ матрица $g(\cdot, \cdot)$ имеет искомый вид $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. В качестве искомого изоморфизма выберем функцию W, которая по определению точке $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \boldsymbol{R}^2$ составляет вектор из $\mathscr{X}$ с координатами $x_{1,2}$ в базисе e'', т. е. $W(x) \rightleftarrows x_1 e_1'' + x_2 e_2''$. Легко убедиться, что эта функция действительно изоморфизм. Для любых точек $x, y \in \boldsymbol{R}^2$, где $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ и $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, получаем $x_1y_1 + x_2y_2 = (x, y)$ и $g(Wx, Wy) = g(x_1e_1'' + x_2e_2'', y_1e_1'' + y_2e_2'') = x_1 \cdot y_1 + x_2 \cdot y_2$. ■

Следствие. Пусть $g(\cdot, \cdot): \mathscr{X}^2 \to \boldsymbol{R}$ — билинейная и симметричная форма, которая необязательно положительно определена. Тогда в $\mathscr{X}$ найдется такой базис $e = (e_1, e_2)$, в котором матрица формы $g(\cdot, \cdot)$ имеет один из следующих видов:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

▷ Если $g(\cdot, \cdot)$ не тождественно равна нулю, то, как и при доказательстве теоремы 3, найдется базис $e' = (e_1', e_2')$, в котором матрица $g(\cdot, \cdot)$ имеет диагональный вид $\begin{pmatrix} g_{11}' & 0 \\ 0 & g_{22}' \end{pmatrix}$. Далее все зависит от знаков g_{11}'. Если $g_{11}' > 0$, то положим $e_1 \rightleftarrows \frac{1}{\sqrt{g_{11}'}} e_1'$ и получим, что $g_{11} = 1$. Если $g_{11}' < 0$, то положим $e_1 \rightleftarrows \frac{1}{\sqrt{-g_{11}'}} e_1'$ и получим, что $g_{11} = \left(\frac{1}{\sqrt{-g_{11}'}}\right)^2 g(e_1', e_1') = -1$. Дальнейшее очевидно. ■

Всякую билинейную, симметричную и положительно определенную форму в векторной плоскости называют *евклидовой*. Векторная плоскость $\mathscr{X}$, в которой фиксирована некоторая евклидова форма, называется *евклидовой векторной плоскостью*. Аффинная плоскость, для которой в соответствующей ей векторной плоскости задана некоторая евклидова форма, называется аффинной евклидовой плоскостью.

Выделение класса евклидовых плоскостей среди всех возможных векторных плоскостей связано с тем, что в евклидовой

плоскости можно с помощью формы определить длину вектора и расстояние между точками.

Предложение 4. ***Пусть $\langle A, \mathscr{X}, f, g\rangle$ — аффинная евклидова плоскость, где $g(\cdot, \cdot)$ — евклидова форма в векторной плоскости $\mathscr{X}$.***

Тогда: а) функция $\|\cdot\| : \mathscr{X} \to \boldsymbol{R}_{\geqslant 0}$, определенная равенством

$$\|x\| \rightleftarrows \sqrt{g(x, x)},$$

является нормой в векторной плоскости $\mathscr{X}$;

б) функция $\rho_A(\cdot, \cdot) : A \times A \to \boldsymbol{R}_{\geqslant 0}$, определенная равенством

$$\rho_A(a, b) \rightleftarrows \|f(a, b)\| = \|\overrightarrow{a, b}\| = \sqrt{g(\overrightarrow{a, b}, \overrightarrow{a, b})},$$

является метрикой в евклидовой плоскости A, инвариантной относительно любого параллельного переноса.

▷ а) По теореме 3 найдется такой изоморфизм $W : \boldsymbol{R}^2 \to \mathscr{X}$, что для всех векторов $\xi, \eta \in \mathscr{X}$ выполняется

$$(W^{-1} \cdot \xi, W^{-1} \cdot \eta) = g(\xi, \eta).$$

Пусть $\|\xi\| = 0$. Тогда $0 = \sqrt{g(\xi, \xi)} = \sqrt{W^{-1} \cdot \xi, W^{-1} \cdot \xi} = \sqrt{x_1^2 + x_2^2}$, где $W^{-1} \cdot \xi = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Следовательно, $x_1 = x_2 = 0$, т. е. $W^{-1}\xi = 0$ и $\xi = W \cdot (W^{-1} \cdot \xi) = 0$. Далее, $\|\lambda \cdot \xi\| = \sqrt{g(\lambda \cdot \xi, \lambda \cdot \xi)} = \sqrt{\lambda^2 g(\xi, \xi)} = |\lambda| \cdot \|\xi\|$ для любого скаляра $\lambda \in \boldsymbol{R}$.

Для проверки того, что функция $\|\cdot\|$ является нормой, осталось доказать неравенство $\|\xi + \eta\| \leqslant \|\xi\| + \|\eta\|$ для всех векторов $\xi, \eta \in \mathscr{X}$.

Пусть $x = W^{-1} \cdot \xi$, $y = W^{-1} \cdot \eta$ и $\|x\|_{R^2}$ — норма элемента x в R^2, т. е. $\|x\|_{R^2} \rightleftarrows \sqrt{x_1^2 + x_2^2}$, где $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Используя неравенство Коши — Буняковского для нормы $\|\cdot\|_{R^2}$, т. е. неравенство $|(x, y)| \leqslant \|x\|_{R^2} \cdot \|y\|_{R^2}$, получим $\|\xi + \eta\|^2 = g(\xi + \eta, \xi + \eta) = (x + y, x + y) = \|x\|_{R^2}^2 + 2(x, y) + \|y\|_{R^2}^2 \leqslant (\|x\|_{R^2} + \|y\|_{R^2})^2 = (g(\xi, \xi) + g(\eta, \eta))^2 = (\|\xi\| + \|\eta\|)^2$.

б) По п. а векторная плоскость $\mathscr{X}$ является векторным пространством с нормой. Поэтому в $\mathscr{X}$, как и во всяком нормированном пространстве, каноническим образом определяется метрика $\rho(\xi, \eta) \rightleftarrows \|\xi - \eta\|$ для любых векторов $\xi, \eta \in \mathscr{X}$. Тот факт, что таким образом определенная функция $\rho : \mathscr{X}^2 \to \boldsymbol{R}_{\geqslant 0}$ является метрикой, непосредственно следует из того, что функция $\|\cdot\|$ является нормой. Более того, функция $\rho(\cdot, \cdot)$ инвариантна относительно параллельных переносов, т. е. для любых векторов $\xi, \eta, \zeta \in X$ выполняется

$$\rho(\xi + \zeta, \eta + \zeta) = \|(\xi + \zeta) - (\eta + \zeta)\| = \|\xi - \eta\| = \rho(\xi, \eta).$$

Докажем теперь, что функция $\rho_A : A^2 \to \boldsymbol{R}_{\geqslant 0}$, определенная равенством $\rho_A(a, b) \rightleftarrows \|\overrightarrow{a, b}\|$ для всех точек $a, b \in A$, является

метрикой в A. Если $\rho(a, b) = 0$, то $\overrightarrow{a, b} = 0$, и по свойствам функции $g(a, b) \rightleftarrows \overrightarrow{a, b} : A^2 \to \mathscr{X}$ получаем, что $a = b$ (см. упражнение 6 в п. 1 этого параграфа).

По свойствам этой же функции $g : A^2 \to X$ имеем также, что $\overrightarrow{b, a} = -\overrightarrow{a, b}$, и, следовательно, $\rho_A(b, a) = \|\overrightarrow{b, a}\| = \|-\overrightarrow{a, b}\| = \|\overrightarrow{a, b}\| = \rho(a, b)$.

Наконец, учитывая равенство $\overrightarrow{a, b} + \overrightarrow{b, c} = \overrightarrow{a, c}$ (снова упражнение 6 в п. 1 этого параграфа), получаем, что

$$\rho_A(a, c) = \|\overrightarrow{a, c}\| = \|\overrightarrow{a, b} + \overrightarrow{b, c}\| \leqslant \|\overrightarrow{a, b}\| + \|\overrightarrow{b, c}\| = \rho_A(a, b) + \rho_A(b, c).$$

Следовательно, $\rho_A(\cdot, \cdot)$ — метрика в A. Если вектор $\xi \in \mathscr{X}$, то $\rho_A(a + \xi, b + \xi) = \rho_A(a, b)$. Действительно, если $a' = a + \xi$ и $b' = b + \xi$, то $\overrightarrow{a, a'} = \xi = \overrightarrow{b, b'}$ и

$$\overrightarrow{a, b} = \overrightarrow{a, a'} + \overrightarrow{a', b'} + \overrightarrow{b', b} = \xi + \overrightarrow{a', b'} - \xi = \overrightarrow{a', b'} \text{ и}$$

$$\|\overrightarrow{a', b'}\| = \|\overrightarrow{a, b}\|.$$

Следовательно, метрика $\rho_A(\cdot, \cdot)$ инвариантна относительно параллельных переносов. ■

В евклидовой плоскости можно обычным образом определить школьные геометрические понятия. Приведем несколько таких определений. Два вектора ξ и η из $\mathscr{X}$ ортогональны, если $g(\xi, \eta) = 0$. Множество всех векторов, ортогональных к фиксированному ненулевому вектору ξ, обозначим $\xi^\perp$, т. е.

$$\xi^\perp \rightleftarrows \{\eta \in X \mid g(\xi, \eta) = 0\}.$$

Упражнение 13. *Докажите, что:*

а) $\xi^\perp$ — подпространство исходного векторного пространства $\mathscr{X}$;

б) подпространство $\xi^\perp$ не состоит из одного нуля и не совпадает со всем пространством $\mathscr{X}$;

в) подпространство $\xi^\perp$ — прямая;

г) пусть g — какая-нибудь (не обязательно евклидова) форма на $\mathscr{X}$; тогда точно так же определим множество $\xi^\perp$; найдите его для всех тех форм g, которые приводились выше в виде примеров.

4. Проективная плоскость. Понятие проективной плоскости наряду с понятиями арифметической плоскости $\boldsymbol{R}^2$, векторной плоскости $\mathscr{X}$ и аффинной плоскости A_2 является отражением интуитивных представлений о «плоскости стола или плоскости доски, неограниченно продолженных во все стороны». Вопрос о том, какое из этих четырех понятий наиболее адекватно отражает основные свойства реально существующих плоскостей и, в частности, тех плоскостей, которые рассматриваются в физике, представляет собой важную и трудную задачу. Она рассматривается, в частности, в теоретической физике.

Напомним, что вещественной проективной плоскостью называется множество всех прямых в $\boldsymbol{R}^3$, проходящих через какую-то

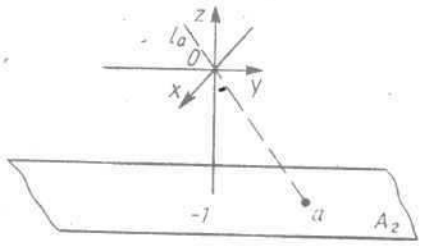


Рис. 27

одну фиксированную точку $\boldsymbol{R}^3$, например через точку $\bar{O}$. Другое эквивалентное определение состоит в том, что вещественной проективной плоскостью называется фактор-множество $(\boldsymbol{R}^3\setminus\{\bar{O}\})/\sim$, где $x\sim y \rightleftarrows \exists\lambda\in\boldsymbol{R}^*(\lambda\cdot x=y)$.

Действительно, любой класс эквивалентности $[x]$ этого фактор-множества состоит из всех точек прямой $l_x \rightleftarrows \{\lambda\cdot x \mid \lambda\in\boldsymbol{R}\}$, кроме начала координат. При этом соответствие $[x]\mapsto l_x$ является биекцией (проверьте).

Точно так же можно определить комплексную проективную плоскость и вообще проективную плоскость над произвольным телом K. А именно, проективной плоскостью $P_2(K)$ над телом K называется фактор-множество $(K^3\setminus\{\bar{0}\})/\sim$, где

$$(x\sim y)\rightleftarrows\exists\lambda\in K^*(\lambda\cdot x=y)$$

и $K^*\rightleftarrows(K\setminus\{0\})$ — мультипликативная группа тела K. Вещественную проективную плоскость $P_2(R)$ будем обозначать P_2. Элементы проективной плоскости будем называть проективными точками.

В P_2 определяется топология; например, как фактор-топология:

$$\mathcal{T}_{P_2}\rightleftarrows\{X\subseteq P_2 \mid \varphi^{-1}(X)\in\mathcal{T}_{(\boldsymbol{R}^3\setminus\{\bar{0}\})}\},$$

где $\varphi: (\boldsymbol{R}^3\setminus\{\bar{0}\})\to P_2$, $\varphi(x)\rightleftarrows[x]$ — каноническое отображение и $\mathcal{T}_{(\boldsymbol{R}^3\setminus\{\bar{0}\})}$ — топология в множестве $\boldsymbol{R}^3\setminus\{\bar{0}\}$, индуцированная топологией в $\boldsymbol{R}^3$.

Для того чтобы пояснить связь между понятиями проективной и аффинной плоскости, рассмотрим в $\boldsymbol{R}^3$ произвольную плоскость, не проходящую через начало координат. Например, плоскость A_2, определяемую уравнением $z=-1$ (см. рис. 27).

Очевидно, структура $\langle A_2, \boldsymbol{R}^2, f\rangle$, где $A_2\rightleftarrows\{(x, y, -1)\mid x, y\in\boldsymbol{R}\}\subseteq\boldsymbol{R}^3$, $f(a, b)\rightleftarrows(b-a)\in\boldsymbol{R}^2$ и $a, b\in A_2$, является аффинной плоскостью с носителем A_2. Здесь точка $b-a$ имеет вид $(x_1-x, y_1-y, 0)$, но мы отождествляем ее с точкой (x_1-x, y_1-y) из $\boldsymbol{R}^2$. Для каждой точки $a\in A_2$ обозначим через l_a прямую, проведенную через точку a и начало координат $\bar{O}$. Тогда соответствие $\psi: a\mapsto l_a$ инъективно, но не сюръективно отображает аффинную плоскость A_2 в проективную плоскость P_2. С помощью соответствия ψ каждую точку аффинной плоскости A_2 можно отождествить с проективной точкой l_a — элементом проективной плоскости P_2 и, следовательно, можно отождествить аффинную плоскость A_2 с некоторым подмножеством проективной плоскости P_2. Именно в этом смысле говорят, что проективная плоскость является «расширением» аффинной плоскости A_2. При этом расширение происходит за счет тех прямых, проходящих через начало координат (проективных точек), которые не являются образами точек a из A_2 при данном соответствии $\psi: a\mapsto l_a$. Другими словами, расширение A_2 происходит за счет проективных точек, которые, как прямые в $\boldsymbol{R}^3$ параллельны плоскости A_2 (в данном случае — за счет всех прямых, проходящих через точку $\bar{O}$ и лежащих в плоскости $z=0$). Такие проективные точки называют бесконечно удаленными.

Множество всех проективных точек, лежащих в какой-то одной плоскости, проходящей через $\bar{O}$, называют проективной прямой. Поэтому множество всех бесконечно удаленных точек, очевидно, образует проективную прямую (в данном случае определяемую плоскостью $z=0$). Эта проективная прямая называется бесконечно удаленной проективной прямой. Если в качестве A_2 вместо $z=-1$ выбрать какую-нибудь иную плоскость и аналогичным образом определить функцию ψ и понятие бесконечно удаленной проективной точки, то, очевидно, бесконечно удаленными окажутся другие проективные точки, которые составят другую бесконечно удаленную проективную прямую.

Поэтому выделение среди проективных точек *бесконечно удаленных* и среди проективных прямых *бесконечно удаленной* зависит от произвола в выборе аффинной плоскости A_2 и, следовательно, субъективно. Иными словами, бесконечно удаленная проективная точка сама по себе, безотносительно к выбору плоскости A_2, ничем не отличается от любой другой проективной точки.

Аналогичным образом обстоит дело с бесконечно удаленной прямой. Для сравнения заметим, что, например, точка $\bar{O}$ в векторной плоскости $\mathcal{X}$ определяется однозначно, независимо ни от каких произвольных выборов и в этом смысле объективно. Каждой прямой l в плоскости A_2 (с уравнением $z=-1$) сопоставим прямую l_∞ в плоскости $z=0$, которая является пересечением плоскости $z=0$ с плоскостью, проходящей через точку $\bar{O}$ и прямую l. Тем самым всякой прямой l в плоскости A_2 соответствует проективная бесконечно удаленная точка l_∞ (рис. 28). Обозначим $\psi_1(l)\rightleftarrows l_\infty$ и получим отображение множества всех прямых в плоскости A_2 в множество всех прямых в плоскости $z=0$, проходящих через точку $\bar{O}$. Отображение $\psi_1(\cdot)$ сюръективно, но не инъективно: всякой проективной точке l_∞ соответствует семейство параллельных прямых l в плоскости A_2: если l параллельно l', то $\psi_1(l)=\psi_1(l')$. Факторизуем мно-

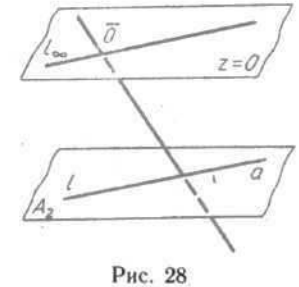


Рис. 28

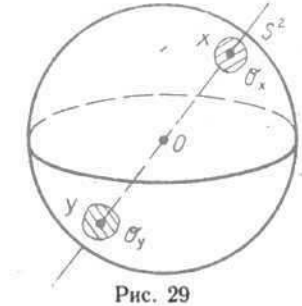


Рис. 29

жество всех прямых в плоскости A_2 отношением эквивалентности $l \| l'$ и на это фактор-множество X опустим отображение ψ_1, т. е. положим $\psi_2([l]) \rightleftarrows \psi_1(l)$. Такое ψ_2 корректно определено и является биекцией между X и множеством всех проективных точек бесконечно удаленной проективной прямой (проверьте!). Классы эквивалентности из X называются еще связками параллельных прямых в A_2. Поэтому бесконечно удаленные проективные точки и связки параллельных прямых в A_2 могут быть отождествлены с помощью отображения ψ_2. Отображение ψ_2 показывает, насколько проективная плоскость P_2 шире аффинной плоскости A_2, а именно P_2 состоит из всех точек A_2 и еще из всевозможных связок параллельных прямых в A_2.

Соответствие $\psi: a \mapsto l_a$ позволяет переносить многие понятия, определенные в аффинной плоскости A_2, в проективную плоскость P_2.

Иногда бывает удобно еще одно определение проективной плоскости. А именно, назовем проективной плоскостью фактор-множество, полученное в результате факторизации сферы S^2 по следующему отношению эквивалентности: $(x \sim_1 y) \rightleftarrows (x = -y)$ (рис. 29). В этом фактор-множестве и определяется фактор-топология; при этом в S_2 рассматривается топология индуцируемая из $\boldsymbol{R}^3$.

Ясно, что каждый класс эквивалентности $[x]$ из фактор-множества $(S^2/\sim_1)$ состоит ровно из двух точек, он однозначно определяет прямую в $\boldsymbol{R}^3$, проходящую через начало координат, т. е. однозначно определяет проективную точку в смысле предыдущего определения проективной плоскости. Эту биекцию вида $(S^2/\sim_1) \mapsto (\boldsymbol{R}^3 \setminus \{\bar{O}\})/\sim$ обозначим f.

Упражнение 14. *Докажите, что:*

а) биекция f — гомеоморфизм; поэтому проективная плоскость P_2 гомеоморфна сфере S^2 с отождествленными диаметрально противоположными точками;

б) канонические отображения φ: $(\boldsymbol{R}^3 \setminus \{\bar{O}\}) \to P_2 \rightleftarrows (\boldsymbol{R}^3 \setminus \{\bar{O}\})/\sim$ и $\varphi_1 : S^2 \to (S^2/\sim_1)$ *непрерывны и открыты;*

в) множество Y, $Y \subseteq P_2$ открыто в P_2 тогда и только тогда, когда множество $X = \{l \cap S^2 / l \in Y\}$ открыто в S^2.

Отметим, что у каждой проективной точки l в P_2 имеется окрестность $\mathcal{O}'$ в P_2, гомеоморфная окрестности в S^2 некоторой точки на сфере. На рисунке 29 такая окрестность $\mathcal{O}'$ состоит из всех прямых, проходящих через точки окрестности $\mathcal{O}_x$ и $\bar{O}$. Соответствующий гомеоморфизм определяется формулой $l \mapsto t_e$, где $l \in \mathcal{O}'$, а t_e точка пересечения прямой l и окрестности $\mathcal{O}_x$. Поэтому проективная плоскость P_2 локально гомеоморфна сфере S^2 и, следовательно, локально гомеоморфна арифметической плоскости $\boldsymbol{R}^2$. Топологические пространства, локально гомеоморфные $\boldsymbol{R}^2$, называются двумерными многообразиями. Например, $\boldsymbol{R}^2$, S^2, P_2 и все обычные поверхности, рассматриваемые в курсе анализа, являются двумерными многообразиями.

Во втором определении проективной плоскости $S^2/\sim_1$ можно заменить сферу S^2 на ее верхнюю половину $S^2_+ \rightleftarrows \{z \in S^2 \mid P_3(z) \geqslant 0\}$. Получится еще одно определение проективной плоскости как фактор-множества $(S^2_+/\sim_1)$ с фактор-топологией в нем. Легко видеть, что проективная плоскость P_2 гомеоморфна этому фактор-пространству, т. е. полусфере S^2_+, у которой на границе отождествлены диаметрально противоположные точки.

Проективные прямые в P_2, как легко видеть, соответствуют в $S^2/\sim_1$ окружностям больших кругов (т. е. сечениям сферы S^2 плоскостями, проходящими через точку $\bar{O}$), у которых диаметрально противоположные точки отождествлены. Такие «окружности» гомеоморфны обычным окружностям, поэтому все проективные прямые гомеоморфны S. В частности, они «замкнуты», в отличие от аффинных прямых: если «идти» по проективной прямой все время «вперед», то когда-нибудь вернешься в ту точку, из которой вышел.

Упражнение 15. *Опишите проективные прямые в проективной плоскости $S^2_+/\sim_1$ и в ее проекции на плоскость Oxy. Эта проекция — гомеоморфизм на круг D, у которого отождествлены противоположные точки границы (так получается еще одно возможное определение проективной плоскости).*

Вероятно, у читателя возник вопрос, в каком смысле эквивалентны пять указанных выше определений проективной плоскости; в каждом из них проективная плоскость определялась как конкретное топологическое пространство, эти пять топологических пространств гомеоморфны между собой; следует ли отсюда, что любое гомеоморфное им топологическое пространство также можно назвать проективной плоскостью. Заметим, что для каждого из этих пяти топологических пространств X определяется и семейство его подмножеств, состоящее из всех проективных прямых в X. При этом указанные гомеоморфизмы переводят проективные прямые одного топологического пространства в проективные прямые другого топологического пространства. Итак, проективная плоскость — нечто большее, чем просто топологическое пространство: в ней определено понятие прямой. Любое топологическое пространство с выделенным в нем семейством подмножеств (элементы которого называются проективными прямыми), гомеоморфное P_2 (причем гомеоморфизм переводит прямые в прямые), можно назвать проективной плоскостью.

Если представить себе, что полусфера S^2_+ сделана из тонкой резины, и попытаться физически осуществить отождествление диаметрально противоположных точек ее границы S, то мы убедимся, что такую операцию нельзя реально осуществить без самопересечений и разрывов. В этом смысле проективная плоскость является примером двумерного многообразия, которое физически не реализуемо в трехмерном пространстве $\boldsymbol{R}^3$.

Чем проективная плоскость предпочтительнее аффинной плоскости? Какая из них более адекватно отражает интуитивное представление о плоскости, а также о реальной физической плоскости?

Не входя подробно в обсуждение этих сложных вопросов, отметим, что в проективной плоскости любые две проективные прямые пересекаются (имеют общие проективные точки). Это непосредственно вытекает из определений проективной точки и проективной плоскости. В случае аффинной плоскости, как и в случаях арифметической и векторной плоскостей, это не так — в них имеются параллельные прямые. Это уникальное свойство проективной плоскости удобно и важно для теории и приложений.

На второй вопрос, казалось бы, нужно сразу ответить, что понятие аффинной плоскости более адекватно отражает интуитивное представление о геометрической плоскости. Однако любая проективная точка в проективной плоскости и проективной прямой имеет окрестность, гомеоморфную окрестности аффинной точки в аффинной плоскости и аффинной прямой. Поэтому локально проективные плоскость и прямая не отличимы от аффинных плоскости и прямой. Глобально они совсем не похожи друг на друга: например, проективная прямая гомеоморфна окружности.

В заключение отметим, что проективную плоскость можно определять и в более абстрактном духе, подобном предшествующему определению аффинной плоскости. А именно, пусть K — произвольное тело (например, $K=\boldsymbol{R}$) и $\mathscr{X}$ — произвольное трехмерное векторное пространство над K. Проективной плоскостью $P_2(K, \mathscr{X})$ назовем множество всех одномерных линейных подпространств в $\mathscr{X}$.

§ 3. АКСИОМАТИЧЕСКИЙ ПОДХОД К ОПРЕДЕЛЕНИЮ ПЛОСКОСТИ

1. Два типа аксиоматического определения плоскости. В предыдущем параграфе понятие плоскости определялось с помощью наперед заданного поля $\boldsymbol{R}$ или арифметической плоскости, которые понимаются как фиксированные множества с фиксированными операциями в них. При таком подходе поле $\boldsymbol{R}$ и арифметическая плоскость $\boldsymbol{R}^2$ считаются заранее известными, индивидуально выделенными математическими объектами, и в этом смысле они являются своего рода константами. Поэтому подход к понятию плоскости в предыдущем параграфе правильно было бы назвать конструктивным или структурным. Он связан с именем Г. Вейля.

Конечно, можно выписать все аксиомы, характеризующие поле $\boldsymbol{R}$ или плоскость $\boldsymbol{R}^2$, и добавить их к аксиомам, характеризующим векторную плоскость, а в случае аффинной плоскости добавить еще две аксиомы Вейля и таким образом получить аксиоматическое (без всяких констант) определение аффинной плоскости. Аналогично можно определить проективную плоскость. Однако такая аксиоматика представляется искусственной. Это проявляется, например, в том, что она будет включать различные мало связанные между собой сорта переменных: переменные по последовательностям, по вещественным числам, по элементам векторной плоскости, по элементам аффинной плоскости и т. д.

Традиция, возникшая после создания общей теории множеств, предлагает определять понятие плоскости с помощью такой аксиоматики, которая с самого начала использует понятия множества и принадлежности одного множества другому множеству (в качестве его элемента). Такой подход к понятию плоскости называют теоретико-множественным. Его иногда путают с евклидовым подходом; однако он принципиально отличается от евклидова подхода тем, что все фигурирующие в нем объекты по определению считаются множествами. А именно, прямые и плоскости считаются множествами точек, прямые — подмножествами плоскостей. Конечно, в евклидовом подходе не имеются в виду никакие множества.

Мы кратко рассмотрим аксиоматические определения плоскости как теоретико-множественного, так и евклидова типа. Начнем с аксиоматических определений первого типа. Они привлекательны замечательной простотой входящих в них аксиом.

Этот параграф помещен вслед за § 2 с целью более полного освещения концепции плоскости. В этом отношении он, конечно, примыкает к предыдущему параграфу. Однако материал этого параграфа не используется в дальнейшем и его можно или вообще опустить или прочесть после изучения главы. Вопросы, затрагиваемые в этом параграфе в связи с логическими аксиомами, выводом, предикатом «истинно», а также теорему 7, можно пропустить.

2. Аксиоматическое теоретико-множественное определение плоскости.

Классической аффинной плоскостью называется пара $\langle X, Y\rangle$, где $Y\subseteq\mathscr{P}(X)$, удовлетворяющая следующим аксиомам (в них для наглядности элементы множества X называются точками, а элементы множества Y называются прямыми).

$\mathbf{A}_1$. Через любые две различные точки можно провести единственную прямую, т. е. $\forall x, y\in X(x\neq y\Rightarrow\exists!l\in Y(x, y\in l))$.

$\mathbf{A}_2$. Для любой прямой l и любой точки x существует единственная прямая l_1, проходящая через x и совпадающая или не пересекающаяся с l (прямая l_1 называется параллельной прямой l), т. е. $\exists!l_1\in Y(x\in l_1\wedge((l=l_1)\vee\neg\exists y\in X(y\in l\wedge y\in l_1)))$.

$\mathbf{A}_3$. Существуют три точки, не лежащие на одной прямой (не коллинеарные), т. е. $\exists x, y, z\in X\forall l\in Y\neg(x, y, z\in l)$.

Например, $\boldsymbol{R}^2$ или $\mathscr{X}_K$, где $\mathscr{X}_K$ — векторное пространство над полем (или телом) K, и, в частности, K^2 — классические аффинные плоскости. В них точками являются элементы этих множеств, а прямыми — подмножества вида $\{\lambda\cdot\xi|\lambda\in K\}$, где ξ — произвольный фиксированный элемент из $\mathscr{X}$, $\xi\neq 0$. Однако существуют классические аффинные плоскости, противоречащие нашему интуитивному представлению о плоскости.

Например, пусть X состоит из четырех точек, Y — из шести прямых, показанных на рисунке 30.

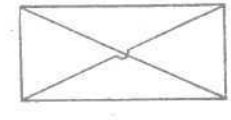


Рис. 30

Каждая из этих прямых по определению состоит ровно из двух точек.

Упражнение 16. а) *Проверьте выполнение аксиом* A_1—A_3 *для приведенной выше структуры, состоящей из четырех точек и шести прямых (аффинная плоскость, у которой носитель* X *конечное множество, называется конечной аффинной плоскостью).*

б) *Проверьте выполнение аксиом* A_1—A_3 *для аффинной плоскости* $\langle A, \mathscr{X}, g, K\rangle$, *где* K *— любое тело (см. § 2 этой главы).*

Существование конечных аффинных плоскостей, с одной стороны, является недостатком приведенной аксиоматики: хорошо, если аксиоматика имеет только такие модели, которые мы подразумеваем, когда пишем ее аксиомы. С другой стороны, получая следствия аксиом A_1—A_3, мы тем самым изучаем не только привычные аффинные плоскости вида $\mathscr{X}_K$, но и конечные аффинные плоскости, расширяя область применений геометрии. С аналогичной ситуацией мы сталкиваемся при изучении аксиоматики Пеано.

Классической проективной плоскостью называется пара множеств $\langle X, Y\rangle$, где $Y \subseteq \mathscr{P}(X)$, которая удовлетворяет следующим аксиомам (в них элементы множества X называются *проективными точками*, а элементы множества Y — *проективными прямыми*):

Π_1. Через любые две различные проективные точки можно провести единственную проективную прямую, т. е. $\Pi_1 \rightleftarrows A_1$.

Π_2. Любые две проективные прямые имеют общую проективную точку, т. е. $\forall l_1, l_2 \in Y\ \exists x \in X (x \in l_1 \wedge x \in l_2)$.

Π_3. Существуют три проективные точки, не лежащие на проективной прямой, т. е. $\Pi_3 \rightleftarrows A_3$.

Π_4. Каждая проективная прямая содержит хотя бы три различные проективные точки, т. е. $\forall l \in Y\ \exists x, y, z \in X (x, y, z \in l \wedge x \neq y \neq z \wedge x \neq z)$.

Например, P_2 или $P_2(K, \mathscr{X})$, где $\mathscr{X}$ — трехмерное векторное пространство над полем или телом K, в частности $P_2(K^3)$ — классические проективные плоскости. Как и в случае аксиоматики аффинной плоскости, эта аксиоматика имеет и неожиданные модели. Например, определим структуру $\langle X, Y\rangle$, в которой X состоит из семи проективных точек и Y — из семи проективных прямых, каждая из которых по определению является конечным множеством проективных точек; чтобы указать эти множества, применим к конечной аффинной плоскости, показанной на рисунке 30, следующий общий прием. Добавим к точкам аффинной плоскости в качестве новых точек связки параллельных прямых, т. е. семейства попарно параллельных прямых. В рассматриваемом случае таких связок три.

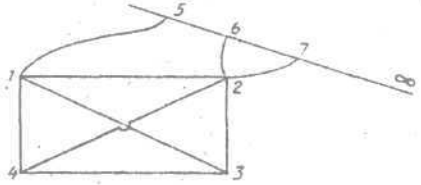


Рис. 31

Поэтому в новой модели будет семь проективных точек. Эти связки называются бесконечно удаленными проективными точками. К прямым, имеющимся в аффинной плоскости, добавим в качестве новой, *бесконечно удаленной* прямой множество, состоящее

из всех бесконечно удаленных точек. Поэтому в новой модели будет семь проективных прямых (рис. 31).

Упражнение 17. *а) Проверьте выполнение аксиом* Π_1 — Π_4 *для приведенной выше структуры из семи проективных точек и семи проективных прямых.*

б) Проверьте выполнение аксиом Π_1 — Π_4 *для проективной плоскости* $P_2(K^3)$, *где* K — *произвольное тело.*

Классическая проективная плоскость, у которой носитель X — конечное множество, называется конечной проективной плоскостью.

Рассмотрим семейство всех классических проективных плоскостей. Это семейство содержит не только все проективные плоскости вида $P_2(K^3)$, где K — любое тело, но и конечные проективные плоскости. Оказывается, что семейство «настоящих» проективных плоскостей, т. е. проективных плоскостей вида $P_2(K^3)$, где K какое-то тело, можно выделить из семейства всех классических проективных плоскостей добавлением всего одной дополнительной аксиомы, называемой аксиомой Дезарга.

Π_5. Пусть p_1, p_2, p_3, q_1, q_2, q_3 — шесть попарно различных точек, и прямые (p_i, q_i) пересекаются в какой-то одной точке a, отличной от каждой из шести точек p_i, q_i, $i = 1, 2, 3$ (см. рис. 32). Тогда три точки S_k, определяемые как пересечения прямых (p_i, p_j) и (q_i, q_j) лежат на какой-то одной прямой, где $\{i, j, k\}$ — всевозможные перестановки трех индексов 1, 2, 3.

Точка a, в которой пересекаются эти три прямые, называется перспективой. Точки p_1, p_2, p_3 и q_1, q_2, q_3 удобно представлять как вершины двух треугольников (они заштрихованы на рис. 32).

Поэтому аксиому Дезарга можно переформулировать: если треугольники (p_1, p_2, p_3) и (q_1, q_2, q_3) перспективны, т. е. имеют точку перспективы, то точки пересечения соответствующих сторон этих треугольников лежат на одной прямой.

Теорема 4. ***Пусть*** $\langle X, Y\rangle$ ***классическая проективная плоскость. Множество*** X ***— биективно (биекцию обозначим*** f***) с*** $P_2(K)$ ***(*** K ***— некоторое тело) так, что*** $\{f(l) \mid l \in Y\}$ ***совпадает с множеством всех проективных прямых в*** $P_2(K)$ ***в том и только в том случае, если в*** $\langle X, Y\rangle$ ***выполняется аксиома Дезарга.***

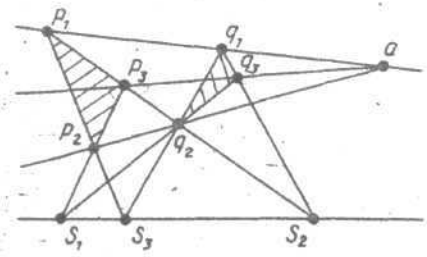


Рис. 32

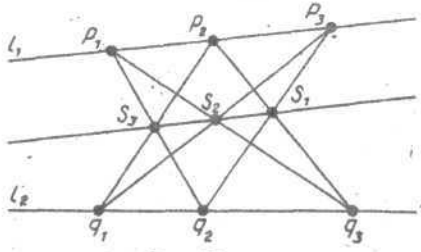


Рис. 33

Любую классическую проективную плоскость, удовлетворяющую аксиоме Дезарга, называют дезарговой проективной плоскостью.

Среди всех дезарговых проективных плоскостей естественно выделить семейство проективных плоскостей вида $P_2(K^3)$, где K — поле. Оказывается, что это семейство можно выделить добавлением к аксиоме Дезарга еще одной, новой аксиомы, называемой аксиомой Паппа.

Π_6. Пусть две различные проективные прямые l_1 и l_2 содержат по две тройки попарно различных точек p_1, p_2, $p_3 \in l_1$ и q_1, q_2, $q_3 \in l_2$. Три точки S_k, определяемые как пересечения прямых (p_i, q_j) и (p_j, q_i), лежат на одной прямой, где $\{i, j, k\}$ — всевозможные перестановки индексов 1, 2, 3 (см. рис. 33).

Теорема 5. ***Пусть $\langle X, Y\rangle$ классическая проективная плоскость. Множество X — биективно (биекцию обозначим f) с $P_2(K)$ (K — некоторое поле) так, что $\{f(l) \mid l \in Y\}$ совпадает с множеством всех проективных прямых в $P_2(K^3)$ в том и только в том случае, если в $\langle X, Y\rangle$ выполняется аксиома Паппа.***

Отметим, что из аксиом $\Pi_1 - \Pi_4$, Π_6 следует аксиома Π_5, т. е. всякая классическая проективная плоскость, удовлетворяющая аксиоме Паппа, удовлетворяет аксиоме Дезарга. Однако из аксиом $\Pi_1 - \Pi_5$ нельзя вывести аксиому Π_6, так как существуют дезарговы проективные плоскости, в которых аксиома Паппа не верна.

Примером такой плоскости является любая проективная плоскость вида $P_2(K^3)$, где K — тело, но не поле. Например, K — тело кватернионов.

В аксиоматиках $A_1 - A_3$ и $\Pi_1 - \Pi_4$ с самого начала предполагается, что объекты X и Y, составляющие пару $\langle X, Y\rangle$, являются множествами. Поэтому приведенные аксиоматики являются теоретико-множественными. Это обстоятельство принципиально отличает их от аксиоматик типа Евклида — Гильберта.

3. Аксиоматики плоскости Евклида — Гильберта, Лобачевского и Римана. Эти аксиоматики не предполагают заранее, как в случае теоретико-множественных аксиоматик, что прямая — это множество точек.

Поэтому нельзя понимать предикат «точка лежит на прямой» как предикат «точка принадлежит (в смысле теории множеств) прямой, являющейся множеством точек». Аналогично нельзя понимать предикат «прямые l_1 и l_2 параллельны» как предикат «множества l_1 и l_2 совпадают или не пересекаются» и т. д.

Аксиоматика Евклида — Гильберта состоит из пяти групп аксиом, включающих, в частности, аксиомы $A_1 - A_3$, в которых теперь не предполагается заранее, что «прямые» являются какими-то множествами. Напомним эту аксиоматику.

I. Аксиомы, касающиеся двухместного предиката «лежит» («проходит», «содержит»). Этот предикат обозначают $x \in l$, где x — точка и l — прямая, и читают: «Точка x лежит на прямой l».

1. Через любые две различные (не равные) точки проходит

ровно одна прямая (можно провести ровно одну прямую), т. е. $I.1 \rightleftarrows A_1$.

2. Всякая прямая содержит хотя бы две различные (не равные) точки.

3. Существуют три точки, не лежащие на одной прямой, т. е. $I.3 \rightleftarrows A_3$.

В этих и следующих аксиомах используется еще один двухместный предикат $x = y$, который читается: «Точки x и y совпадают (являются одной и той же точкой)». Тогда выражение «различные точки» записывается в виде $x \neq y$.

II. Аксиомы, касающиеся трехместного предиката «между». Его обозначают $[x, y, z]$, где x, y, z — точки, и читают «x, y, z — три различные точки, лежащие на одной прямой, и точка y лежит между точками x и z».

1. Если точка x лежит между двумя точками y и z, то эти три точки x, y, z попарно различны, лежат на одной прямой и x лежит между z и y.

2. Среди трех точек прямой существует не более одной точки, лежащей между двумя другими.

3. Если две различные точки x и y лежат на одной прямой, то существует лежащая на этой прямой точка z, для которой y лежит между x и z.

4. Пусть x, y, z три точки, не лежащие на одной прямой, и l прямая, не содержащая этих точек. Если l проходит через внутреннюю точку отрезка $[x, y]$, то она проходит или через внутреннюю точку отрезка $[x, z]$, или через внутреннюю точку отрезка $[y, z]$.

Из первых трех аксиом этой группы легко вытекает (проверьте), что для любых двух различных точек x и y на прямой, проходящей через них, между точками x и y лежит какая-то точка z, отличная от x и y. Назовем отрезком и обозначим $[x, y]$ множество всех точек z, лежащих между x и y. В четвертой аксиоме внутренней точкой отрезка $[x, y]$ называется такая его точка z, которая не равна x или y; такая точка существует. Впрочем, появление множеств в четвертой аксиоме не по существу: ее легко сформулировать без использования отрезков. А именно,

$$\forall x, y, z(\mid \exists l(x, y, z \in l) \Rightarrow (\exists u([x, u, y] \wedge u \neq x \wedge u \neq y)) \Rightarrow$$
$$\Rightarrow \exists v(([x, v, z \wedge v \neq x \wedge v \neq y) \vee ([y, v, z] \wedge v \neq y \wedge v \neq z))).$$

III. Аксиомы, касающиеся двух предикатов «конгруэнтности». Первый из них обозначают $[x, y] \sim [z, u]$ и читают «$x \neq y$ и $z \neq u$ и отрезок $[x, y]$ конгруэнтен отрезку $[z, u]$». Второй из них обозначают $\angle xyz \sim \angle x'y'z'$ и читают «x, y, z — три точки, не лежащие на одной прямой, и x', y', z' — три точки, не лежащие на одной прямой, и угол $\angle xyz$ конгруэнтен углу $\angle x'y'z'$». Первый из этих предикатов четырехместен, а второй — шестиместен.

1. Для любой прямой и любой точки на ней и любого отрезка существуют ровно два отрезка, лежащих на этой прямой по разные

стороны от этой точки и конгруэнтных данному отрезку. Запишем эту аксиому точнее:

$$\forall l \forall x \forall y,\ z(x \in l \wedge y \neq z \Rightarrow \forall u,\ v \in l([u,\ x,\ v] \mapsto \exists! r \exists! s$$

$$(r,\ s \in l \wedge ([u,\ r,\ x] \vee [r,\ u,\ x]) \wedge [r,\ x] \sim [y,\ z] \wedge$$

$$\wedge([x,\ s,\ v] \vee [x,\ v,\ s]) \wedge [x,\ s] \sim [y,\ z]).$$

2. Если два отрезка конгруэнтны какому-то третьему отрезку, то они конгруэнтны между собой.

3. Если $[x, y]$ и $[y, z]$ два отрезка на одной прямой, не имеющие общих внутренних точек, и $[x', y']$ и $[y', z']$ — два отрезка также на одной прямой, не имеющие общих внутренних точек, и пары отрезков $[x, y]$ и $[x', y']$, $[y, z]$ и $[y', z']$ конгруэнтны, то отрезки $[x, z]$ и $(x', z']$ конгруэнтны.

4. Для любого луча и любого угла (пары лучей, имеющих общую вершину) и любой стороны плоскости, определяемой первым лучом, существует ровно один луч, расположенный по данную сторону от первого луча, и такой, что он вместе с первым лучом образует угол, конгруэнтный данному углу (см. рис. 34). Любой угол конгруэнтен себе.

5. Пусть x, y, z три точки, не лежащие на одной прямой и x', y', z' также три точки, не лежащие на одной прямой. Если $[x, y] \sim [x', y']$, $[x, z] \sim [x', z']$ и $\angle yxz \sim \angle y'x'z'$, то $\angle xyz \sim \sim \angle x'y'z'$ и $\angle xzy \sim \angle x'z'y'$ (рис. 35).

В четвертой аксиоме «сторона плоскости, определяемая данным лучом» задается точкой, не лежащей на прямой, определяемой этим лучом; например, точкой z, показанной на рисунке 34.

В пятой аксиоме тройки точек x, y, z и x', y', z' можно назвать треугольниками. Тогда эта аксиома выражает один из признаков равенства треугольников.

Запишите все эти аксиомы аналогично аксиоме III_1 и убедитесь, что они не используют других предикатов, кроме тех, которые названы выше.

IV. Аксиома параллельности.

Для всякой прямой и всякой точки, не лежащей на ней, существует единственная прямая, проходящая через эту точку и не пересекающаяся с исходной прямой, т. е. $IV \Rightarrow A_2$.

Кроме этих четырех групп аксиом, в аксиоматику Евклида — Гильберта часто включают еще одну группу аксиом.

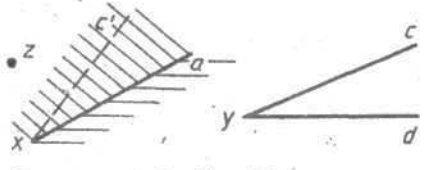


Рис. 34

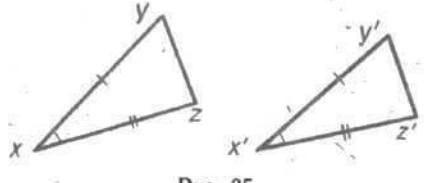


Рис. 35

V_1. *Аксиома Архимеда.* Для любой прямой и любого отрезка и любых двух точек x и y, лежащих на этой прямой, существуют натуральное число n и последовательность длины n из точек этой прямой, для которой $[x, x_1, x_2]$, $[x_1, x_2, x_3]$ и т. д., и все отрезки $[x, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$, ..., $[x_{n-1}, x_n]$ конгруэнтны исходному отрезку, и точка y находится между точками x и x_n.

V_2. *Аксиома Кантора.* Для любой последовательности вложенных друг в друга отрезков существует точка, которая им всем принадлежит.

Понятия «луч», «не пересекаться», «быть по разные стороны», «быть вложенными» и т. п., использованные в формулировках аксиом, опишите самостоятельно.

Эти две аксиомы называют аксиомами второго порядка, так как в них фигурируют новые объекты (иными словами, новые переменные): натуральные числа и последовательности, а также новый двуместный предикат «последовательность длины n».

Рассмотрим язык Γ_1, содержащий предикатные знаки вида $\cdot \in \cdot$; $[\cdot, \cdot, \cdot]$; $[\cdot, \cdot] \sim [\cdot, \cdot]$; $\angle \ldots \sim \angle \ldots$; $(\cdot, \cdot) \| (\cdot, \cdot)$, соответствующей арности (их аргументные места показаны точками), а также содержащий переменные первого сорта x, y, z, u, v, s, ... и второго сорта l, a, ... и два вида кванторов $\exists x$; $\forall x$; $\exists l$; $\forall l$, и, наконец, содержащий знаки $\wedge$; $\vee$; $\neg$; $\Rightarrow$; $\Leftrightarrow$ и скобки (;). Отметим, что все предикатные знаки кроме одного $\in$ на аргументных местах могут иметь переменные только первого сорта. Предикатный знак $\cdot_1 \in \cdot_2$ на аргументном месте $\cdot_1$ имеет переменную первого сорта, а на аргументном месте $\cdot_2$ переменную второго сорта. Формулой в языке Γ_1 называется слово, полученное в результате применения следующего индуктивного правила:

1) Формулой называется любое из следующих слов: $x \in l$; $[x, y, z]$; $[x, y] \sim [z, u]$; $\angle xyz \sim \angle x^1y^1z^1$,
где переменные x, l, y, z, u, x^1, y^1, z^1 могут быть заменены на любые другие переменные того же сорта;

2) Если φ и ψ формулы, то формулой называется любое из следующих слов: $(\varphi) \wedge (\psi)$, $(\varphi) \vee (\psi)$, $(\neg \varphi)$, $(\varphi) \Rightarrow (\psi)$, $(\varphi) \Leftrightarrow (\psi)$, $\exists x(\varphi) \forall x(\varphi)$, $\exists l(\varphi)$, $\forall x(\varphi)$, где переменные x и l могут быть заменены на любые переменные того же сорта.

Упражнение 18. *Запишите формулами языка Γ_1 все аксиомы групп I—IV.*

Получившийся список аксиом (аксиоматическая система) называется элементарной планиметрией.

Аксиоматической теорией Евклида — Гильберта (в языке Γ_1) называется семейство всех формул (в языке Γ_1), которые могут быть получены (вытекают, выводятся) из аксиом групп I—IV и перечисленных ниже логических аксиом по следующим правилам вывода: $\dfrac{A;\ A \Rightarrow B}{B}$ *(modus ponens)* и $\dfrac{A}{\forall \xi A}$ *(правило обобщения)*. Здесь и в ниже следующих логических аксиомах A, B, C — любые формулы нашего языка и ξ, η — любые переменные одного и того же вида. Сами эти правила понимаются традиционно: если формулы, перечисленные над чертой, выводимы,

то формула, написанная под чертой, по определению считается выводимой. Напомним обычный список логических аксиом:

1) $A \Rightarrow (B \Rightarrow A)$;
2) $(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$;
3) $A \Rightarrow (B \Rightarrow A \wedge B)$;
4) $A \wedge B \Rightarrow A$;
5) $A \wedge B \Rightarrow B$;
6) $(A \Rightarrow C) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \vee B \Rightarrow C))$;
7) $A \Rightarrow A \vee B$;
8) $B \Rightarrow A \vee B$;
9) $(A \Rightarrow B) \Rightarrow ((A \Rightarrow \neg B) \Rightarrow \neg A)$;
10) $\neg \neg A \Rightarrow A$;
11) $\forall \xi A \Rightarrow A(\varphi)$;
12) $\forall \xi (C \Rightarrow A) \Rightarrow (C \Rightarrow \forall \xi A)$, где C не содержит ξ;
13) $A(\varphi) \Rightarrow \exists \xi A$;
14) $\forall \xi (A \Rightarrow C) \Rightarrow (\exists \xi A \Rightarrow C)$, где C не содержит ξ.

Ясно, что аксиомы второго порядка не могут быть записаны в языке Γ_1. Для этого нужно, по крайней мере, два новых сорта переменных: один — переменные n, m, k, ..., которые пробегают натуральные числа, другой — переменные α, β, γ, ..., которые пробегают последовательности точек. Кроме того, нужны соответствующие новым переменным кванторы $\exists n$, $\forall n$, $\exists \alpha$, $\forall \alpha$ и два функциональных знака: двуместный функциональный знак $\alpha(n)$ и одноместный функциональный знак n' — и, наконец, еще новый двуместный предикат $n \leqslant m$. Первый из функциональных знаков читают «α — последовательность точек, и n — натуральное число; $\alpha(n)$ — точка, являющаяся значением последовательности α на натуральном числе n». Второй функциональный знак читают: «n — натуральное число, и n' — непосредственно следующее за ним число, т. е. $n' = n + 1$». Предикат $n \leqslant m$ читают: «n и m — натуральные числа, и число n меньше или равно числу m». Читателю предлагается самостоятельно определить соответствующий язык. Обозначим его Γ_2. Запишем на этом языке аксиомы пятой группы:

V. 1. $\forall l \forall z, u \forall x, y (z \neq u \wedge x, y \in l) \Rightarrow \exists n \exists \alpha (\forall m \leqslant n (\alpha(m) \in l) \wedge \alpha(1) = x \wedge \wedge \forall m < n ([\alpha(m), \alpha(m')] \sim [z, u]) \wedge [u, y, \alpha(n)])$.

V. 2. $\forall \alpha \forall \beta (\forall n \forall x ([\alpha(n'), x, \beta(n')] \Rightarrow [\alpha(n), x, \beta(n)]) \Rightarrow \exists y, \forall n ([\alpha(n), y, \beta(n)]))$.

Запись $m < n$ означает $(m \leqslant n) \wedge (m \neq n)$.

Конечно, в первой из этих аксиом можно считать, что α — произвольная конечная последовательность точек. Мы не делаем этого, чтобы не вводить еще одного сорта переменных «конечные последовательности точек».

В языке Γ_2 определяется соответствующая аксиоматическая теория Евклида — Гильберта точно так же, как и в случае языка Γ_1. Она обозначается Γ_2. Сформулируйте ее аналогично определению теории Γ_1.

Аксиоматика Лобачевского по определению состоит из всех аксиом групп I—V за единственным исключением: аксиома IV заменена на аксиому

IVа. Для любой прямой и любой точки, не лежащей на ней, существуют две различные прямые, проходящие через эту точку и не пересекающиеся с данной прямой.

Аксиоматика Римана по определению состоит из всех аксиом групп I, III—V, за исключением того, что аксиома IV заменена на аксиому

IVб. Любые две прямые пересекаются.

Обратите внимание, что в аксиомах Евклида — Гильберта и в аксиомах IVа и IVб смысл знака $\in$ отличается от смысла внешне такого же знака в аксиомах предыдущего пункта и в аксиомах теории множеств.

Множество всех формул в языке Γ_1, которые выводимы из аксиоматики Лобачевского (из аксиоматики Римана), называется аксиоматической теорией геометрии Лобачевского (соответственно аксиоматической теорией геометрии Римана) в языке Γ_1. Аналогично определяются аксиоматические теории Лобачевского и Римана в языке Γ_2.

Как и в случае любой аксиоматической системы, возникает первостепенный вопрос: существуют ли модели этих трех аксиоматик? Опишем в общих чертах, как строятся все обычные модели этих аксиоматик. В следующем пункте определяется специальный вид метрических пространств $\langle X, \rho \rangle$ (которые называются двумерными римановыми многообразиями). Элементы множества X называются точками. За счет особого вида этих пространств определяются такие их подмножества, которые называются прямыми. Семейство всех прямых в X обозначим Y. В произвольном метрическом пространстве нет оснований назвать некоторые его подмножества прямыми, с тем чтобы для этих прямых выполнялась большая часть упомянутых выше аксиом. Таким образом в двумерном римановом многообразии определяются точки и прямые. Предикат «лежит» означает, что точка x является элементом прямой l как множества; «между» $[x, y, z]$ означает, что $\exists l (x, y, z \in l)$ и выполняется $x \neq y \neq z$, $\rho(x, z) = \rho(x, y) + \rho(y, z)$. Движением, как всегда, называется биекция $f : X \leftrightarrow X$, для которой $\forall x, y \in X (\rho(x, y) = \rho(f(x), f(y))$. Любые два множества w_1 и w_2, содержащиеся в X, называются конгруэнтными, если существует движение f, для которого $f(w_1) = w_2$. В частности, таким образом определяется конгруэнтность отрезков и углов. (Угол $\angle yxz$ определяется как объединение лучей $[x, y)$ и $[x, z)$. Луч, как и отрезок, определяется с помощью понятия «между».) Натуральные числа и последовательности точек из пятой группы аксиом определяются как обычно. Таким образом, все исходные знаки и, следовательно, все формулы языка Γ_2 интерпретируются (получают явное определение) в римановом многообразии $\langle X, \rho \rangle$. Если при этом все аксиомы данной аксиоматики «истинны», то риманово многообразие (метрическое пространство) $\langle X, \rho \rangle$ называется соответственно моделью аксиоматик Евклида — Гильберта, Лобачевского или Римана.

Итак, мы указали интерпретацию всех формул языка Γ_2 с помощью метрического пространства $\langle X, \rho\rangle$ и некоторого (пока еще не определенного) семейства Y подмножеств множества X. Набор $\langle X, Y, \rho\rangle$ назовем евклидовой системой.

Подчеркнем, что наравне с понятием прямой осталось не определенным понятие «истинно». Последнее понятие не является чем-то самоочевидным и наряду со всеми другими не определяемыми, исходными понятиями этих аксиоматик требует интерпретации (явного определения). Приведем одно из возможных определений двуместного предиката «истинно», обозначая его $\cdot_1 \models \cdot_2$. На первое аргументное место подставляются произвольные евклидовы системы $M \rightleftarrows \langle X, Y, \rho\rangle$, а на второе аргументное место подставляются объекты, которые получаются из формул φ с переменными $x_1, \ldots, x_n$ подстановкой вместо переменных первого сорта «точек» из X и вместо переменных второго сорта «прямых» из Y. Здесь переменные $x_1, \ldots, x_n$ — это все свободные переменные, входящие в формулу φ; последнее записывается в виде $\varphi(x_1, \ldots, x_n)$. (Мы считаем известным понятие свободной переменной: переменной x, не входящей в φ в виде $\exists x(\ldots)$, $\forall x(\ldots)$.) После такой подстановки получается объект $\varphi(x_1, \ldots, x_n)$, где $x_i \in X$ или $x_i \in Y$ в зависимости от сорта переменной x_i. Этот объект, конечно, не удовлетворяет определению формулы, но, допуская вольность речи, будем называть его также формулой. Данное определение $\varphi(x_1, \ldots, x_n)$ буквально годится только для языка Γ_1, так как в языке Γ_2 имеются еще два сорта переменных; но точно так же оно переносится на этот более широкий случай (уточните).

Определим предикат $\langle X, Y, \rho\rangle \models \varphi(\bar{x}_1, \ldots, \bar{x}_n)$ индукцией по длине формулы φ.

Пусть $\varphi \rightleftarrows (x \in l)$. По определению выполняется $M \models \varphi$, если x, как элемент, принадлежит множеству l в смысле теории множеств. Пусть $\varphi \rightleftarrows [x_1, x_2, x_3]$; по определению он выполняется, если $\rho(x_1, x_3) = \rho(x_1, x_2) + \rho(x_2, x_3)$. Пусть $\varphi \rightleftarrows (w_1 \sim w_2)$, где w_1 и w_2 отрезки или углы; по определению выполняется предикат $M \models \varphi$, если существует такое движение $f : X \leftrightarrow X$, что $f(w_1) = w_2$.

Если $\varphi \rightleftarrows (\varphi_1 \wedge \varphi_2)$, то $(M \models \varphi) \rightleftarrows (M \models \varphi_1) \wedge (M \models \varphi_2)$. Если $\varphi \rightleftarrows (\varphi_1 \vee \varphi_2)$, то $(M \models \varphi) \rightleftarrows (M \models \varphi_1) \vee (M \models \varphi_2)$. Если $\varphi \rightleftarrows |\varphi_1$, то $(M \models \varphi) \rightleftarrows |(M \models \varphi_1)$. Случаи $M \models (\varphi_1 \Rightarrow \varphi_2)$ и $M \models (\varphi_1 \Leftrightarrow \varphi_2)$ определяются аналогично.

Если $\varphi \rightleftarrows (\exists x \varphi_1(x))$, то $(M \models \varphi) \rightleftarrows \exists x \in X (M \models \varphi_1(x))$.

Если $\varphi \rightleftarrows (\forall x \varphi_1(x))$, то $(M \models \varphi) \rightleftarrows \forall x \in X (M \models \varphi_1(x))$.

Случаи $M \models \exists l \varphi_1(l)$ и $M \models \forall l \varphi_1(l)$ определяются аналогично.

В аксиоматику Римана не входит предикат «между», в этом случае он не участвует в формулах и в определении предиката $M \models \varphi$.

Предикат $M \models \varphi$ определяется в рамках теории множеств и, следовательно, его выполнение или невыполнение зависит от того, что истинно и что ложно в этой теории. Таким образом теория множеств оказалась необходимой и для евклидова аксиоматического подхода в геометрии. Правда, в отличие от аксиоматики из п. 2 здесь теория множеств возникает на другом уровне: в связи с определением предиката истинности. Итак, с учетом определения предиката «истинно» мы приходим к следующему понятию.

Теоретико-множественной моделью аксиоматики Евклида — Гильберта (соответственно аксиоматики Лобачевского и Римана) называется евклидова система вида $M = \langle X, Y, \rho\rangle$, в которой все аксиомы Евклида — Гильберта (соответственно двух других аксиоматик) истинны.

Теоретико-множественную модель любой из этих аксиоматик называют еще аксиоматически определенной плоскостью соответственно Евклида — Гильберта, Лобачевского или Римана.

Теорема 6. ***Евклидова система $M_0 \rightleftarrows \langle R^2, Y, \rho\rangle$, где Y семейство всех (обычных) прямых в R^2 и $\rho(\cdot, \cdot)$ — обычная метрика в R^2, является теоретико-множественной моделью аксиоматики Евклида — Гильберта и, следовательно, аксиоматически определенная плоскость Евклида — Гильберта существует.***

▷ Доказательство этой теоремы состоит в непосредственной и очевидной проверке всех аксиом. Она оставляется читателю в качестве упражнения. ■

Следствие. Все формулы, выводимые в аксиоматической теории Евклида — Гильберта, истинны в M_0.

▷ В теореме 6 говорится об истинности в M_0 самих аксиом. Легко проверяется, что все логические аксиомы истинны в M_0, и если формулы над чертой в двух правилах вывода истинны в M_0, то формулы под чертой также истинны в M_0. Эти утверждения непосредственно вытекают из определения предиката $M \models \varphi$ и не зависят от конкретного вида формул и системы M_0. Отсюда индукцией по длине вывода получается следствие. ■

В конце книги в приложении 4 строятся теоретико-множественные модели аксиоматик Лобачевского и Римана (ими оказываются верхняя половина двуполостного гиперболоида и проективная плоскость) и, следовательно, существуют аксиоматически определенные плоскости Лобачевского и Римана.

Предыдущее следствие, как видно из его доказательства, универсально: если рассматривается любая аксиоматическая теория Λ и M структура, в которой истинны все аксиомы теории Λ (т. е. M модель Λ), то в M истинны и все формулы, выводимые в теории Λ. Обратное утверждение почти всегда не верно. Обратное утверждение имеет вид: если Λ — некоторая аксиоматическая теория и M — структура, в которой истинны все аксиомы теории Λ (т. е. M модель Λ), то всякая формула, истинная в M, выводима в теории Λ. В тех редких случаях, когда для аксиоматической теории Λ существует хотя бы одна модель M, для которой выполняется последнее утверждение, теория Λ называется полной. В противном случае теория Λ называется неполной. Итак, почти все аксиоматические теории не полны. Но аксиоматическая теория Евклида — Гильберта (в языке Γ_1) относится к числу редких исключений: она полна.

Теорема 7.

а) *Элементарная планиметрия, т. е. аксиоматическая теория Евклида — Гильберта (в языке Γ_1) полна.*

б) *Существует такой алгоритм $\mathfrak{A}$, определенный на множестве всех формул языка Γ_1, что*

$$\mathfrak{A}(\varphi)=\begin{cases}1, & \text{если } \varphi \text{ истинна в евклидовой системе } M_0,\\ 0, & \text{если } \varphi \text{ ложна в } M_0.\end{cases}$$

Аксиоматическая теория, для которой выполняется утверждение 8 б), называется разрешимой. Итак, элементарная планиметрия Γ_1 — полная и разрешимая теория. Оба эти утверждения не верны для аксиоматической теории Евклида — Гильберта в языке Γ_2. Поскольку обычная школьная планиметрия совпадает с аксиоматической теорией Евклида — Гильберта в языке Γ_1, то ЭВМ может ответить на любой вопрос, возникающий в ней. Для этого нужно запрограммировать на ЭВМ алгоритм $\mathfrak{A}$. Это на самом деле осуществлено. По-видимому, к такой ситуации применимы слова «машинная математика».

4. Двумерные римановы многообразия как модели аксиоматических определений плоскости. Кратко рассмотрим вопрос о том, как могут определяться структуры (системы), в которых интерпретируются (получают смысл) все исходные понятия, встречающиеся в аксиомах Евклида — Гильберта, Лобачевского и Римана, и, следовательно, интерпретируются сами аксиомы одноименных аксиоматических систем. При этом одни аксиомы могут оказаться истинными, а другие ложными. Среди этих систем найдутся и такие, в которых все аксиомы соответствующих аксиоматических систем истинны, они называются моделями этих аксиоматических систем.

Напомним, что мы рассматриваем только двумерный случай, т. е. случай планиметрии.

В качестве множества X в евклидовой системе выбирается двумерное многообразие (поверхность) X (см. с. 92). Типичными примерами таких множеств являются важные и во многих других отношениях поверхности: арифметическая плоскость $\boldsymbol{R}^2$, сфера S^2, верхняя половина двуполостного гиперболоида вращения $\begin{cases}x_1^2-x_2^2-x_3^2=1\\ x_1>0\end{cases}$, (которую мы обозначим ${}^+S_1^2$) и проективная плоскость P_2. Гиперболоид вращения обозначим S_1^2. Отметим, что поверхности S^2 и S_1^2 определяются аналогично, как множества вида $\{x\in\boldsymbol{R}^3\mid\rho(x,\,0)=1\}$, где в первом случае $\rho(x,\,0)=\sqrt{x_1^2+x_2^2+x_3^2}$, а во втором случае $\rho(x,\,0)=\sqrt{x_1^2-x_2^2-x_3^2}$. Очевидно, гиперболоид вращения S_1^2 состоит из двух связных частей, гомеоморфных между собой (рис. 36).

Знак $+$ в записи ${}^+S_1^2$ означает связную компоненту множества S_1^2, т. е. одну из двух его «одинаковых половинок». Не важно, какую именно из них выбрать, и мы произвольно выбираем верхнюю половинку.

Три из указанных поверхностей, а именно $\boldsymbol{R}^2$, S^2, S_1^2, являются подмножествами множества $\boldsymbol{R}^3$. Напротив, четвертая поверхность P_2 не является подмножеством $\boldsymbol{R}^3$ и не гомеоморфна такому подмножеству.

В $\boldsymbol{R}^3$, как и в любом векторном пространстве над полем $\boldsymbol{R}$, можно выбирать разные билинейные формы (см. § 2 п. 3). Если в векторном пространстве $\boldsymbol{R}^3$ выбирается форма

$$(x,\,y)=x_1y_1+x_2y_2+x_3y_3,$$

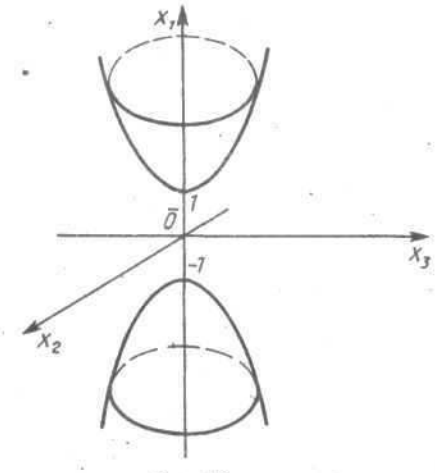


Рис. 36

то получившееся пространство с формой обозначают также $\boldsymbol{R}^3$. Если в том же векторном пространстве $\boldsymbol{R}^3$ выбирается другая форма

$$(x,\,y)_1=x_1y_1-x_2y_2-x_3y_3,$$

то получившееся пространство с формой обозначают $\boldsymbol{R}_1^3$.

В соответствии с общим определением метрики в пространстве с формой, в $\boldsymbol{R}^3$ определяется метрика $\rho(x,\,y)=\sqrt{(x-y,\;x-y)}$ и, в частности, $\rho(x,\,0)=\sqrt{x_1^2+x_2^2+x_3^2}$, а в $\boldsymbol{R}_1^3$ определяется метрика $\rho_1(x,\,y)=\sqrt{(x-y,\;x-y)_1}$ и, в частности,

$$\rho_1(x,\,0)=\sqrt{x_1^2-x_2^2-x_3^2}.$$

Итак, S_1^2 — сфера с центром в точке 0 радиуса 1 в пространстве $\boldsymbol{R}_1^3$. Точки $(1,\,0,\,0)$ и $(-1,\,0,\,0)$ называются соответственно северным и южным полюсами сфер S^2 и S_1^2.

Следующая часть этого пункта посвящена обсуждению важного понятия касательной плоскости к двумерному многообразию в одной из его точек.

Пусть X одна из трех поверхностей $\boldsymbol{R}^2$, S^2, ${}^+S_1^2$ и $a\in X$. Обычным образом определяется понятие касательной плоскости к поверхности X в точке a. Обозначим ее $T_a(X)$. В случае $X=\boldsymbol{R}^2$ касательная плоскость $T_a(\boldsymbol{R}^2)$ — это само $\boldsymbol{R}^2$, у которого начало координат перенесено в точку a. Элементами касательной плоскости $T_a(X)$ являются обычные векторы в $\boldsymbol{R}^3$ с началом в точке a; они называются касательными векторами в точке a к поверхности X.

Множество всех векторов в $\boldsymbol{R}^3$ с началом в точке a обозначим $T_a(\boldsymbol{R}^3)$. Поэтому $T_a(X)\subseteq T_a(\boldsymbol{R}^3)$ для любой точки $a\in X$. На рисунке

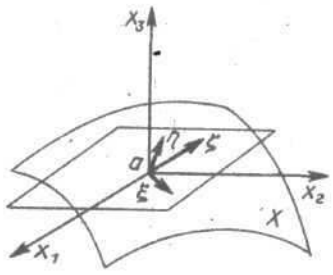


Рис. 37

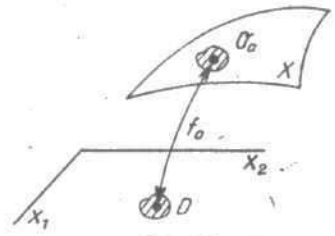


Рис. 39

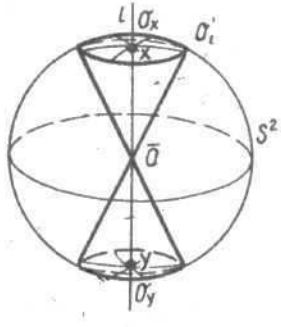


Рис. 38

37 отражено взаимоотношение объектов X, a, $T_a(X)$ и $T_a(\boldsymbol{R}^3)$, на нем векторы ξ, η, ζ принадлежат $T_a(X)$ и $T_a(\boldsymbol{R}^3)$.

Элементы касательной плоскости $T_a(X)$ складывают и умножают на вещественные числа обычным образом, поэтому $T_a(X)$ — двумерное векторное пространство над полем $\boldsymbol{R}$.

В любой из четырех перечисленных поверхностей X определена топология. Для первых трех из них топологии совпадают с топологиями, индуцированными из $\boldsymbol{R}^3$, а для последней — с топологией проективной плоскости (рис. 38). Для любой точки $a \in X$ можно указать ее окрестность $\mathcal{O}_a$, гомеоморфную открытому кругу $\mathring{D}$. На рисунке 39 $f_a : \mathring{D} \leftrightarrow \mathcal{O}_a$ соответствующий гомеоморфизм.

Упражнение 19. *Укажите такие $\mathcal{O}_a$, f_a, $\mathring{D}$ для перечисленных выше четырех поверхностей.*

Поскольку в X определена топология, то возникает семейство всех непрерывных функций $f : X \to \boldsymbol{R}$. Оно обозначается $C(X, \boldsymbol{R})$ и образует относительно поточечно определенных операций сложения, умножения, а также умножения на вещественные числа (бесконечномерную) алгебру над полем $\boldsymbol{R}$.

Каждая из трех первых поверхностей X определяется уравнением $f(x, y, z) = 0$ в $\boldsymbol{R}^3$ с бесконечно дифференцируемой левой частью, которое для любой фиксированной точки $a \in X$ имеет ненулевую производную $f'(a)$ от своей левой части f. Производная (градиент), как всегда, определяется равенством $f'(a) \rightleftarrows (f'_x(a), f'_y(a), f'_z(a))$. Пусть, например, $f'_z(a) \neq 0$. По теореме о неявной функции из курса анализа можно разрешить уравнение $f(x, y, z) = 0$ относительно переменной z в некоторой окрестности $\mathcal{O}_a$ точки a. Здесь $\mathcal{O}_a$ — открытое множество из топологии в X. Иными словами, найдется функция $z = g(x, y)$, определенная, например, в открытом

круге $\mathring{D}$, которая бесконечно дифференцируема и задает гомеоморфизм $G_a(x, y) \rightleftarrows (x, y, g(x, y))$ вида $G_a : \mathring{D} \leftrightarrow \mathcal{O}_a$. Бесконечно дифференцируемый гомеоморфизм $G_a : D(G_a) \leftrightarrow \mathcal{O}_a$, область определения которого $D(G_a)$ любое открытое связное множество в $\boldsymbol{R}^2$, называется картой на X в точке a. Следовательно, вся поверхность X покрывается областями значений некоторых карт.

Функцию $f \in C(X, \boldsymbol{R})$ назовем бесконечно дифференцируемой, если для любой точки $a \in X$ функция $(f \circ G_a) : D(G_a) \to \boldsymbol{R}$ бесконечно дифференцируема.

Чтобы получить пример бесконечно дифференцируемой функции на X, достаточно рассмотреть любую бесконечно дифференцируемую функцию $h : \boldsymbol{R}^3 \to \boldsymbol{R}$ от трех переменных и сузить ее на X. Тогда для любой точки $a \in X$ композиция $(h \circ G_a) : D(G_a) \to \boldsymbol{R}$ будет, очевидно, бесконечно дифференцируемой функцией.

Семейство всех бесконечно дифференцируемых функций на X обозначим $C^\infty(X, \boldsymbol{R})$; конечно $C^\infty(X, \boldsymbol{R}) \subseteq C(X, \boldsymbol{R})$. Ясно, что элементы семейства $C^\infty(X, \boldsymbol{R})$, как и семейства $C(X, \boldsymbol{R})$, можно поточечно складывать, перемножать, а также умножать на вещественные числа, т. е. оба эти семейства — (бесконечномерные) алгебры над полем $\boldsymbol{R}$.

Обозначим $B_a(X)$ семейство всех функций f вида $f : \mathcal{O}_f \to \boldsymbol{R}$, где $\mathcal{O}_f$ некоторая (зависящая от f) окрестность точки a топологии в X, для которых композиция $(f \circ G_a) : D(G_a) \to \boldsymbol{R}$ бесконечно дифференцируема. Две функции из $B_a(X)$ по определению не различаются, если они совпадают в какой-нибудь окрестности точки a. Ясно, что семейство $B_a(X)$ совершенно аналогично семейству B, используемому в определении 1г. Для любой функции f из $C^\infty(X, \boldsymbol{R})$ сужение f на любую окрестность точки a принадлежит $B_a(X)$. В частности, $C^\infty(X, \boldsymbol{R}) \subseteq B_a(X)$.

Точно, как в определении 1г, сформулируем понятие вектора на X в точке a. А именно, вектором на X в точке a назовем любое отображение (дифференцирование) вида $d(\cdot) : B_a(X) \to \boldsymbol{R}$, обладающее свойствами из определения 1 г (где точка 0 заменена на точку a).

Упражнение 20. *Докажите, что все векторы на X в точке a образуют векторную плоскость, изоморфную $T_a(X)$.*

Указание. Чтобы построить этот важный изоморфизм, нужно любую функцию f из $B_a(X)$, определенную, например, в окрестности $\mathcal{O}_a$, продолжить по нормалям к поверхности X в некоторую окрестность уже в топологии $\boldsymbol{R}^3$ точки a до функции f_1, а затем для фиксированного вектора ξ из $T_a(X)$ положить $d_\xi(f) \rightleftarrows d_\xi(f_1)$ — производной по направлению ξ от функции f_1. Тогда $d_\xi : B_a(X) \to \boldsymbol{R}$. Положим $\psi(\xi) \rightleftarrows d_\xi$; отображение ψ переводит $T_a(X)$ в множество всех дифференцирований на X в точке a. Это отображение ψ является искомой биекцией и изоморфизмом.

Другое описание этого изоморфизма также полезно: пусть I_a —

якобиан отображения G_a, т. е. матрица $I_a \rightleftarrows \begin{pmatrix} 1 & 0 & g'_x(a) \\ 0 & 1 & g'_y(a) \end{pmatrix}$. Эта матрица задает линейную биекцию между $T_b(\boldsymbol{R}^2)$ и $T_a(X)$ вида $\xi \mapsto I_a \cdot \xi$, где $G_a(b)=a$. Можно положить: $d_\xi(f) \rightleftarrows d_\eta(f \circ G_a)$, где $\eta \rightleftarrows \rightleftarrows I_a^{-1}(\xi)$, а d_η — дифференцирование по направлению η в точке b. ■

Спрашивается, зачем рассматривать дифференцирования и алгебру $B_a(X)$, если множество всех дифференцирований в точке a с помощью изоморфизма ψ отождествляется с гораздо более наглядным объектом $T_a(X)$? Дело в том, что для поверхности $X=P_2$ не ясно, что такое $T_a(P_2)$, а дифференцирования в точке $a \in P_2$ определяются для поверхности P_2 так же, как для трех предыдущих поверхностей. А именно, согласно первому определению проективной плоскости P_2 совпадает с фактор-множеством $(\boldsymbol{R}^3 \setminus \overline{|O|})/\sim$. Каждый класс эквивалентности имеет вид $\{\langle \lambda \cdot x, \lambda \cdot y, \lambda \cdot z \rangle \mid \lambda \in \boldsymbol{R}^*\}$, где $\langle x, y, z \rangle \in \boldsymbol{R}^3 \setminus \overline{|O|}$. Если хотя бы один представитель этого класса имеет ненулевой член на одном из своих трех мест, то и любой представитель этого класса на этом же месте имеет ненулевой член. Поэтому множество P_2 представимо в виде $\{[\langle x, y, z \rangle] \mid x \neq 0\} \cup \{[\langle x, y, z \rangle] \mid y \neq 0\} \cup \{[\langle x, y, z \rangle] \mid z \neq 0\}$. Эти три множества обозначим соответственно X_1, X_2, X_3; они попарно пересекаются.

Упражнение 21. *Докажите, что множества X_1, X_2, X_3 открыты в топологии плоскости P_2.*

Между множествами X_1 и $\boldsymbol{R}^2$ фиксируем гомеоморфизм $f_1 : X_1 \leftrightarrow \boldsymbol{R}^2$, определяемый равенством $f_1([\langle x, y, z \rangle]) \rightleftarrows \left(\frac{y}{x}, \frac{z}{x}\right) \in \boldsymbol{R}^2$. Аналогично $f_2([\langle x, y, z \rangle]) \rightleftarrows \left(\frac{x}{y}, \frac{z}{y}\right) \in \boldsymbol{R}^2$, $f_2 : X_2 \leftrightarrow \boldsymbol{R}^2$ и $f_3([\langle x, y, z \rangle]) \rightleftarrows \left(\frac{x}{z}, \frac{y}{z}\right) \in \boldsymbol{R}^2$, $f_3 : X_3 \leftrightarrow \boldsymbol{R}^2$.

Упражнение 22. *Проверьте, что функции f_1, f_2, f_3 действительно являются гомеоморфизмами.*

Если проективная точка $l=[\langle x, y, z \rangle]$ одновременно принадлежит X_1 и X_2, то $f_1(l)=\left(\frac{y}{x}, \frac{z}{x}\right)$, $f_2(l)=\left(\frac{x}{y}, \frac{z}{y}\right)$ и возникает отображение $f_{21} : \left(\frac{y}{x}, \frac{z}{x}\right) \mapsto \left(\frac{x}{y}, \frac{z}{y}\right)$ вида $f_{21} : \boldsymbol{R}^2 \to \boldsymbol{R}^2$, т. е. $f_{21}((u, v)) = \left(\frac{1}{u}, \frac{v}{u}\right)$, где $(u, v) \in \boldsymbol{R}^2$ и $u \neq 0$. Якобиан (производная) f'_{21} отображения f_{21} равен $\begin{pmatrix} -\frac{1}{u^2} & 0 \\ -\frac{v}{u^2} & \frac{1}{u} \end{pmatrix}$ и невырожден во всей области определения $D(f_{21})$. Аналогично определяются отображения $f_{12}, \ldots, f_{23}$.

Упражнение 23. *Докажите, что все отображения f_{21}, f_{31}, f_{23} гомеоморфизмы, у которых, как и у обратных к ним отображений, существуют частные производные их компонент любого*

порядка, а якобианы всюду невырождены. (Такие отображения называются диффеоморфизмами.)

Для любой точки $l \in P_2$ фиксируем в качестве ее окрестности одно из множеств X_1, X_2, X_3, содержащее l, и в качестве соответствующего этой окрестности гомеоморфизма фиксируем одно из отображений f_1^{-1}, f_2^{-1}, f_3^{-1} с тем же номером.

Так же, как выше (l в роли a, X_l в роли $\mathcal{O}_a$, f_l^{-1} в роли G_a), определяем алгебру $B_l(P_2)$ и ее дифференцирования в точке l. Полученное семейство дифференцирований (векторов) назовем касательной плоскостью к поверхности P_2 в точке l и обозначим $T_l(P_2)$.

Упражнение 24. *Пусть точка l принадлежит одновременно, например, X_1 и X_2 и d — дифференцирование в $B_l(P_2)$. Оно соответствует какому-то вектору ξ в смысле $d(h)=d_\xi(h \circ f_1^{-1})$ и вектору η в смысле $d(h)=d_\eta(h \circ f_2^{-1})$, где h пробегает $B_l(P_2)$, определяемое соответственно f_1^{-1} и f_2^{-1}. Докажите, что η получается из ξ умножением на якобиан f'_{21}.*

Если на двумерном многообразии X для каждой из его точек a заданы окрестность этой точки $\mathcal{O}_a$ и гомеоморфизм G_a и при этом отображения $G_{ba} : D(G_a) \to D(G_b)$, определенные как $G_b^{-1} \circ G_a$, являются диффеоморфизмами, то говорят, что X — двумерное гладкое многообразие. Не входя в детальное разъяснение этого понятия, ограничимся тем, что интересующие нас четыре поверхности, как было довольно подробно разобрано, являются гладкими многообразиями, т. е. для них определены все упомянутые объекты $\mathcal{O}_a$, G_a, G_{ba} и они обладают нужными свойствами. Главное для нас, что для такого X и каждой его точки a определена касательная плоскость $T_a(X)$.

Теперь перейдем к описанию формы в касательных плоскостях к интересующим нас поверхностям.

Если поверхность X содержится в $\boldsymbol{R}^3$ и в $\boldsymbol{R}^3$ фиксирована билинейная форма g, то можно сузить эту форму на X: если $\xi, \eta \in T_a(X)$, то, учитывая $T_a(X) \subseteq T_a(\boldsymbol{R}^3)$, положим $g_a(\xi, \eta) \rightleftarrows g(\xi, \eta)$. Хотя форма g одна и та же для всех точек $a \in \boldsymbol{R}^3$, но форма $g_a(\xi, \eta)$ может зависеть от точки $a \in \boldsymbol{R}^3$. Для любой фиксированной точки $a \in X$ форма $g_a(\xi, \eta) : (T_a(X))^2 \to \boldsymbol{R}$ очевидно, билинейная. Если X не содержится в $\boldsymbol{R}^3$, как, например, в случае $X=P_2$, то менее ясно, откуда может возникнуть такая билинейная форма $g_a(\xi, \eta)$ (зависящая кроме своих обычных аргументов ξ, η из $T_a(X)$ еще и от точки $a \in X$).

Пусть для двумерного гладкого многообразия X определена функция g от трех аргументов $a \in X$ и $\xi, \eta \in T_a(X)$, которая при фиксированной точке a является евклидовой формой на $T_a(X)$. Если при этом для любой карты $G_a : D(G_a) \to \mathcal{O}_a$ функция $g_1(u, \xi, \eta) \rightleftarrows \rightleftarrows g(G_a(u), d_1, d_2)$, где $d_1(h) \rightleftarrows d_\xi(h \circ G_a)$ и $d_2(h) \rightleftarrows d_\eta(h \circ G_a)$. $\forall h \in B_a(X)$, бесконечно дифференцируемая по своим аргументам u, ξ, η всюду в $D(G_a) \times \boldsymbol{R}^2 \times \boldsymbol{R}^2$, то пара $\langle X, g \rangle$ называется двумерным римановым многообразием.

В случае поверхностей $\boldsymbol{R}^2$, S^2 и ${}^+S_1^2$ форма в них по определению является сужением формы в содержащем их векторном пространстве $\boldsymbol{R}^3$. Точнее, в $\boldsymbol{R}^2$ — это обычное скалярное произведение (не зависящее от точки $a \in \boldsymbol{R}^2$) или, что то же самое, сужение в $\boldsymbol{R}^2$ обычного скалярного произведения в $\boldsymbol{R}^3$; в S^2 и ${}^+S_1^2$ — это сужения форм, определенных соответственно в $\boldsymbol{R}^3$ и $\boldsymbol{R}_1^3$.

Упражнение 25. *Проверьте, что сужение (не евклидовой) формы* $(\xi, \eta)_1$ *в* $\boldsymbol{R}_1^3$ *на* ${}^+S_1^2$ *является евклидовой формой.*

Итак, приведены три примера двумерных римановых многообразий. Четвертым примером будет проективная плоскость P_2. Нужно только определить форму в P_2: если два вектора (дифференцирования) d_1, d_2 принадлежат касательной плоскости $T_a(P_2)$, то a принадлежит хотя бы одному из множеств X_1, X_2, X_3. Пусть, например, $a \in X_1$. Тогда в касательном пространстве $T_{f_1(a)}(\boldsymbol{R}^2)$ найдутся векторы ξ и η (т. е. попросту найдутся векторы ξ и η в $\boldsymbol{R}^2$ с началом в точке a), для которых $d_1(h) = d_\xi(h \circ f_1^{-1})$ и $d_2(h) = d_\eta(h \circ f_1^{-1})$, $\forall h \in B_a(P_2)$. Векторы ξ и η однозначно определяются по d_1 и d_2 (почему?). Положим

$$g_a(d_1, d_2) \rightleftarrows (\xi, \eta),$$

где $(\cdot, \cdot)$ — обычное скалярное произведение в $\boldsymbol{R}^2$. Проверьте: если $a \in X_2$ и все то же самое проделать для $f_2(a)$ и соответствующих ξ_1 и η_1 в $T_{f_2(a)}(\boldsymbol{R}^2)$, то численное значение $g_a(d_1, d_2)$ не изменится. Далее X — одно из этих четырех многообразий.

Пусть $a, b \in X$ и α — произвольная инъективная непрерывная кривая, соединяющая точки a и b, т. е. $\alpha : [t_0, t_1] \to X$, где $t_0, t_1 \in \boldsymbol{R}$ и $\alpha(t_0) = a$, $\alpha(t_1) = b$, а также для любой точки $c \in R(\alpha)$ композиция $(G_c^{-1} \circ \alpha) : [t_0, t_1] \to \boldsymbol{R}^2$ является гладкой кривой в своей области определения. Так как X связно, то нетрудно доказать, что для любых точек a и b из X такое α существует, но далеко не единственно. Такое α называют путем (кривой) на множестве X из точки a в точку b.

В X следующим образом определяется метрика:

$$\rho(a, b) \rightleftarrows \inf \int_{t_0}^{t_2} \sqrt{g_{\alpha(t)}(\alpha'(t), \alpha'(t))} \cdot dt, \qquad (*)$$

где α любой путь на X из a в b. Эта точная нижняя грань, очевидно, существует и функция ρ действительно удовлетворяет трем аксиомам метрики. Более того, можно доказать, что эта точная нижняя грань достигается на некотором (для многих пар точек a, b — единственном) пути на многообразии X, соединяющем точки a и b. Такой путь (кривая) называется кратчайшей геодезической между точками a и b из X.

В случае $X = \boldsymbol{R}^2$ эта точная нижняя грань достигается на прямолинейном отрезке, соединяющем точки a и b, и метрика ρ совпадает с обычной метрикой в $\boldsymbol{R}^2$. В этом случае кратчайшими геодезическими являются отрезки. В случае $X = S^2$ эта точная нижняя грань достигается на меньшей дуге окружности большого круга, определяемого как пересечение сферы с плоскостью, проходящей через точки O и a, b; следовательно кратчайшими геодезическими являются эти дуги. Говоря в предыдущем случае об отрезке и в этом случае о меньшей дуге большого круга, мы подразумеваем их обычную параметризацию. Уточните ее и проверьте прямыми подсчетами эти утверждения.

В случае ${}^+S_1^2$ эта нижняя грань достигается на дуге, которая определяется как пересечение ${}^+S_1^2$ с плоскостью, проходящей через точки a, b и O (проверьте). Здесь также подразумевается обычная параметризация этой дуги.

В случае P_2 путь α, соединяющий проективные точки l' и l'' из P_2 определяет путь α_1 на сфере S^2. А именно, $\alpha_1(t) \rightleftarrows \alpha(t) \cap S^2$. Однако α_1 — не функция в обычном смысле этого слова, а «двузначная функция». Для любого t «значение» $\alpha_1(t)$ состоит из двух диаметрально противоположных точек.

При этом $R(\alpha_1)$ несвязное множество, которое состоит из двух компонент связности. Сопоставляя паре диаметрально-противоположных точек $\alpha_1(t)$ точку на сфере в одной (произвольно выбранной и фиксированной) из этих компонент связности, получим путь $\alpha_2(t)$ на сфере S^2. Значение интеграла в формуле $(*)$ для пути α равно значению такого же интеграла для пути α_2. Таким образом получаем, что в случае P_2 нижняя грань в формуле $(*)$ достигается на дуге, соединяющей проективные точки l' и l'' и состоящей из проективных точек l, для которых $l \cap S^2$ содержится в образах двух меньших дуг окружности большого круга на S^2, соединяющих точки $l' \cap S^2$ и $l'' \cap S^2$ в S^2. Такие дуги являются кратчайшими в P_2. Сама метрика равна длине любой из этих двух меньших дуг, соединяющих точки $l' \cap S^2$ и $l'' \cap S^2$ и равна радианной мере угла, о котором говорилось выше.

Прямой в римановом многообразии называется такая кратчайшая геодезическая α, что не существует кратчайшей геодезической α_1, для которой $R(\alpha) \subset R(\alpha_1)$. Интуитивно это определение означает, что α нельзя продолжить на X за пределы $R(\alpha)$.

Итак, евклидовы системы $\langle X, Y, \rho \rangle$ обычно определяются следующим образом: X — двумерное риманово многообразие, ρ — метрика в этом многообразии, определяемая по формуле $(*)$, и Y — все прямые в X.

§ 4. ОСНОВНЫЕ ГРУППЫ ШКОЛЬНОЙ ПЛАНИМЕТРИИ И ИХ ДЕЙСТВИЕ В ПЛОСКОСТИ

1. Аффинные отображения. В школе функцию $y = a \cdot x + b$ называют линейной. В высшей математике линейной называют функцию, обладающую свойствами:

$$f(x_1+x_2)=f(x_1)+f(x_2),$$
$$f(\lambda\cdot x)=\lambda\cdot f(x).$$

Функция $y=a\cdot x+b$ в случае $b\neq 0$ не обладает ни одним из этих свойств и в этом смысле не может быть названа линейной. Откуда же возникло школьное название? Дело в том, что $y=a\cdot x+b$ — линейная функция относительно аффинной структуры в арифметической прямой $\boldsymbol{R}$; поэтому ее правильнее называть аффинно-линейной или короче — аффинной функцией. Далее рассматриваются аффинные плоскости только над полем $\boldsymbol{R}$.

Определение 4. **Пусть $\langle A, X, g\rangle$ аффинная плоскость. Отображение $f: A\to A$ называется аффинным, если найдется вспомогательное отображение $h: X\to X$, обладающее двумя свойствами:**

1) для любых точек $a, b\in A$ выполняется

$$h(\overline{a, b})=\overline{f(a), f(b)},$$

2) h — линейное отображение из X в X.

Как мы увидим, для аффинного отображения f существует ровно одно отображение h, обладающее этими двумя свойствами. Оно называется дифференциалом отображения f.

Рассмотрим первое из этих свойств подробнее. Если X произвольное множество и f произвольное отображение из X в X, то через $f\otimes f$ обозначим отображение из $X\times X$ в себя, определяемое равенством

$$(f\otimes f)(\langle a, b\rangle)\rightleftarrows\langle f(a), f(b)\rangle.$$

Пусть $\langle A, \mathscr{X}, g\rangle$ — аффинная плоскость и f произвольное отображение из A в A. Первое из свойств в определении 4 заключается в том, что следующий набор функций: $f\otimes f$, g, g — коммутативен относительно какой-то функции h вида h: $\mathscr{X}\to\mathscr{X}$ (рис. 40). Иными словами, существует функция $h: \mathscr{X}\to\mathscr{X}$, для которой

$$h(\overline{a, b})=\overline{f(a), f(b)}.$$

Чтобы найти значение $h(\xi)$, где $\xi\in\mathscr{X}$, из условия коммутативности этой диаграммы, т. е. из приведенного выше равенства, нужно сделать два шага. Первый шаг состоит в том, что выбираются любые точки a и b из A, для которых $a, b=\xi$. Второй шаг — к паре элементов $\langle a, b\rangle$ применяется композиция отображений $g\circ(f\otimes f)$, то есть образуется $\overline{f(a), f(b)}$. Второй шаг (после выбора точек a и b) выполняется однозначно, но первый шаг, т. е. выбор самих точек a и b, для которых $a, b=\xi$, неоднозначен: существует бесконечно много различных точек a и b, для которых $\overline{a, b}=\xi$. Свойство коммутативности диаграммы как раз означает, что при фиксированном векторе ξ значение $g(\overline{f(a), f(b)})$ при условии $a, b=\xi$ не зависит от выбора точек a и b. Иначе говоря, если взять другие точки a' b', для которых $a', b'=\xi$, и применить к паре $\langle a', b'\rangle$ композицию отображений $g\circ(f\otimes f)$, то коммутативность диаграммы означает, что результат будет таким же, как при исходном выборе точек a и b.

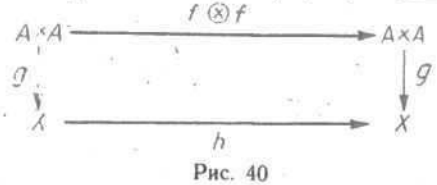


Рис. 40

В нашем случае, равенство $\overline{a, b}=\xi$ означает, что $b=a+\xi$, поэтому коммутативность диаграммы означает, что $\overline{f(a), f(a+\xi)}$ не зависит от выбора точки $a\in A$ при фиксированном векторе ξ. Следовательно, вектор $\overline{f(a), f(a+\xi)}$ зависит только от вектора ξ и поэтому можно определить отображение вида $\xi\mapsto\overline{f(a), f(a+\xi)}$, которое и обозначено на диаграмме h. Отсюда ясно, что первое свойство в определении 4 уже обеспечивает, что h по f образуется единственным образом, а именно, $h(\xi)=\overline{f(a), f(a+\xi)}$. Итак, первое свойство в определении 4 аффинного отображения f заключается в существовании такого отображения $h: X\to X$, что

$$f(a+\xi)=f(a)+h(\xi) \qquad (*)$$

для любых точек $a\in A$ и любых векторов $\xi\in\mathscr{X}$.

Второе свойство в определении 4 состоит в том, что это однозначно определяемое по f отображение h линейно.

В геометрических терминах равенство $(*)$ означает, что переносу точки $a\in A$ на вектор $\xi\in\mathscr{X}$ соответствует при отображении $f: A\to A$ перенос точки $f(a)$ на вектор $h(\xi)\in\mathscr{X}$.

Итак, если f — аффинное отображение аффинной плоскости $\langle A, X, g\rangle$ в себя, то h, однозначно определяемое по f, называется дифференциалом отображения f и обозначается $D(f)$. Иными словами, $(Df)(a, b)\rightleftarrows\overline{f(a), f(b)}$ для всех точек $a, b\in A$. Такое $(Df)(\cdot): X\to X$ всюду определено и является линейным отображением. Используя дифференциал, можно переписать равенство $(*)$ в красивой форме

$$f(a+\xi)=f(a)+(Df)(\xi),\ \forall a\in A,\ \forall\xi\in\mathscr{X}.$$

Приведем примеры аффинных отображений и их дифференциалов.

а) В аффинной плоскости $\langle X, X, g\rangle$, где первое X — множество, а второе X — векторная плоскость и $g(a, b)\rightleftarrows b-a$, рассмотрим отображение $f: X\to X$, которое линейно, как отображение векторной плоскости X в себя. Тогда f является аффинным отображением аффинной плоскости X в себя и $D(f)=f$. Действительно,

$$f(\xi)=f(a+\xi)-f(a)=g(f(a), f(a+\xi))=\overline{f(a), f(a+\xi)},$$

поэтому $(Df)(\xi)=f(\xi)$.

б) Пусть $\xi\in\mathscr{X}$ и $t_\xi(\cdot)$ — параллельный перенос в аффинной

плоскости $\langle A, \mathscr{X}, g\rangle$, т. е. $t_\xi : a \mapsto a+\xi$. Тогда t_ξ аффинное отображение и $D(t_\xi) = \mathrm{id}_X$ — тождественному отображению X на себя.

Действительно, если положить $(Dt_\xi) \rightleftarrows \mathrm{id}_X$, то Dt_ξ линейное отображение X на себя и $(Dt_\xi)(\eta) = \eta = ((a+\eta)+\xi)-(a+\xi) = t_\xi(a+\eta)-t_\xi(a) = \overline{t_\xi(a),\ t_\xi(a+\eta)}$.

в) Пусть $f : A \to A$ аффинное отображение аффинной плоскости $\langle A, X, g\rangle$ в себя и $\xi \in \mathscr{X}$. Определим $f_1(a) \rightleftarrows f(a)+\xi$ для всех $a \in A$. Тогда f_1 аффинное отображение и $Df_1 = Df$.

Действительно, Df_1 — линейно и $(Df_1)(\eta) = \overline{f_1(a),\ f_1(a+\eta)} = \overline{f(a),\ f(a+\eta)}$, так как $\overline{f_1(a),\ f_1(a+\eta)} = \overline{f_1(a),\ f(a)} + \overline{f(a),\ f(a+\eta)} + \overline{f(a+\eta),\ f_1(a+\eta)} = -\xi + \overline{f(a),\ f(a+\eta)} + \xi = \overline{f(a),\ f(a+\eta)}$.

г) Пусть f линейное отображение векторной плоскости X в себя и ξ — фиксированный элемент из X. Из примеров а) — в) видно, что в аффинной плоскости $\langle X, X, g\rangle$, определенной в примере а), всякое отображение $f_1(x) \rightleftarrows f(x)+\xi$ является аффинным отображением.

Покажем, что верно и обратное утверждение: всякое аффинное отображение f_1 аффинной плоскости $\langle X, X, g\rangle$ в себя представимо в виде $f_1(x) = f(x)+\xi$ для некоторого вектора $\xi \in X$ и некоторого линейного отображения f векторной плоскости X в себя. Действительно, положим $\xi \rightleftarrows f_1(0)$ и $f(x) \rightleftarrows f_1(x)-\xi$. Тогда $f_1(x) = f(x)+\xi$ для всех $x \in X$ и f — линейное отображение X в себя, так как в силу аффинности f_1 получаем:

$$f_1(a+x) = f_1(a) + (Df_1)(x).$$

Положив $a \rightleftarrows 0$, получаем $f_1(x) = f_1(0) + (Df_1)(x)$,

$$f(x) = f_1(x) - \xi = f_1(x) - f_1(0) = (Df_1)(x),$$

т. е. $f(x) = (Df_1)(x)$, а D_{f_1} — по определению линейная функция.

д) Рассмотрим одномерный вариант предыдущего примера. Любое линейное отображение f одномерного векторного пространства $\boldsymbol{R}$ в себя имеет вид (см. главу I, § 1):

$$f(x) = a \cdot x, \text{ где } a \rightleftarrows f(1).$$

Поэтому множество всех функций вида $y = a \cdot x + b$, где a и b из $\boldsymbol{R}$, совпадает с множеством всех аффинных отображений аффинной прямой $\boldsymbol{R}$ в себя. Поэтому функции вида $y = ax+b$ правильно назвать аффинными.

е) В примере в) показано: если отображение f_1 является композицией аффинного отображения f и переноса $t_\xi(\cdot)$, т. е. $f_1(x) \rightleftarrows f(x)+\xi$, то f_1 также аффинное отображение и $Df_1 = Df$. Покажем, что выполняется и обратное утверждение: из совпадения дифференциалов двух аффинных отображений f и f_1 следует, что каждое из них является композицией другого с некоторым параллельным переносом.

Пусть f и f_1 аффинные отображения с одинаковыми дифференциалами, т. е. $Df = Df_1$. Фиксируем точку $a \in A$ и положим

$\xi \rightleftarrows \overline{f(a),\ f_1(a)}$. Докажем, что для любой точки $b \in A$ выполняется равенство $f_1(b) = f(b)+\xi$.

Обозначая $\eta \rightleftarrows \overline{a,\ b} \in \mathscr{X}$, получаем:

$$f_1(b) = f_1(a+\eta) = f_1(a) + (Df_1)(\eta) = f_1(a) + (Df)(\eta) = f_1(a) + \overline{f(a),\ f(a+\eta)} - \xi + \xi) = f_1(a) + \overline{f_1(a),\ f(a+\eta)} + \xi = f(a+\eta) + \xi = f(b) + \xi.$$

Упражнение 26. *Докажите, что отображение аффинной плоскости в себя является аффинным тогда и только тогда, когда оно переводит любую прямую снова в прямую.*

Предложение 5. ***а) Композиция $f \circ h$ двух аффинных отображений f и h аффинной плоскости в себя снова является аффинным отображением и $D(f \circ h) = D(f) \circ D(h)$.***

б) Аффинное отображение f аффинной плоскости $\langle A, \mathscr{X}, g\rangle$ в себя является биекцией множества A на себя тогда и только тогда, когда дифференциал Df является биекцией множества $\mathscr{X}$ на себя.

▷ а) Так как Df и Dh линейны, то композиция $Df \circ Dh$ — линейное отображение X в себя.

Кроме того, $(f \circ h)(a+\xi) = f(h(a+\xi)) = f(h(a) + Dh(\xi)) = f(h(a)) + Df(Dh(\xi)) = (f \circ h)(a) + (Df \circ Dh)(\xi)$.

Следовательно, отображение $f \circ h$ аффинно и его дифференциал равен $(Df) \circ (Dh)$.

б) Выполняется $f(a) = f(b) \Leftrightarrow \overline{f(a),\ f(b)} = 0 \Leftrightarrow (Df)(\overline{a,\ b}) = 0 \Leftrightarrow a,\ b \in \mathrm{Ker}\,(Df)$.

Если Df инъективно, то в силу линейности Df, $\mathrm{Ker}\,(Df) = \{0\}$, и поэтому $f(a) = f(b) \Leftrightarrow a = b$, т. е. f инъективно.

Если f инъективно, т. е. $f(a) = f(b) \Leftrightarrow a = b$, то $Df(\overline{a,\ b}) = 0 \Leftrightarrow a = b$, т. е. $Df(\overline{a,\ b}) = 0 \Leftrightarrow a,\ b = 0$. Поэтому $\mathrm{Ker}\,(Df) = \{0\}$ и Df инъективно.

Пусть Df сюръективно и a — некоторая фиксированная точка из A. Для любой точки $b \in A$ вектор $\overline{f(a),\ b} \in X$ и найдется такой вектор $\eta \in X$, что $(Df)(\eta) = \overline{f(a),\ b}$. Поэтому $f(a+\eta) = f(a) + Df(\eta) = f(a) + \overline{f(a),\ b} = b$, т. е. f сюръективно.

Если f сюръективно, то для любого вектора $\xi \in X$ найдется такая точка $b \in A$, что $f(b) = f(a)+\xi$, где a — некоторая фиксированная точка из A. Тогда, обозначая $\eta \rightleftarrows \overline{a,\ b}$, получим:

$$(Df)(\eta) = \overline{f(a),\ f(a+\eta)} = \overline{f(a),\ f(b)} = \xi,$$

т. е. Df сюръективно. ■

Упражнение 27. *Докажите, что:*

а) множество всех линейных биекций векторной плоскости X на себя с композицией в качестве бинарной операции является группой;

б) множество всех аффинных биекций аффинной плоскости $\langle A, X, g\rangle$ с композицией в качестве бинарной операции является группой.

Пусть $h: X \to X$ — линейное отображение и e — произвольный базис в X. Обозначим $T_e(h)$ матрицу отображения h в базисе e, т. е. $T_e(h)\, W_e^{-1}(\xi) = W_e^{-1}(h(\xi))$, $\forall \xi \in X$. Обозначим $\det h$ определитель матрицы $T_e(h)$.

Предложение 6. ***а) Для любого линейного отображения $h: X \to X$ число $\det h$ не зависит от выбора базиса e.***

б) Аффинное отображение $f: A \to A$ является биекцией в том и только в том случае, если $\det Df \neq 0$.

Пункт а доказывается прямым подсчетом. Пункт б непосредственно вытекает из предложения 5б.

Определение 5. **Множество всех аффинных биекций аффинной плоскости $\langle A, X, g \rangle$ с композицией в качестве групповой операции называется аффинной группой и обозначается $\mathrm{Aff}(A)$.**

2. Основные группы школьной планиметрии, действующие в арифметической плоскости. До конца этого параграфа мы перечислим основные группы биекций плоскости на себя и одновременно отметим некоторые их свойства. Эти группы с определенной точки зрения лежат в основании школьной планиметрии.

В этом пункте рассмотрим случай арифметической плоскости $\boldsymbol{R}^2$, а в следующем пункте перенесем указанные для $\boldsymbol{R}^2$ группы биекций в векторную и аффинную плоскости.

Любое линейное отображение W из $\boldsymbol{R}^2$ в себя однозначно определяет матрицу $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = A$, для которой $\forall x \in \boldsymbol{R}^3 (W(x) = A \cdot x)$, так как в $\boldsymbol{R}^2$ имеется выделенный базис $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$. А именно, $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \rightleftarrows W\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$ и $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \rightleftarrows W\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$.

Обычно само линейное отображение W и соответствующая ему матрица A обозначаются одной и той же буквой. При этом композиции линейных отображений соответствует произведение этих матриц. Поэтому группа, состоящая из всех линейных отображений $\boldsymbol{R}^2$ в себя, и любая ее подгруппа называются матричными группами.

Множество всех вещественных (2×2) — матриц обозначим $\mathrm{Mat}(2, \boldsymbol{R})$ или кратко M_2. Множество M_2 отождествляется с множеством $\boldsymbol{R}^4$ с помощью биекции: $A \leftrightarrow (a_{11}, a_{12}, a_{21}, a_{22})$. По этой биекции в M_2 переносится структура метрического пространства, а именно: $\rho((a_{ij}), (b_{ij})) \rightleftarrows \left(\sum_{i,j=1}^{2} (a_{ij} - b_{ij})^2 \right)^{\frac{1}{2}}$. Значит, предыдущая биекция — изоморфизм метрических пространств M_2 и $\boldsymbol{R}^4$. Из неравенства $|a_{11} - b_{11}| \leqslant \rho((a_{ij}), (b_{ij}))$ следует, что функция $(a_{ij}) \mapsto a_{11}$ вида $M_2 \to \boldsymbol{R}$ является непрерывной. Аналогичным образом непрерывны и функции

$$(a_{ij}) \mapsto a_{12},\ (a_{ij}) \mapsto a_{21},\ (a_{ij}) \mapsto a_{22}$$

вида $M_2 \to \boldsymbol{R}$.

Функция $(a_{ij}) \mapsto \det(a_{ij})$ непрерывна, так как $\det(a_{ij}) = a_{11} \cdot a_{22} -$

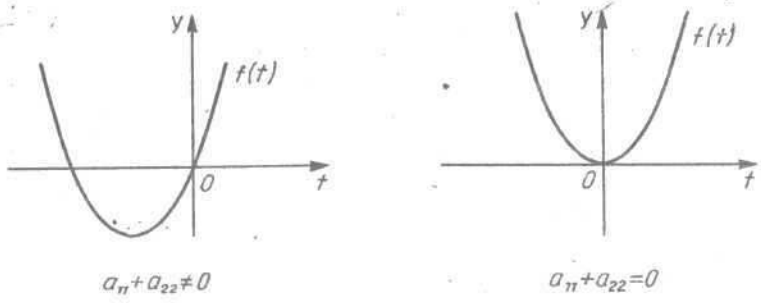


Рис. 41

$- a_{12} \cdot a_{21}$ и правая часть этого равенства является разностью произведений непрерывных функций.

а) Общая линейная группа по определению состоит из всех невырожденных матриц вместе с обычной операцией умножения двух матриц. Она обозначается GL_2. Так как функция $\det: (a_{ij}) \mapsto \mapsto \det(a_{ij})$ непрерывна и $\mathrm{GL}_2 = \{(a_{ij}) | \det(a_{ij}) \neq 0\}$, то GL_2 является прообразом открытого множества $\boldsymbol{R} \setminus \{0\}$ относительно непрерывной функции $\det(\cdot)$. Поэтому множество GL_2 открыто в метрическом пространстве M_2.

Покажем, что множество GL_2 плотно в пространстве M_2, т. е. в любой окрестности любой матрицы (a_{ij}) точки из M_2 найдется матрица точка из множества GL_2.

Если $(a_{ij}) \in GL_2$, то в силу открытости GL_2 элемент (a_{ij}) имеет окрестность, целиком содержащуюся в GL_2, т. е. целиком состоящую из невырожденных матриц.

Пусть $(a_{ij}) \notin GL_2$, т. е. $\det(a_{ij}) = a_{11}a_{22} - a_{12}a_{22} = 0$.

Определим функцию $f(t) \rightleftarrows \det \begin{pmatrix} a_{11}+t & a_{12} \\ a_{21} & a_{22}+t \end{pmatrix}$, где $t \in \boldsymbol{R}$. Тогда $f(t) = (a_{11}+t)(a_{22}+t) - a_{12} \cdot a_{21} = t^2 + t(a_{11}+a_{22}) + a_{11}a_{22} - a_{12} \cdot a_{21} = = t^2 + t(a_{11}+a_{22})$.

Функция $f(t)$ является параболой, проходящей через начало координат, ветви которой направлены вверх (рис. 41).

Для малых чисел t, отличных от нуля, $f(t)$ также отлично от нуля и соответствующие этим t матрицы $\begin{pmatrix} a_{11}+t & a_{12} \\ a_{21} & a_{22}+t \end{pmatrix}$ принадлежат GL_2.

Поскольку

$$\rho\left(\begin{pmatrix} a_{11}+t & a_{12} \\ a_{21} & a_{22}+t \end{pmatrix}, \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right) = \sqrt{2} \cdot t,$$

то матрица $A(t) \rightleftarrows \begin{pmatrix} a_{11}+t & a_{12} \\ a_{21} & a_{22}+t \end{pmatrix}$ при $t \to 0$, $t > 0$ стремится к матрице $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Следовательно, в любой окрестности матрицы $A = (a_{ij})$ найдется невырожденная матрица A_t.

б) Линейная группа по определению состоит из всех матриц определителем, равным $+1$ или -1 вместе с обычной операцией умножения двух матриц. Она обозначается L_2 и является подгруппой общей линейной группы GL_2. Она замкнута в M_2, так как является прообразом замкнутого множества $\{-1, 1\}$ относительно непрерывной функции $\det(\cdot)$. Линейная группа L_2, как мы увидим в главе III, играет решающую роль в понятиях площади и меры плоской фигуры. А именно L_2 — наибольшая из матричных групп, под действием которых площадь фигуры не меняется. Отметим ее подгруппу SL_2 (специальная линейная группа), по определению состоящую из всех матриц $A \in M_2$, для которых $\det A = +1$.

в) Ортогональная группа по определению состоит из всех матриц, для которых совпадают транспонированная и обратная матрицы, вместе с обычной операцией умножения двух матриц. Она обозначается O_2. Носитель группы O_2 можно определить и как множество всех матриц, сохраняющих обычное скалярное произведение в $\boldsymbol{R}^2$ (см. гл. I, § 7).

Если матрица $A \in O_2$, то $A^t = A^{-1}$ и $\det A = \det(A^t) = \det(A^{-1}) = \frac{1}{\det A}$.

Следовательно, $(\det A)^2 = 1$ и $\det A = \pm 1$. Поэтому O_2 — подгруппа L_2.

Отметим подгруппу SO_2 ортогональной группы (группа поворотов), которая по определению состоит из всех матриц $A \in O_2$ с определителем, равным $+1$. Напомним (гл. I, § 7), что всякий элемент $A \in SO_2$ имеет вид $\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$, где t — некоторое фиксированное число из R.

Следовательно, SO_2 — ограниченное множество в метрическом пространстве M_2. Нетрудно показать, что всякий элемент $A \in (O_2 \setminus SO_2)$ имеет вид $\begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}$, где t — некоторое фиксированное число из $\boldsymbol{R}$.

Следовательно, $O_2 = (SO_2) \cup (O_2 \setminus SO_2)$ — ограниченное множество в метрическом пространстве M_2.

Из непрерывности отображений $W \mapsto W^t$ и $W \mapsto W^{-1}$ вытекает, что O_2 и SO_2 — замкнутые множества в пространстве M_2. Поэтому O_2 и SO_2 — компактные множества и, следовательно, компактные матричные группы.

г) Группа гомотетий по определению состоит из всех матриц вида $t \cdot E$, где t — фиксированное число из $\boldsymbol{R}^*$ и E — единичная матрица, вместе с обычным умножением двух матриц в качестве групповой операции. Эта группа обозначается H_2. Группа H_2, очевидно, изоморфна группе $\boldsymbol{R}^*$ — ненулевых вещественных чисел с умножением в качестве групповой операции.

Ее подгруппа SH_2 (группа собственных гомотетий) по определению состоит из всех матриц вида $t \cdot E$, где $t > 0$. Она изоморфна группе $\boldsymbol{R}^*$ — всех строго положительных вещественных чисел с

операцией умножения чисел (см. гл. I, § 2). Группы SH_2 и H_2 замкнутые подгруппы общей линейной группы GL_2.

д) Группа линейных подобий по определению состоит из всех матриц, получаемых в результате умножения матриц из ортогональной группы O_2 на матрицы из группы гомотетий H_2, вместе с операцией умножения матриц. Группа линейных подобий, очевидно, изоморфна прямому произведению $O_2 \times \boldsymbol{R}^*$.

Группа собственных линейных подобий состоит из всех матриц, получаемых в результате умножения матриц из SO_2 на матрицы из SH_2; она изоморфна прямому произведению $O_2 \times \boldsymbol{R}^*$. Итак,

$$\begin{array}{ccccccc} SL_2 & \subset & L_2 & \subset & & & GL_2 \subset M_2 \\ \cup & & \cup & & & & \cup \\ SO_2 & \subset & O_2 & \subset & O_2 \times H_2 & \supset & H_2 \\ & & \cap & & & & \cup \\ & & O_2 \times SH_2 & & \supset & & SH_2 \end{array}$$

Здесь все включения, кроме $GL_2 \subset M_2$, являются включениями подгруппы в группу и все группы, кроме GL_2, замкнуты в пространстве M_2. Из этих групп только SO_2 и O_2 — компактные группы.

Опишем действие матричных групп в множестве $\boldsymbol{R}^2$. Пусть G — любая из приведенных выше матричных групп и $A \in G$. Кратко говоря, действие матриц A из группы G на точку x из $\boldsymbol{R}^2$ по определению состоит в умножении матрицы A на эту точку x, т. е. $A \mapsto A \cdot x$. Здесь функция, состоящая в умножении фиксированной матрицы A на переменную точку (вектор) из $\boldsymbol{R}^2$, обозначается $A \cdot x$ подобно тому, как функция синус обозначается $\sin x$.

Подробнее действие группы G определяется как отображение F, задаваемое равенством $F(A) \rightleftarrows A \cdot x$ для всех матриц $A \in G$. Значение отображения F, т. е. $F(A)$, также является функцией, но уже вида $\boldsymbol{R}^2 \to \boldsymbol{R}^2$. Функция $A \cdot x$ является линейной биекцией, так как G группа и, следовательно, вместе с любым A она содержит A^{-1}, т. е. $\det A \neq 0$. Итак, F — отображение матричной группы G в группу $\mathrm{Is}(\boldsymbol{R}^2)$, т. е. F имеет вид $F : G \to \mathrm{Is}(\boldsymbol{R}^2)$. При этом F — гомоморфизм, так как

$$F(A \cdot B) = (A \cdot B) \cdot x = A \cdot (Bx) = F(A) \circ F(B).$$

Обозначим $F(A) \rightleftarrows f_A$.

Предложение 7. ***Пусть G — матричная группа и F построенное выше отображение G в $\mathrm{Is}(\boldsymbol{R}^2)$. Тогда F является эффективным действием группы G в множестве $\boldsymbol{R}^2$.***

▷ Как проверялось выше, произведению матриц A_1 и A_2 соответствует линейное отображение, являющееся композицией линейных отображений f_{A_1} и f_{A_2}. Поэтому F — гомоморфизм из G в $\mathrm{Is}(\boldsymbol{R}^2)$. Если $f_A = \mathrm{id}_{\boldsymbol{R}^2}$, то $A = E$, т. е. ядро гомоморфизма F состоит из $\{E\}$. Следовательно, F инъективный гомоморфизм (эффективное действие) G в $\boldsymbol{R}^2$. ■

Мы будем отождествлять матрицы и соответствующие им линей-

ные функции и поэтому будем отождествлять любую матричную группу G с ее образом $F(G)\subseteq \mathrm{Is}(\boldsymbol{R}^2)$, где F — описанное выше эффективное действие G в $\boldsymbol{R}^2$.

Действие $F: A\mapsto f_A$ матричной группы G в множестве $\boldsymbol{R}^2$ не является транзитивным. Например, все отображения f_A линейны и, следовательно, все они переводят начало координат в себя. Существует простой способ «усилить» действие F таким образом, чтобы оно стало транзитивным. Согласно § 2 этой главы группа $(\boldsymbol{R}^2)^+$ транзитивно и эффективно действует в множестве $\boldsymbol{R}^2$ при помощи параллельных переносов $t_\xi(\cdot): x\mapsto x+\xi$, «добавление» группы $\boldsymbol{R}^2$ к любой матричной группе G определяет уже эффективное и транзитивное действие новой группы в множестве $\boldsymbol{R}^2$. А именно, пусть G — матричная группа. В множестве $G\times\boldsymbol{R}^2$ рассмотрим следующую бинарную операцию:

$$(A_1,\ \xi_1)\cdot(A_2,\ \xi_2)\rightleftarrows(A_1\cdot A_2;\ A_1\cdot\xi_2+\xi_1).$$

Упражнение 28. *Докажите, что:*

а) множество $G\times\boldsymbol{R}^2$ замкнуто относительно этой операции;

б) эта бинарная операция ассоциативна;

относительно этой бинарной операции

в) пара $(E,\ 0)$ является нейтральным элементом;

г) пара $(A^{-1};\ -A^{-1}\xi)$ является обратным элементом к элементу $(A,\ \xi)$.

Из упражнения 28 следует, что множество $G\times\boldsymbol{R}^2$ с определенной в нем бинарной операцией является группой. Эта группа называется полупрямым произведением групп G и $(\boldsymbol{R}^2)^+$. Определим действие этой группы в множестве $\boldsymbol{R}^2$:

$$f_{\langle A,\xi\rangle}\rightleftarrows(A\cdot x+\xi).$$

Функция $F(\langle A,\ \xi\rangle)\rightleftarrows f_{\langle A,\xi\rangle}$ является композицией линейного отображения A и параллельного переноса $t_\xi(\cdot)$. Следовательно, $f_{\langle A,\xi\rangle}$ — биекция множества $\boldsymbol{R}^2$ на себя и отображение F вида $F: G\times\boldsymbol{R}^2\to\mathrm{Is}(\boldsymbol{R}^2)$.

Предложение 8. ***Отображение $F:\langle A,\ \xi\rangle\mapsto f_{\langle A,\xi\rangle}$ является эффективным и транзитивным действием группы $G\times(\boldsymbol{R}^2)^+$ в множестве $\boldsymbol{R}^2$.***

▷ Проверим гомоморфность F и тем самым докажем, что F действие:

$$f_{\langle A_1,\xi_1\rangle}\circ f_{\langle A_2,\xi_2\rangle}(x)=f_{\langle A_1,\xi_1\rangle}(f_{\langle A_2,\xi_2\rangle}(x))=f_{\langle A_1,\xi_1\rangle}(A_2\cdot x+\xi_2)=A_1A_2x+A_1\xi_2+\xi_1=\langle A_1\cdot A_2,\ A_1\cdot\xi_2+\xi_1\rangle.$$

Докажем эффективность действия F, т. е. инъективность гомоморфизма F или, что эквивалентно, проверим равенство $\mathrm{Ker}\,F=\{(E,\ \bar{0})\}$.

Пусть $f_{\langle A,\xi\rangle}$ совпадает с единицей в группе $\mathrm{Is}(\boldsymbol{R}^2)$, т. е. $f_{\langle A,\ \xi\rangle}=\mathrm{id}_{\boldsymbol{R}^2}$. Следовательно, для всех точек $x\in\boldsymbol{R}^2$ выполняется $f_{\langle A,\xi\rangle}(x)=A\cdot x+\xi=\mathrm{id}_{\boldsymbol{R}^2}(x)=x$. Положив $x=0$, получим $\xi=0$ и тогда из равенства $Ax=x$, $\forall x\in\boldsymbol{R}^2$, получим $A=E$.

Докажем транзитивность f, т. е. для любых точек x и y из $\boldsymbol{R}^2$ найдется такой элемент $\langle A,\ \xi\rangle$ из $G\times\boldsymbol{R}^2$, что $f_{\langle A,\xi\rangle}(x)=y$. Для этого положим $A\rightleftarrows E\in G$ и $\xi\rightleftarrows y-x$. ■

Будем отождествлять элемент $\langle A,\ \xi\rangle\in(G\times\boldsymbol{R}^2)$ и соответствующую ему биекцию $f_{\langle A,\ \xi\rangle}$ множества $\boldsymbol{R}^2$ на себя. В главе III действие F используется для случаев, когда группа G совпадает с L_2 и O_2. Группу $G_1\rightleftarrows L_2\times(\boldsymbol{R}^2)^+$ будем называть эквиаффинной группой.

Рассмотрим аффинную плоскость $\langle\boldsymbol{R}^2,\ \boldsymbol{R}^2,\ g\rangle$, где

$$g(a,\ b)\rightleftarrows b-a.$$

Из примера г) в п. 1 следует, что множество всех аффинных отображений аффинной плоскости $\boldsymbol{R}^2$ в себя совпадает с множеством отображений $\boldsymbol{R}^2$ в себя вида $x\mapsto A\cdot x+\xi$, где A — линейное отображение и $\xi\in\boldsymbol{R}^2$. Следовательно, $\mathrm{GL}_2\times(\boldsymbol{R}^2)^+$ является как раз аффинной группой плоскости $\boldsymbol{R}^2$, т. е.

$$\mathrm{Aff}(\boldsymbol{R}^2)=\mathrm{GL}_2\times(\boldsymbol{R}^2)^+.$$

Группа $O_2\times(\boldsymbol{R}^2)^+$ называется группой движений аффинной плоскости $\boldsymbol{R}^2$. Можно доказать, что она состоит из всех биекций $\boldsymbol{R}^2$ на себя, сохраняющих обычную метрику в $\boldsymbol{R}^2$. Группа $\mathrm{SO}_2\times\boldsymbol{R}^2$ называется группой собственных движений аффинной плоскости $\boldsymbol{R}^2$. Группы $(O_2\times H_2)\times(\boldsymbol{R}^2)^+$ и $(O_2\times SH_2)\times(\boldsymbol{R}^2)^+$ называются соответственно группами подобий и собственных подобий аффинной плоскости $\boldsymbol{R}^2$.

3. Поднятие группы биекций в арифметической плоскости в векторную и аффинную плоскости. Пусть X — векторная плоскость над полем $\boldsymbol{R}$. Фиксируем в X некоторый базис $e=\{e_1,\ e_2\}$ и через W_e обозначим функцию вида $\boldsymbol{R}^2\to X$, определяемую равенством

$$W_e(\langle x_1,\ x_2\rangle)\rightleftarrows x_1e_1+x_2e_2.$$

Легко проверить, что эта функция — изоморфизм. С помощью изоморфизма $W_e:\boldsymbol{R}^2\to X$ любая биекция T плоскости $\boldsymbol{R}^2$ на себя «переносится» в X. А именно отображение $T\mapsto W_e\circ T\circ W_e^{-1}\rightleftarrows T_e$ каждой биекции $T\in\mathrm{Is}(\boldsymbol{R}^2)$ ставит в соответствие биекцию T_e множества X на себя. Это соответствие удобно показать на следующей коммутативной диаграмме (рис. 42).

Определение 6. **Пусть G — некоторая группа биекций R^2 на себя и e — некоторый базис в векторной плоскости X. Поднятием группы G в X относительно базиса e называется множество всех биекций плоскости $\mathscr{X}$ на себя вида $W_e\cdot T\cdot W_e^{-1}$, где $T\in G$, вместе с бинарной операцией композиции. Это поднятие будем обозначать G_e.**

Предложение 9. ***Поднятие G_e группы G в векторную плоскость $\mathscr{X}$ относительно произвольного базиса e является подгруппой группы $\mathrm{Is}(\mathscr{X})$ изоморфной исходной группе G.***

▷ Пусть $T_{1e},\ T_{2e}\in G_e$. Тогда $T_{1e}\circ T_{2e}=(W_e\circ T_1\circ W_e^{-1})\cdot(W_e\circ T_2\circ W_e^{-1})=W_e\circ(T_1\circ T_2)\circ W_e^{-1}=(T_1\circ T_2)_e\in G_e$ и $\mathrm{id}\,X=W_e\circ\mathrm{id}_{\boldsymbol{R}^2}\circ W_e^{-1}\in G_e$. Следо-

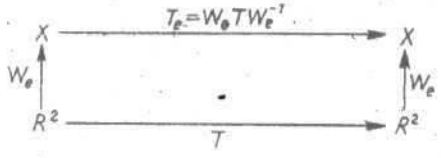


Рис. 42

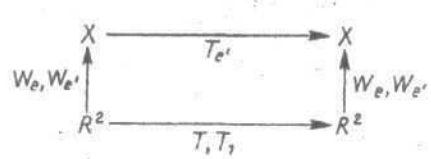


Рис. 43

вательно, множество G_e содержит единицу группы $\mathrm{Is}(X)$ и замкнуто относительно композиции, которая является групповой операцией в $\mathrm{Is}(X)$, т. е. G_e — подгруппа группы $\mathrm{Is}(X)$.

Докажем, что отображение вида $T \mapsto T_e$ является изоморфизмом групп G и G_e. Его гомоморфность следует из уже доказанного равенства $(T_1 \circ T_2)_e = (T_1)_e \circ (T_2)_e$. Это соответствие по определению сюръективно. Докажем его инъективность. Если $T_e = \mathrm{id}\, X$, т. е. $W_e \cdot T \cdot W_e^{-1} = \mathrm{id}_X$, то $T = W_e^{-1} \circ \mathrm{id}\, X \circ W_e = \mathrm{id}_{R^2}$. Следовательно, ядро соответствия $T \mapsto T_e$ состоит только из элемента id_{R^2}. ■

Если выбрать другой базис $e' = \{e_1', e_2'\}$ в векторной плоскости X, то возникает два поднятия в X группы G: одно — относительно базиса e и другое — относительно базиса e'. Возникает вопрос: совпадают ли группы G_e и $G_{e'}$ для различных базисов e и e'? Для одних групп G ответ положителен, для других — отрицателен.

Тем не менее по предложению 9 группы G_e и $G_{e'}$ всегда изоморфны между собой, так как обе они изоморфны группе G. Иначе говоря, поднятие G_e группы G в векторную плоскость X однозначно с точностью до изоморфизма групп.

В этом смысле группа G_e не зависит от выбора базиса e, и ее называют поднятием в плоскости X группы G.

Разберем подробнее несколько примеров.

а) Пусть $G = GL_2$ и e, e' — два базиса в X. Докажем, что в этом случае G_e (как множество!) совпадает с $G_{e'}$, т. е. множества G_e и $G_{e'}$ равны, $G_e = G_{e'}$.

Возьмем биекцию $T_{e'} \in G_{e'}$ и докажем, что $T_{e'} \in G_e$. Иными словами, найдем матрицу $T_1 \in GL_2$, для которой $T_{e'} = W_e T_1 W_e^{-1}$ (рис. 43).

Поскольку биекция $T_{e'} \in G_{e'}$, то для некоторой матрицы $T \in GL_2$ выполняется $T_{e'} = W_{e'} T W_{e'}^{-1}$. Обозначая $W_{e',e}$ изоморфизм $W_e^{-1} \cdot W_{e'}$, получим $T_{e'} = W_{e'} T W_{e'}^{-1} = W_e (W_e^{-1} W_{e'} T W_{e'}^{-1} W_e) W_e^{-1} = W_e (W_{e',e} T W_{e',e}^{-1}) W_e^{-1} = W_e T_1 W_e^{-1}$, где $T_1 \rightleftarrows W_{e',e} T W_{e',e}^{-1}$ является изоморфизмом $\boldsymbol{R}^2$ на себя как композиция изоморфизмов $W_{e',e}^{-1}$ и T и $W_{e',e}$ плоскости $\boldsymbol{R}^2$ на себя.

Следовательно, $T_1 \in GL_2$, и тем самым доказано включение $G_{e'} \subseteq G_e$.

Аналогично проверяется и обратное включение $G_e \subseteq G_{e'}$.

б) Пусть группа G совпадает с группой L_2 или группой SL_2. Докажем, что и в этом случае группа G_e (как множество) совпадает с группой $G_{e'}$. Аналогично предыдущему примеру, если $T_{e'} = W_{e'} T W_{e'}^{-1}$, где $T \in G$, то $T_{e'} = W_e T_1 W_e^{-1}$, где $T_1 = W_{e',e} T W_{e',e}^{-1}$, и

$W_{e',e}$ — изоморфизм $\boldsymbol{R}^2$ на себя. Тогда $\det T_1 = \det(W_{e',e}) \cdot (\det T) \cdot \frac{1}{\det W_{e',e}} = \det T$.

Следовательно, если $\det T = 1$, то $\det T_1 = 1$, и если $\det T = -1$, то и $\det T_1 = -1$. Поэтому если $T \in L_2$ или $T \in SL_2$, то $T_1 \in L_2$ или $T_1 \in SL_2$.

Значит, $T_{e'} = (T_1)_e$, где $T_1 \in G$. Поэтому $G_{e'} \subseteq G_e$. Обратное включение проверяется аналогично.

в) Пусть T — параллельный перенос в плоскости $\boldsymbol{R}^2$ на фиксированный вектор $\xi = \langle x_1, x_2 \rangle$. Тогда T_e — параллельный перенос в плоскости X на вектор $x_1 \cdot e_1 + x_2 \cdot e_2$, а $T_{e'}$ — перенос в X на вектор $x_1 e_1' + x_2 e_2'$.

Если $e_1' = a_{11} e_1 + a_{12} e_2$ и $e_2 = a_{21} e_1 + a_{22} e_2$, то $T_{e'}$ — перенос на вектор $(x_1 a_{11} + x_2 a_{21}) e_1 + (x_1 a_{12} + x_2 a_{22}) e_2$.

Следовательно, $T_{e'}$ является параллельным переносом в X, который получается поднятием с помощью $W_{e'}$ параллельного переноса в $\boldsymbol{R}^2$ на вектор $\xi' = \langle x_1 a_{11} + x_2 a_{21}, x_1 a_{12} + x_2 a_{22} \rangle$.

Поэтому группа $G = (\boldsymbol{R}^2)^+$ также дает пример группы биекций в $\boldsymbol{R}^2$, для которой поднятия G_e и $G_{e'}$ совпадают (как множества) для любых базисов e и e'.

г) Пусть $G = O_2$ или $G = SO_2$ и e, e' — два базиса в векторной плоскости X. В этом случае поднятия G_e и $G_{e'}$ группы G относительно базисов e и e', вообще говоря, не совпадают между собой (как множества). Однако если рассматривать евклидову плоскость X, т. е. векторную плоскость X с фиксированной в ней положительно определенной, симметричной и билинейной формой $g(\cdot, \cdot)$, то среди всех базисов в X выделяется некоторое особое множество базисов, относительно которых поднятия группы G в X совпадают между собой (как множества). А именно семейство таких базисов состоит из всех таких $e = \{e_1, e_2\}$, для которых $g(e_1, e_1) = g(e_2, e_2) = 1$ и $g(e_1, e_2) = 0$. Такие базисы называются ортонормированными относительно фиксированной формы $g(\cdot, \cdot)$.

Пусть $e = \{e_1, e_2\}$ — какой-то ортонормированный базис относительно евклидовой формы $g(\cdot, \cdot)$ в векторной плоскости X. Тогда изоморфизм $W_e : \boldsymbol{R}^2 \to X$ переводит обычное скалярное произведение в $\boldsymbol{R}^2$ в евклидову форму $g(\cdot, \cdot)$ в X. Точнее, для любых точек $x, y \in \boldsymbol{R}^2$ выполняется $(x, y) = g(W_e \cdot x, W_e \cdot y)$.

Действительно, $g(W_e \cdot x, W_e \cdot y) = g(x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = x_1 \cdot y_1 \cdot g(e_1, e_1) + x_1 \cdot y_2 \cdot g(e_1, e_2) + x_2 \cdot y_1 \cdot g(e_2, e_1) + x_2 y_2 \cdot g(e_2, e_2) = x_1 y_1 + x_2 y_2 = (x, y)$.

Докажем, что если e и e' — какие-то ортонормированные базисы в векторной плоскости X с евклидовой формой $g(\cdot, \cdot)$, то множества G_e и $G_{e'}$ совпадают, $G_e = G_{e'}$. Пусть $T_{e'} = W_{e'} T W_{e'}^{-1}$, где $T \in G$. Тогда, как и в п. а и б, выполняется:

$$T_{e'} = W_e (W_e^{-1} W_{e'} T W_{e'}^{-1} W_e) W_e^{-1} = W_e T_1 W_e^{-1},$$

где $T_1 \rightleftarrows W_{e',e} T W_{e',e}^{-1}$ и $W_{e',e} \rightleftarrows W_e^{-1} W_{e'}$.

Так как W_e и $W_{e'}$ переводят скалярное произведение в $\boldsymbol{R}^2$ в евклидову форму $g(\cdot, \cdot)$ в X, то изоморфизм $W_{e', e}$ сохраняет скалярное произведение в $\boldsymbol{R}^2$, и тем самым $W_{e', e} \in O_2$.

Следовательно, $T_1 = W_{e', e} T W_{e', e}^{-1} \in O_2$.

Если $T \in SO_2$, т. е. $T \in O_2$ и $\det T = 1$, то $T_1 \in SO_2$. Поэтому $T_1 \in G$ и $T_{e'} = W_e T_1 W_e^{-1} \in G_e$. Следовательно, $G_{e'} \subseteq G_e$. Обратное включение следует из того, что все рассуждения можно повторить, поменяв местами ортонормированные базисы e и e'.

д) Аналогичным образом поднимаются в X и группы биекций в $\boldsymbol{R}^2$ вида $G \times (\boldsymbol{R}^2)^+$, где G — матричная группа.

Нетрудно показать, что поднятие группы $G \times (\boldsymbol{R}^2)^+$ относительно базиса e совпадает с группой $G_e \times X^+$, где G_e — поднятие группы G в векторную плоскость X относительно базиса e.

Поэтому, например, поднятие аффинной группы $GL_2 \times (\boldsymbol{R}^2)^+$ аффинной плоскости $\boldsymbol{R}^2$ относительно любого базиса e не зависит от выбора этого базиса и совпадает с аффинной группой векторной плоскости $\langle X, X, g \rangle$.

Поднятие группы биекций векторной плоскости $\mathscr{X}$ в аффинную плоскость $\langle A, \mathscr{X}, g \rangle$ выполняется при помощи биекций f_a вида $f_a : A \to \mathscr{X}$. Точнее, если T — биекция в $\mathscr{X}$, то определяется биекция T_a в A вида $T_a \rightleftarrows f_a^{-1} \cdot T \cdot f_a$.

Если G группа биекций в $\mathscr{X}$, то в множестве $\{T_a | T \in G\}$, конечно, рассматривается операция композиции; полученная группа обозначается G_a.

В зависимости от выбора группы биекций в плоскости $\mathscr{X}$ ее поднятие в аффинную плоскость $\langle A, X, g \rangle$ зависит или не зависит от выбора точки a из A.

Иными словами, множества $\{T_a | T \in G\}$ и $\{T_b | T \in G\}$, где G группа биекций в $\mathscr{X}$, могут совпадать (как два подмножества) для любых точек a, b из A или для некоторых точек a и b.

В любом случае группы G_a и G_b изоморфны и в этом смысле группа G_a не зависит от выбора точки a из A.

§ 5. ПОНЯТИЕ ПЛАНИМЕТРИИ

Цель этого параграфа — изложить представление о планиметрии, предложенное Ф. Клейном.

1. Клейновский подход в геометрии: понятие о планиметрии данной группы. Рассмотрим совершенно произвольные множества X и Y, группу G, действие $F(\cdot)$ группы G на множестве X, т. е. гомоморфизм $F(\cdot) : G \to \mathrm{Is}(X)$, где $\mathrm{Is}(X)$ — группа всех биекций множества X на себя.

Определение 7а. **Функция $\psi : X^n \to Y$, где $n \in N$, называется G-инвариантом (или точнее инвариантом действия F группы G на множестве), если $\forall g \in G \forall x_1, \ldots, x_n \in X (\psi(x_1, \ldots, x_n) = \psi(f_g(x_1), \ldots, f_g(x_n)))$.**

Если Y — числовое множество, то ψ называется скалярным G-инвариантом. Если $X = U^m$ и U является носителем некоторого кольца, а ψ — многочлен от m переменных с коэффициентами из этого кольца, то ψ называется *алгебраическим G-инвариантом.*

Например, рассмотрим в качестве ψ обычное расстояние между двумя точками в $\boldsymbol{R}^2$. В этом случае положим $X = \boldsymbol{R}^2$, $n = 2$, $\psi = \rho(\cdot, \cdot)$, $Y = \boldsymbol{R}_{\geqslant 0}$, $G = O_2 \times \boldsymbol{R}^2$. Функция F определяется как всегда:

$$F(\langle A, \xi \rangle) = A \cdot x + \xi : \boldsymbol{R}^2 \to \boldsymbol{R}^2.$$

Тогда функция ψ является G-инвариантом, т. е.

$$\forall \langle A, \xi \rangle \in G \ \forall x, y \in R^2 (\psi(x, y) = \psi(A \cdot x + \xi, A \cdot y + \xi)).$$

Рассмотрим следующий пример. Пусть ψ — обычное скалярное произведение в $\boldsymbol{R}^2$, т. е. $\psi(x, y) = (x, y) \rightleftarrows x_1 \cdot y_1 + x_2 \cdot y_2$. Тогда положим $X = R^2$ и $n = 2$. Функция $\psi(x_1, x_2, y_1, y_2) \rightleftarrows x_1 \cdot y_1 + x_2 \cdot y_2$ есть многочлен над $\boldsymbol{R}$ от четырех переменных. Пусть $Y = \boldsymbol{R}$ и $G = O_2$ и, как всегда, $F(A) \rightleftarrows A \cdot x : \boldsymbol{R}^2 \to \boldsymbol{R}^2$. Тогда функция ψ является O_2-инвариантом, так как согласно одному из определений ортогональной матрицы выполняется соотношение $\forall A \in O_2 \ \forall x, y \in \boldsymbol{R}^2 (\psi(x, y) = \psi(A \cdot x, A \cdot y))$. Более того, ψ является алгебраическим O_2-инвариантом.

Бывают полезны G-инварианты иного рода.

Определение 7б. **Пусть ψ — произвольная функция вида $P(X) \to Y$, которая, может быть, частично определена. Функция $\psi : \mathscr{P}(X) \to Y$ называется G-инвариантом (или, точнее, инвариантом действия F группы G на множестве X), если**

$$\forall g \in G \ \forall w \in D(\psi) (\psi(w) = \psi(f_g(w))),$$

где $D(\psi)$ — область определения функции $\psi(\cdot)$, $D(\psi) \subseteq \mathscr{P}(X)$.

Например, обозначим U множество всех многоугольников (или для простоты — всех треугольников) в плоскости $\boldsymbol{R}^2$. Площадь $T(\cdot)$ многоугольника (соответственно треугольника) не меняется под действием группы $G_s \rightleftarrows L_2 \times (\boldsymbol{R}^2)^+$ на множестве $\boldsymbol{R}^2$. А именно $\forall \langle A, \xi \rangle \in G_s \ \forall u \in U (T(u) = T(A \cdot u + \xi))$, где $A \cdot u + \xi \rightleftarrows \{A \cdot x + \xi | x \in U\}$. Для случая многоугольников это утверждение доказывается в главе III, для случая треугольников его легко и полезно проверить непосредственно. Итак, площадь многоугольника или треугольника также является примером G_s-инварианта.

Другие аналогичные примеры можно построить, если в качестве U выбрать семейство всех прямоугольников или еще какое-нибудь семейство элементарных плоских фигур.

Наконец, приведем определение еще одного вида G-инвариантов.

Определение 7в. **Пусть $P(x_1, \ldots, x_n)$ — n-местный предикат на множестве X, т. е. функция вида $P : X^n \to \{0, 1\}$. Предикат P называется G-инвариантом (G-инвариантным предикатом), если**

$$\forall g \in G \forall x_1, \ldots, x_n \in X (P(x_1, \ldots, x_n) \Leftrightarrow P(f_g(x_1), \ldots, f_g(x_n))),$$

где F — действие группы G на множестве X.

Можно рассматривать G-инвариантный предикат P как частный случай G-инвариантной скалярной функции в смысле определения 7а. В этом случае нужно положить $Y \rightleftarrows \{0, 1\}$.

Например, $X = \boldsymbol{R}^2$ и $n = 4$ и $P(x_1, x_2, x_3, x_4) \rightleftarrows x_1 \neq x_2 \wedge x_3 \neq x_4 \wedge (x_1x_2) \| (x_2x_4)$, где (x_1x_2) — прямая, проходящая через точки x_1 и x_2, а (x_3x_4) — прямая, проходящая через точки x_3 и x_4, и знак $\|$ означает «параллельно». Очевидно, что этот предикат является $(O_2 \times \boldsymbol{R}^2)$-инвариантом. Можно заменить знак $\|$ на знак $\perp$ — перпендикулярно.

Следующий пример. Пусть $X = P(R^2)$, $n = 1$ и $P(w)$ — предикат «$w \subseteq \boldsymbol{R}^2$ и w — треугольник (или многоугольник и т. п.)». Тогда $P(\cdot)$ есть GL_2-инвариантный предикат. Этот пример подробно рассматривается в следующей главе.

Аналогично определяется понятие G-инварианта в случаях, когда $\psi(\cdot)$ или $P(\cdot)$ заданы на множествах, которые получаются из X с помощью операций взятия множества всех подмножеств и декартова произведения множеств. Например, если ψ — функция из множества $(P(X))^k \times X^n$ в множество Y, то ψ называется G-инвариантом, если

$$\forall w_1, \ldots, w_k \in P(X)\ \forall x_1, \ldots, x_n \in X (\psi(w_1, \ldots, w_k, x_1, \ldots, x_n) = \psi(f_g(w_1), \ldots, f_g(x_n))).$$

Например, $\psi(w, x)$ — расстояние между точкой x и замкнутым множеством w. Здесь $w \subseteq \boldsymbol{R}^2$, $x \in \boldsymbol{R}^2$, т. е. $\psi(w, x) \rightleftarrows \inf\limits_{z \in w} (z, x)$.

Следующий пример: $P(w_1, w_2)$ — множества w_1 и w_2 пересекаются. Здесь $w_1, w_2 \subseteq \boldsymbol{R}^2$. Очевидно, функция $\psi(\cdot, \cdot)$ является $O_2 \times \boldsymbol{R}^2$-инвариантом, предикат $P(\cdot, \cdot)$ является M_2-инвариантом.

Если фиксировано некоторое действие F группы G на множестве X, то будем называть G-геометрией совокупность всех G-инвариантов всевозможных указанных выше видов. Если G есть матричная группа или группа вида $G \times (\boldsymbol{R}^2)^+$, где G — матричная группа, то ее действие в $\boldsymbol{R}^2$ раз и навсегда фиксировано. Оно состоит просто в умножении точки x из $\boldsymbol{R}^2$ или каждой точки x из фигуры w на матрицу A из G (или соответственно в преобразовании вида $A \cdot x + \xi$ точки или фигуры). Поэтому для таких групп можно не указывать, о каком действии идет речь; всегда подразумевается указанное действие.

Знаменитое клейновское понимание геометрии состоит в том, что предметом геометрии является изучение всевозможных G-геометрий относительно различных групп G, их действий F и множеств X.

В случае школьной планиметрии в качестве множества X выступает арифметическая плоскость $\boldsymbol{R}^2$, или векторная плоскость $\mathscr{X}$, или аффинная плоскость A_2. Если рассматривается векторная или аффинная плоскость, то обычно подразумевается модель аксиоматики векторной плоскости — арифметическая плоскость $\boldsymbol{R}^2$, в которой не выделен никакой базис, или модель аксиоматики аффинной плоскости вида $\langle R^2, \boldsymbol{R}^2, g \rangle$, где первое R^2 — множество, второе $\boldsymbol{R}^2$ — арифметическая или векторная плоскость и $g(a, b) \rightleftarrows b - a$. Другими словами, в случае векторной плоскости в $\boldsymbol{R}^2$ «забывают» о стандартном базисе, а в случае аффинной плоскости в $\boldsymbol{R}^2$ «забывают» еще и о начале координат.

В качестве группы G в школьной планиметрии выбирается одна из групп биекций в $\boldsymbol{R}^2$, или в векторной, или в аффинной плоскости, о которых подробно говорилось в предыдущем параграфе. Вопрос о выборе действия этих групп не возникает: каждая из них действует в той плоскости, из биекции которой она состоит. Действие сводится просто к вычислению значения какой-то из этих биекций.

В соответствии с клейновским пониманием геометрии полное изучение школьной планиметрии состоит в описании всех инвариантов этих групп.

Реализуем клейнову программу для случая группы $G = O_2$ и арифметической плоскости $\boldsymbol{R}^2$.

2. Евклидова планиметрия — планиметрия ортогональной группы. Пусть $(\cdot, \cdot)$ — обычное скалярное произведение в $\boldsymbol{R}^2$. Рассмотрим функцию $\|\cdot\|^2 : \boldsymbol{R}^2 \to \boldsymbol{R}_{\geqslant 0}$, определяемую равенством $\|x\|^2 \rightleftarrows (x, x)$. Так как $\|x\|^2 \rightleftarrows (x, x) = x_1^2 + x_2^2$, то функция $\|\cdot\|^2$ является многочленом над полем $\boldsymbol{R}$ относительно двух координат x_1 и x_2 вектора x. При этом функция $\|\cdot\|^2$ является инвариантом группы O_2, так как согласно одному из определений O_2 выполняется соотношение $\forall A \in O_2\ \forall x \in \boldsymbol{R}^2 (\|x\|^2 = \|A \cdot x\|^2)$. Поэтому она является алгебраическим (второй степени) O_2-инвариантом от одной векторной переменной x или от двух вещественных переменных x_1 и x_2.

Нетрудно построить другие примеры алгебраических O_2-инвариантов, используя функцию $\|\cdot\|^2$. Например, таковыми являются следующие функции:

$$x \mapsto (x, x)^2 - 8 \cdot (x, x) + 1,$$
$$x \mapsto (x, x)^{1983} + (x, x)^{1001} + 2.$$

Вообще, любая функция вида $x \mapsto P_k((x, x))$, где $x \in \boldsymbol{R}^2$ и P_k — произвольный многочлен над полем $\boldsymbol{R}$ какой-то k-ой степени от одной числовой переменной, является O_2-инвариантом. Оказывается, верно и обратное утверждение: всякий алгебраический над вещественным полем O_2-инвариант от одной векторной переменной x равен композиции многочлена над полем $\boldsymbol{R}$ от одной переменной и функции $\|x\|^2 \rightleftarrows (x, x)$, где $x \in \boldsymbol{R}^2$.

Предложение 10. ***Пусть $f : \boldsymbol{R}^2 \to \boldsymbol{R}$ — многочлен (над полем $\boldsymbol{R}$) от двух вещественных переменных x_1 и x_2 (и следовательно, от одной векторной переменной $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$), который O_2-инвариантен, т. е.***

$$\forall A \in O_2\ \forall x \in \boldsymbol{R}^2 (f(x) = f(A \cdot x)).$$

Тогда найдется такой многочлен (над тем же полем R) $P(\alpha)$ от одной вещественной переменной α, что $\forall x \in R^2$ $(f(x) = P((x, x)))$.

▷ Пусть $f(x_1, x_2)$ — многочлен от двух вещественных переменных и $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in R^2$. Фиксируем какой-то один вектор $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Найдется такой поворот $W \in O_2$, который переводит точку $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ в точку $\begin{pmatrix} \|x\| \\ 0 \end{pmatrix}$. Тогда по условию O_2-инвариантности функции $f(\cdot)$ выполняется $f(x_1, x_2) = f(\|x\|, 0)$. Если в многочлен $f(\cdot, \cdot)$ всюду подставить вместо его второй переменной число нуль и затем выполнить арифметические действия, то получится некоторый многочлен над полем R от одной переменной x_1. Этот многочлен обозначим $P(\cdot)$, он является искомым многочленом. Итак, $f(x_1, x_2) = P(\|x\|)$. Теперь рассмотрим поворот $A \in O_2$, переводящий точку $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ в точку $\begin{pmatrix} -\|x\| \\ 0 \end{pmatrix}$ и аналогично получим:

$$f(x_1, x_2) = f(-\|x\|, 0) = P(-\|x\|).$$

Учитывая предыдущее равенство, получаем $P(-\|x\|) = P(\|x\|)$. Вектор x пробегает всю плоскость R^2, поэтому значение функции $\|\cdot\|$ пробегает все множество $R_{\geqslant 0}$. Значит $\forall \alpha \in R_{\geqslant 0} (P(-\alpha) = P(\alpha))$. Следовательно, многочлен $P(\cdot)$ является четной функцией: последнее возможно только в том случае, если он ни разу не содержит свою переменную x_1 в нечетной степени. Действительно, соберем вместе все четные одночлены, входящие в P, их сумма является четной функцией. Вычтем эту функцию из четной функции $P(\cdot)$. Разность двух четных функций является четной функцией. Следовательно, эта разность является постоянной. Определим многочлен $Q(\cdot)$, который получается из многочлена $P(\cdot)$ понижением всех входящих в него степеней переменной x_1 вдвое, т. е. заменой x_1^n на $x_1^{n/2}$. Очевидно, $Q(x_1^2) = P(x_1)$. Итак, $f(x) = P(\|x\|) = P(\sqrt{x_1^2 + x_2^2}) = Q(x_1^2 + x_2^2) = Q((x, x)) = Q \circ (x, x)$. ■

Теперь обобщим предложение 10 на случай алгебраических O_2-инвариантов от двух векторных переменных x и y и соответственно от четырех вещественных переменных x_1, x_2, y_1, y_2, где $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Простейшими алгебраическими O_2-инвариантами от двух векторных переменных являются следующие функции:

$$\langle x, y \rangle \mapsto (x, x), \ \langle x, y \rangle \mapsto (x, y), \ \langle x, y \rangle \mapsto (y, y),$$

где $x, y \in R^2$. Оказывается, что любой другой алгебраический (над полем R) O_2-инвариант от двух векторных переменных $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ и $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ выражается через эти простейшие инварианты

с помощью подстановки их в некоторый многочлен над полем R от трех переменных α_1, α_2, α_3.

Теорема 8. ***Пусть $f: R^4 \to R$ — многочлен (над полем R) от четырех вещественных переменных x_1, x_2, y_1, y_2, который O_2-инвариантен, т. е.***

$\forall A \in O_2 \ \forall x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in R^2 (f(x, y) = f(A \cdot x, A \cdot y))$. ***Тогда найдется такой многочлен $P(\alpha, \beta, \gamma)$ (над полем R) от трех переменных α, β, γ, что $\forall x, y \in R^2 (f(x, y) = P((x, x), (x, y), (y, y))) = P \circ \langle (x, x), (x, y), (y, y) \rangle$.***

▷ Фиксируем пару векторов $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ и $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ в плоскости R^2. Рассмотрим поворот $W \in O_2$, переводящий вектор x в вектор $\begin{pmatrix} \|x\| \\ 0 \end{pmatrix}$. При этом $W \cdot y$ есть какой-то вектор y' с координатами y_1' и y_2'. Тогда $(x, y) = (W \cdot x, W \cdot y) = \|x\| \cdot y_1'$ и $\|y\|^2 = (y, y) = (Wy, Wy) = (y_1')^2 + (y_2')^2$.

Следовательно, $y_1' = \dfrac{(x, y)}{\|x\|}$ и

$$y_2' = \sqrt{\|y\|^2 - \frac{(x, y)^2}{\|x\|^2}} = \frac{\sqrt{\|x\|^2 \cdot \|y\|^2 - (x, y)^2}}{\|x\|}.$$

Можно считать, что $y_2' \geqslant 0$, так как в противном случае рассмотрим композицию преобразования W и симметрии относительно оси Ox. При этом вектор $\begin{pmatrix} y_1' \\ y_2' \end{pmatrix}$ перейдет в вектор $\begin{pmatrix} y_1' \\ -y_2' \end{pmatrix}$ и вектор $\begin{pmatrix} \|x\| \\ 0 \end{pmatrix}$ останется на месте.

В силу O_2-инвариантности функции $f(\cdot)$ получаем:

$$f(x, y) = f(W \cdot x, W \cdot y) = f\left(\begin{pmatrix} \|x\| \\ 0 \end{pmatrix}, \begin{pmatrix} y_1' \\ y_2' \end{pmatrix}\right) = F\left(\sqrt{(x, x)}, \right.$$

$$\left. \frac{(x, y)}{\sqrt{(x, x)}}, \sqrt{\frac{(x, x)(y, y) - (x, y)^2}{(x, x)}}\right), \qquad (*)$$

где $F(t_1, t_2, t_3)$ — многочлен, который получается из многочлена $f(x_1, x_2, x_3, x_4)$ аннулированием переменной x_2. Оставшиеся три вещественные переменные переобозначим t_1, t_2, t_3, т. е. t_1 — это переменная x_1, t_2 — переменная x_3, t_3 — переменная x_4.

Итак, мы уже выразили функцию $f(x, y)$ через функции (x, x), (x, y), (y, y), но пока что с использованием нежелательных операций деления и извлечения квадратного корня.

Если рассмотреть композицию преобразования W и симметрии относительно оси Oy, то в силу O_2-инвариантности функции $f(\cdot)$ получим, что значения многочлена $F(\cdot)$ не меняются при одновременном изменении знака у переменных t_1 и t_2, т. е.

$$F(-t_1, -t_2, t_3) = F(t_1, t_2, t_3).$$

По аналогичной причине выполняется $F(t_1, t_2, -t_3) = F(t_1, t_2, t_3)$.

Многочлен $F(\cdot)$ является линейной комбинацией одночленов вида $C \cdot t_1^n \cdot t_2^m \cdot t_3^k$. Из равенства $F(t_1, t_2, -t_3) = F(t_1, t_2, t_3)$ следует, что k должно быть четным числом для всех одночленов в многочлене $F(\cdot)$.

Действительно, фиксируя $t_1 = t_2 = 1$, получим многочлен $F(1, 1, t_3)$ от одной переменной, который является четной функцией и поэтому содержит четные степени k. В свою очередь из равенства $F(-t_1, -t_2, t_3) = F(t_1, t_2, t_3)$ следует, что для любого одночлена $C \cdot t_1^n \cdot t_2^m \cdot t_3^k$ многочлена $F(\cdot)$ число $n+m$ четно, т. е. числа n и m одновременно четны или нечетны.

Действительно, полагая $t_1 = t_2 = t$ и $t_3 = 1$, получим многочлен $F(t, t, 1)$ от одной переменной t. В каждый одночлен этого многочлена переменная t входит со степенью $(n+m)$, и сам многочлен четный.

Итак, для каждого одночлена $C \cdot t_1^n \cdot t_2^m \cdot t_3^k$ возможны два случая: либо все числа n, m, k четны, т. е.

$$t_1^n \cdot t_2^m \cdot t_3^k = (t_1^2)^{\frac{n}{2}} (t_2)^{\frac{m}{2}} (t_3^2)^{\frac{k}{2}},$$

либо числа n, m нечетны и число k четно, т. е.

$$t_1^n \cdot t_2^m \cdot t_3^k = (t_1 \cdot t_2)(t_1^2)^{n_1} (t_2^2)^{m_1} (t_3^2)^{k_1}, \text{ где } n_1, m_1, k_1 \in \boldsymbol{N}.$$

Возвращаясь к равенству $(*)$, получаем, что функция $f(x, y)$ является многочленом относительно функций (x, x), $\frac{(x, y)^2}{(x, x)}$, $\frac{(x, x)(y, y) - (x, y)^2}{(x, x)}$ и (x, y).

Приводя все одночлены в $F(\cdot)$ к общему знаменателю, получаем, что

$$f(x, y) \equiv \frac{F_1((x, x), (x, y), (y, y))}{(x, x)^l}, \qquad (**)$$

где $F_1(\cdot, \cdot, \cdot)$ — новый многочлен от трех вещественных переменных и $l \in \boldsymbol{N}$. Аналогичным образом можно доказать, что

$$f(x, y) \equiv \frac{G_1((x, x), (x, y), (y, y))}{(y, y)^s},$$

где $G(\cdot, \cdot, \cdot)$ — многочлен от трех вещественных переменных и $s \in \boldsymbol{N}$. Следовательно, разность вида $(y, y)^s \cdot F_1((x, x), (x, y), (y, y)) - (x, x)^l G_1((x, x), (x, y), (y, y))$ тождественно равна нулю при всех x и y из $\boldsymbol{R}^2$.

Рассмотрим многочлен $H(\cdot, \cdot, \cdot)$ от трех переменных:

$$H(\alpha, \beta, \gamma) \rightleftarrows \gamma^s \cdot F_1(\alpha, \beta, \gamma) - \alpha^l \cdot G_1(\alpha, \beta, \gamma).$$

Для любых чисел $\alpha \geqslant 0$, $\gamma \geqslant 0$, β, если $\alpha\beta \geqslant \gamma^2$, можно подобрать такие x и y из $\boldsymbol{R}^2$, что $\alpha = (x, x)$, $\beta = (x, y)$, $\gamma = (y, y)$. Поэтому область, в которой многочлен H тождественно равен нулю, содержит внутренние точки, т. е. многочлен H тождественно равен нулю при всех α, β, γ.

Пусть r — максимальная степень, с которой переменная α входит одновременно во все одночлены многочлена $F_1(\alpha, \beta, \gamma)$, т. е. $F_1 = \alpha^r \cdot P(\alpha, \beta, \gamma)$, где многочлен $P(\alpha, \beta, \gamma)$ не делится нацело на α в кольце многочленов $R[\alpha, \beta, \gamma]$. Если бы r было меньше l, то $H(\alpha, \beta, \gamma) = \alpha^r(\gamma^s \cdot P(\alpha, \beta, \gamma) - \alpha^{l-r} \cdot G_1(\alpha, \beta, \gamma)) \rightleftarrows \alpha^r \cdot H_1(\alpha, \beta, \gamma) \equiv 0$.

Многочлен H_1 обращается в нуль в области $\{(\alpha, \beta, \gamma) | \alpha \neq 0\}$, в которой имеются внутренние точки, и поэтому $H_1(\alpha, \beta, \gamma)$ тождественно равен нулю. Поэтому $P(\cdot, \cdot, \cdot)$ нацело делится на α в кольце многочленов $R[\alpha, \beta, \gamma]$. Полученное противоречие доказывает, что $r \geqslant l$.

Поэтому частное $\frac{F_1(\alpha, \beta, \gamma)}{\alpha^l}$ является многочленом относительно переменных α, β, γ.

Возвращаясь к равенству $(**)$, получаем, что $f(x, y)$ является многочленом относительно (x, x), (x, y), (y, y). ■

Алгебраические O_2-инварианты от любого числа векторных переменных также выражаются через скалярные произведения с помощью многочленов. Однако доказательство этого факта значительно сложнее приведенных здесь доказательств для случаев $n = 1, 2$.

Глава III

ИЗМЕРЕНИЕ ВЕЛИЧИН. ПЛОЩАДЬ И МЕРА ПЛОСКИХ ФИГУР

ВВЕДЕНИЕ

Связанные между собой понятия измерения и величины весьма сложны; их полное обсуждение далеко выходит за рамки нашего курса. В школьной математике имеют дело с такими частными случаями измерений, как измерения длины, площади и объема, и с таким частным случаем величин, как положительные скалярные величины.

Все же сделаем несколько замечаний, касающихся общей концепции измерения и величины.

Когда речь идет об измерении, то явно или неявно предполагаются известными ответы на следующие вопросы:

что измеряется?

каким образом измеряется?

что является результатом измерения?

Когда мы говорим об измерении, то способ измерения предполагается явно реализованным, например, в виде алгоритма или схемы каких-то вычислений, или в виде физического прибора, или сочетания прибора и некоторых вычислений. Кратко говоря, предполагается известной некоторая инструкция для измерения.

Инструкция для измерения задает (определяет) сюръективное отображение T из семейства всех измеряемых объектов U на семейство всех возможных результатов измерения R_T. Конечно, нет оснований думать, что произвольное отображение может быть реализовано некоторым измерением.

Пусть фиксировано какое-то измерение, т. е. в конечном счете фиксировано некоторое сюръективное отображение вида $T: U \to R_T$. Величиной относительно измерения T интуитивно хотелось бы назвать то общее у измеряемых объектов u и v, что обеспечивает одинаковый результат их измерения посредством инструкции (отображения) T. Более точно величиной относительно фиксированного измерения T будем называть всякое множество вида $T^{-1}(\lambda)$, где λ фиксировано и принадлежит R_T. Системой однотипных величин относительно фиксированного измерения T будем называть совокупность всех множеств вида $T^{-1}(\lambda)$, т. е. совокупность $\mathfrak{U} \rightleftarrows \{T^{-1}(\lambda) \mid \lambda \in R_T\}$. Это множество, очевидно, совпадает с фактор-множеством множества U по отношению эквивалентности $u \sim u \rightleftarrows (T(u) = T(v))$. Классы эквивалентности записываются в виде

$[u]$, где $u \in U$; поэтому будем записывать величины в виде $[u]$, где u — один из измеряемых объектов. Ясно, что $[u] = T^{-1}(T(u))$.

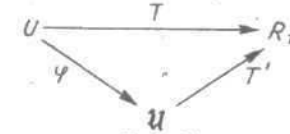


Рис. 44

Измеряемые объекты u и v будем называть равновеликими, если они принадлежат одной величине, т. е. если $[u] = [v]$. Конечно, последнее эквивалентно тому, что $T(u) = T(v)$.

Что представляет собой семейство R_T? Распространенная точка зрения, по-видимому, состоит в том, что R_T — некоторое подмножество множества $\boldsymbol{R}$, но тогда концепция измерения и величины с самого начала предполагает известным понятие вещественного числа. Последнее кажется нежелательным по многим причинам; например, концепции измерения и величины возникли и плодотворно изучались (скажем, во времена Евклида) задолго до возникновения точных определений вещественного числа и множества $\boldsymbol{R}$. Скорее наоборот, понятие вещественного числа сформировалось в результате изучения многих разных измерений и соответствующих им величин и на основе того подмеченного в математической практике наблюдения, что многие разные измерения приводят по существу к одинаковым системам величин; на современном языке можно сказать — приводят к системам величин, изоморфным R.

На систему однотипных величин $\mathfrak{U}$ «переносится» измерение T, первоначально определенное на множестве измеряемых объектов U. Точнее, измерением системы однотипных величин $\mathfrak{U}$ называется отображение $T': \mathfrak{U} \to R_T$, определяемое равенством $T'([u]) \rightleftarrows T(u)$. Тогда $T(u) = (T' \circ \varphi)(u)$, где $\varphi: U \to \mathfrak{U}$ — каноническое отображение множества U на фактор-множество $\mathfrak{U} \rightleftarrows U/\sim$, т. е. $\varphi(u) \rightleftarrows [u]$. Как всегда такое отображение T' называется опусканием отображения T на $\mathfrak{U}$ (рис. 44).

Отображение T' отличается от T тем, что T' биективно, и поэтому существует обратное отображение $(T')^{-1}: R_T \to \mathfrak{U}$. Итак, распространенное выражение «измерение величины объекта u» в точности означает вычисление композиции $T'(\varphi(u))$.

В конкретных ситуациях иногда удобно считать первоначальными понятия U и T и по ним определять понятия $\mathfrak{U}$ и T', а иногда наоборот — удобнее считать первоначальными $\mathfrak{U}$ и T' и по ним определять U и T.

Примерами измерений в физике являются измерения:

скорости, ускорения и массы материальной точки (в классической механике);

импульса, энергии и момента импульса частицы (в специальной теории относительности);

электрической и магнитной напряженности (в классической теории электромагнитного поля);

массы покоя, спина, заряда и времени жизни частицы (в теории элементарных частиц).

Вопросы о том, как фактически осуществлять эти измерения и да-

же как определять соответствующие этим измерениям объекты U, R_T и T, являются предметом глубоких физических и математических теорий.

В школьной математике традиционными являются измерения величин углов (относительно измерения углов), отрезков и кривых (относительно измерения длины), многоугольников, некоторых криволинейных фигур и поверхностей некоторых тел (относительно измерения площади), многогранников и тел вращения (относительно измерения объема) и т. п.

Итак, измерение — это некоторое отображение, система величин — некоторое фактор-множество (а величина — класс эквивалентности или иными словами — элемент соответствующего фактор-множества). Конечно, такая схема очень поверхностна, так как вся нетривиальная часть концепции измерения и величины состоит как раз в том, какие именно отображения являются измерениями (и как эти отображения задаются) и какие именно фактор-множества образуют системы величин. На этот счет мы ограничимся рядом примеров измерений и величин; они приводятся ниже в § 1. В § 2 приводятся такие примеры систем величин, для которых можно определить операцию сложения и отношение порядка, похожие по их свойствам на сложение и порядок в множестве вещественных чисел $\boldsymbol{R}$ (правда, операция и отношение в системе величин определяются с помощью самой алгебраической системы $\boldsymbol{R}$). Обобщая эти примеры, мы приводим в § 2 строгое, математическое определение частного случая понятия величины — определение положительной скалярной величины. Это определение аксиоматическое и не использует $\boldsymbol{R}$. Как и в случае любого аксиоматического определения, лучше приступать к его изучению после тщательного знакомства с примерами — моделями этого аксиоматического определения.

Возникает интересный вопрос, можно ли определить понятия величины, системы величин, операцию сложения величин, порядок между величинами, не используя измерение, которое в конкретных примерах с необходимостью апеллирует к понятию вещественного числа? Более того, если бы удалось определить систему величин $\mathfrak{U}$, обладающую рядом естественных свойств без использования измерения T, то, может быть, функция измерения этих величин T' возникла также без использования T как изоморфизм между $\mathfrak{U}$ и $\boldsymbol{R}$ вида $T':\mathfrak{U}\to\boldsymbol{R}$: тогда с помощью T' можно было бы определить само измерение T. Ответ на первый вопрос кратко намечается в примерах а), б) из § 2 (где действительно все определяется с помощью группы преобразований евклидовой плоскости вместо измерения T). Ответ на второй вопрос содержится в теореме 1.

После этого речь пойдет об измерении плоских фигур. В § 3, 4 напоминаются конструктивное и аксиоматическое определения площади многоугольника, доказывается в сущности совпадение этих определений, на этой основе устанавливается эквивалентность различных конкретных способов измерения площади многоугольника.

В § 5 продолжается изучение измерения плоских фигур, а именно, в нем рассматривается конкретное измерение (называемое мерой Лебега), которое применимо не только для всех многоугольников, являющихся, так сказать, прямолинейными фигурами, но и для очень многих криволинейных фигур, например для круга, внутренности эллипса, сектора, сегмента и т. п. Это новое измерение (обозначаемое M) также определяется конструктивно и аксиоматически и затем доказывается в сущности совпадение этих двух определений. Естественно, что для многоугольников мера и площадь совпадают, т. е. мера является продолжением функции площади с множества всех многоугольников на более широкое множество плоских фигур. Такое продолжение называется «мерой», термин «площадь» мы относим к случаю многоугольников.

§ 1. ПРИМЕРЫ ИЗМЕРЕНИЙ И ВЕЛИЧИН

а) Отрезок на арифметической прямой $\boldsymbol{R}$ можно определить как упорядоченную пару точек $\langle a, b\rangle$, где $a, b\in\boldsymbol{R}$ и $a\leqslant b$. Другое определение отрезка на $\boldsymbol{R}$ (или в $\boldsymbol{R}$) можно получить, сказав, что отрезком называется любое множество вида $[a, b]\rightleftarrows\{(1-\lambda)\cdot a+\lambda\cdot b \mid 0\leqslant\lambda\leqslant 1\}$, где $a, b\in\boldsymbol{R}$ и $a\leqslant b$. Мы будем пользоваться без специальных оговорок любым из этих двух эквивалентных определений. (Может быть, первое из них предпочтительнее в том отношении, что $\langle a, b\rangle$ — конечное множество, а $[a, b]$ — бесконечное множество, если $a\neq b$). Множество всех отрезков в смысле первого определения обозначим U.

Положим $T(u)\rightleftarrows b-a$, где $u=\langle a, b\rangle\in U$. Такое T назовем измерением длины отрезка. Тогда $R_T=\boldsymbol{R}_{\geqslant 0}$.

Можно в качестве множества измеряемых объектов выбрать $U\rightleftarrows\boldsymbol{R}^2$. Тогда положим $T(u)\rightleftarrows b-a$, где $u=\langle a, b\rangle\in\boldsymbol{R}^2$. Получим $R_T=\boldsymbol{R}$.

В соответствии с общим определением, величиной отрезка (относительно измерения длины) называется совокупность всех отрезков, имеющих одинаковую длину. Ясно, что $T(\langle a, b\rangle)=T(\langle c, d\rangle)\Leftrightarrow(\exists h\in R^+((a+h=c)\wedge(b+h=d)))$. Поэтому оказывается (!), что величину отрезка можно определить и не используя самого измерения T. А именно, величиной отрезка $\langle a, b\rangle$ называется класс эквивалентности

$$[\langle a, b\rangle]\rightleftarrows\{\langle c, d\rangle\in\boldsymbol{R}^2 \mid \exists h\in\boldsymbol{R}^+(\langle a, b\rangle+h=\langle c, d\rangle)\},$$

где $a, b\in\boldsymbol{R}$ и $a\leqslant b$. В этом определении вместо отображения T используется группа $\boldsymbol{R}^+$ (точнее — действие группы $\boldsymbol{R}^+$ в множестве $\boldsymbol{R}$, определяемое отображением $F(h)\rightleftarrows f_h$, где $h\in\boldsymbol{R}^+$ и $f_h(x)\rightleftarrows x+h$; здесь $F:\boldsymbol{R}\to\mathrm{Is}(\boldsymbol{R})$, $f_h:\boldsymbol{R}\to\boldsymbol{R}$). Система величин $\mathfrak{A}$ состоит из всех классов эквивалентности $[\langle a, b\rangle]$, т. е. совпадает с фактор-множеством множества отрезков относительно отношения эквивалентности:

$$(\langle a, b\rangle\sim\langle c, d\rangle)\rightleftarrows\exists h\in\boldsymbol{R}^+(\langle a, b\rangle+h=\langle c, d\rangle).$$

Здесь, конечно, $(\langle a, b\rangle + h) \rightleftarrows \langle a+h, b+h\rangle$, поэтому условие $\langle a, b\rangle + h = \langle c, d\rangle$ эквивалентно условию $\begin{cases} a+h=c, \\ b+h=d. \end{cases}$

б) Напомним, что аффинной евклидовой прямой $A^{(1)}$ называется структура вида $\langle A, \mathscr{X}^{(1)}, f, g\rangle$, где A — множество, любым двум точкам a и b которого сопоставляется вектор $f(a, b) \rightleftarrows \overline{a, b}$ из одномерного векторного пространства $\mathscr{X}^{(1)}$ таким образом, что одноместная функция (при фиксированном a) $f(a, \cdot): A \to \mathscr{X}^{(1)}$ является биекцией и $\overline{a, b} + \overline{b, c} = \overline{c, a} = 0$. При этом в $\mathscr{X}^{(1)}$ фиксирована некоторая симметричная, билинейная, положительно определенная форма $g(\cdot, \cdot)$. (Сравните это определение с определением аффинной евклидовой плоскости в главе II, § 2, п. 2—3.)

Пусть $U = A^2$ и $T(\langle a, b\rangle) \rightleftarrows \sqrt{g(\overline{a, b}, \overline{a, b})}$. Легко доказать, что при фиксированном векторе $\bar{e}$ из $\mathscr{X}^{(1)}$ выполняется $T(\langle a, b\rangle) = c \cdot |\lambda|$, где c некоторая постоянная, зависящая от $\bar{e}$, и $\overline{a, b} = \lambda \cdot e$.

Такое T называют измерением длины отрезка в аффинной евклидовой прямой.

в) Этот пример предполагает, что читатель знаком с понятиями кривой и ориентации кривой. При первом чтении его можно пропустить. Если на кривой α в $\boldsymbol{R}^2$ задана некоторая ориентация σ, то α в паре с этой ориентацией называется ориентированной кривой и обозначается $\langle \alpha, \sigma\rangle$. Пусть U — множество всех ориентированных кривых в R^2.

Определим $T(\langle \alpha, \sigma\rangle) \rightleftarrows (\pm \int_a^b \sqrt{(x_t')^2 + (y_t')^2}\, dt)$; здесь $\langle \alpha, \sigma\rangle \in U$, т. е. $\alpha(t) = \begin{cases} x = x(t), \\ y = y(t), \end{cases} a \leqslant t \leqslant b$ — кривая и, следовательно, по определению $x(\cdot)$ и $y(\cdot)$ — непрерывно дифференцируемые функции вида $[a, b] \to \boldsymbol{R}$, а σ — ее ориентация. Перед интегралом по определению ставится знак $+$ в том случае, если ориентация σ совпадает с ориентацией, определяемой на кривой α отображением $(x'(\cdot), y'(\cdot)): [a, b] \to R^2$ (и стандартной ориентацией $\boldsymbol{R}$) и ставится знак — в противном случае.

Тогда T называется измерением длины ориентированной кривой и $R_T = \boldsymbol{R} \cup \{-\infty; +\infty\}$. Для такого измерения каждая величина состоит из весьма разнообразных объектов.

г) Пусть U — множество всех параллельных переносов в арифметической плоскости $\boldsymbol{R}^2$, т. е. множество всех отображений вида $f_\xi: \boldsymbol{R}^2 \to \boldsymbol{R}^2$, $f_\xi: x \mapsto x + \xi$, где ξ какой-то фиксированный элемент из $\boldsymbol{R}^2$. Интуитивно «размер» параллельного переноса можно измерить вектором $\xi = f_\xi(0)$. Определим $T: U \to \boldsymbol{R}^2$ равенством $T(f) \rightleftarrows f(0)$. Тогда $R_T = \boldsymbol{R}^2$. Отметим, что в данном случае отображение T биективно и величина параллельного переноса f является классом эквивалентности $[f]$, состоящим ровно из одного объекта f.

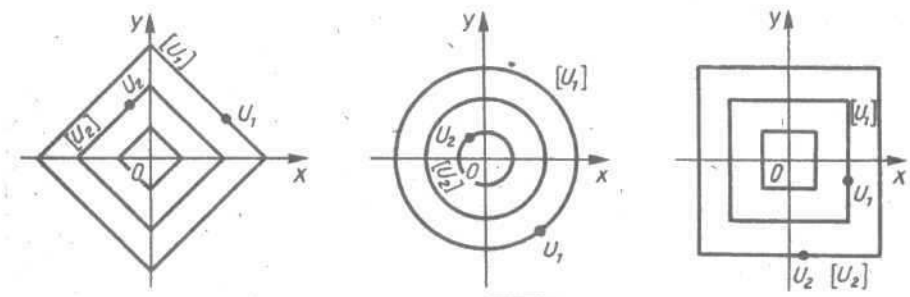


Рис. 45

д) Пусть U — множество углов в $\boldsymbol{R}^2$ с вершиной в точке $\overline{0} \rightleftarrows (0, 0)$, т. е $U = \mathrm{An}$. Определим значение $T(u)$, где $u \in U$, как главную меру угла u по основанию 2π. Тогда $R_T = [0, 2\pi[$.

е) Пусть $U = \boldsymbol{R}^2$ и $\|\cdot\|$ — некоторая норма в $\boldsymbol{R}^2$, т. е. $\|\cdot\|$ произвольная функция из $\boldsymbol{R}^2$ в $\boldsymbol{R}_{\geqslant 0}$ со свойствами:

1) $\|u\| = 0 \Leftrightarrow u = 0$,
2) $\|\lambda \cdot u\| = |\lambda| \cdot \|u\|$,
3) $\|u + v\| \leqslant \|u\| + \|v\|$.

Положим $T(u) \rightleftarrows \|u\|$. Тогда величина вектора u является классом эквивалентности $[u] = T^{-1}(T(u)) = T^{-1}(\|u\|)$. Он состоит из всех векторов v из $\boldsymbol{R}^2$, которые имеют одну и ту же норму, равную норме вектора u. Поэтому $[u]$ — «сфера» радиуса $\|u\|$ с центром в точке $\overline{0}$, т. е. $[u] = \{v \in R^2 \mid \|v\| = \|u\|\}$. Семейство $\mathfrak{U}$ всех величин относительно такого измерения состоит из всех гомотетичных «сфер», т. е. две такие сферы получаются друг из друга умножением на некоторое фиксированное число $\lambda > 0$ (рис. 45).

ж) Пусть U — семейство всех интегрируемых (по Лебегу) функций, определенных на отрезке $[0, 1]$. Определим $T: U \to \boldsymbol{R}$ равенством $T([u]) \rightleftarrows \int_0^1 u(x)dx$, где $u(\cdot) \in U$. Тогда $R_T = \boldsymbol{R}$ и величина $[u]$ состоит из всех функций, отличающихся от данной функции u на множестве меры нуль. В этом случае семейство всех величин совпадает с пространством $L([0; 1])$, которое изучается в курсе анализа.

з) Пусть U — множество всех событий в некотором пространстве элементарных событий и $T(u)$ по определению равно вероятности события $u \in U$. Тогда $R_T = [0, 1]$ и $\mathfrak{U}$ состоит из классов равновероятных событий.

§ 2. ПОЛОЖИТЕЛЬНАЯ СКАЛЯРНАЯ ВЕЛИЧИНА

Напомним, что системой (однотипных) величин относительно фиксированного измерения $T: U \to R_T$ называется фактор-множество $\mathfrak{U}$ в множестве U по следующему отношению эквивалентности: $u \sim v \rightleftarrows T(u) = T(v)$. Множество $\mathfrak{U}$ можно определить и как $\mathfrak{U} \rightleftarrows \{T^{-1}(\lambda) \mid \lambda \in R_T\}$. Элементы такого множества, т. е. отдельные величины, будем обозначать буквой w с индексами или без них. В каждом из нижеследующих примеров будет показано, что в системе величин можно определить операцию сложения $+$ и отношение порядка $\leqslant$ (что, в частности, оправдывает термин «система»), относительно которых это семейство изоморфно алгебраической системе $\langle \boldsymbol{R}_{\geqslant 0}, +, \leqslant \rangle$ или ее подсистеме некоторого одного и того же простого вида. Эти примеры приведут нас к аксиоматическому описанию одного класса семейств величин. Входящие в этот класс величины называются положительными скалярными величинами. После этого будет доказано, что соответствующая аксиоматика имеет модели только одного, простого вида, связанные с системой $\boldsymbol{R}$.

а) Рассмотрим измерение длин отрезков, т. е. функцию $T(\langle a, b \rangle) \rightleftarrows b - a$, где $a, b \in \boldsymbol{R}$ и $a \leqslant b$. Величиной длины отрезка называется любое множество вида $[\langle a, b \rangle] \rightleftarrows \{\langle c, d \rangle \mid c, d \in \boldsymbol{R} \wedge c \leqslant d \wedge \exists h \in \boldsymbol{R}^+ (\langle c, d \rangle + h = \langle a, b \rangle)\}$, где $a, b \in \boldsymbol{R}$ и $a \leqslant b$.

В соответствии с общим определением все такие множества образуют семейство величин $\mathfrak{U}$ относительно данного измерения T. Отображение T опускается на фактор-множество $\mathfrak{U}$ (см. введение) и получается биекция вида $T': \mathfrak{U} \leftrightarrow R_{\geqslant 0}$. С помощью этой биекции и системы $\boldsymbol{R}_{\geqslant 0}$ определим операцию сложения $+$, порядок $\leqslant$ и элемент 0 в множестве $\mathfrak{U}$: $(w_1 + w_2) \rightleftarrows (T')^{-1}(T'(w_1) + T'(w_2))$, $0_{\mathfrak{U}} \rightleftarrows (T')^{-1}(0)$, где 0 — число из $\boldsymbol{R}$.

Получилась алгебраическая система $\langle \mathfrak{U}, +, \leqslant, 0_{\mathfrak{U}} \rangle$, которая, очевидно, изоморфна системе $\langle R_{\geqslant 0}, +, \leqslant, 0 \rangle$. Изоморфизм определяется биекцией $T': \mathfrak{U} \leftrightarrow R_{\geqslant 0}$.

С помощью группы $\boldsymbol{R}^+$ (точнее — с помощью действия этой группы в множестве $\boldsymbol{R}$) можно определить не только семейство величин $\mathfrak{U}$, но и структуру в $\mathfrak{U}$. А именно, суммой двух величин $[\langle a, b \rangle]$ и $[\langle c, d \rangle]$ назовем величину $[\langle a, d' \rangle]$, в которой $d' \rightleftarrows f_h(d)$ и f_h определяется условием $f_h(c) = b$. Попросту говоря, в группе $\boldsymbol{R}^+$ находится такое h, что $b = c + h$, и затем полагают сумму величин, равной классу эквивалентности с представителем $\langle a, d + h \rangle$ (рис. 46).

Нулем в $\mathfrak{U}$ назовем такую величину, которая содержит отрезок $\langle a, a \rangle$, где $a \in \boldsymbol{R}$.

Аналогично положим $[\langle a, b \rangle] \leqslant [\langle c, d \rangle]$, если $d' \leqslant b$, где $d' \rightleftarrows f_h(d)$ и f_h определяется условием $f_h(c) = b$.

Упражнение 1. *Проверьте, что операции сложения и порядки в $\mathfrak{U}$, определяемые с помощью биекции T' и группы $\boldsymbol{R}^+$, соответственно совпадают.*

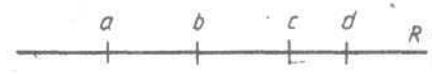


Рис. 46

Итак, операцию и порядок в $\mathfrak{U}$ можно определить с помощью одних только параллельных переносов на прямой $\boldsymbol{R}$, как это часто делается в школе (где прямая $\boldsymbol{R}$ заменяется интуитивным представлением об «евклидовой прямой»). По-видимому, в таком же духе понимали операции над величинами, «возникающими из отрезков», и в те времена, когда не существовало понятие вещественного числа.

Пусть X теоретико-множественная модель прямой в аксиоматике Евклида — Гильберта, т. е. X — множество, в котором выполняются все аксиомы Евклида — Гильберта для случая прямой. Пусть фиксирована группа G, которая транзитивно действует в множестве X, т. е. $F: G \to \mathrm{Is}(X)$ и $F(h) \rightleftarrows f_h: X \to X$ и выполняется $\forall x, y \in X \exists h \in G (f_h(x) = y)$.

Точно так же, как и для случая $G = \boldsymbol{R}^+$ и $X = \boldsymbol{R}$, можно определить фактор-множество множества всех отрезков в X (т. е. определить семейство величин $\mathfrak{A}$) и определить операцию сложения в $\mathfrak{A}$, элемент $0_{\mathfrak{A}}$ и порядок в $\mathfrak{A}$. Эти определения вместо измерения T и множества $\boldsymbol{R}$ используют группу преобразований $F(G)$ прямой. Получилась алгебраическая система $\langle \mathfrak{A}, +, \leqslant, 0_{\mathfrak{A}} \rangle$. Если наложить условие на группу G, состоящее в том, что определяемая по этой группе система $\langle \mathfrak{A}, +, \leqslant, 0_{\mathfrak{A}} \rangle$ удовлетворяет определению 1 (прочтите его, полагая, что $V = \mathrm{I}$, и, следовательно, множество I в этом определении можно опустить), то по теореме 1 (прочтите ее) возникает сюръективное отображение вида $T': \mathfrak{U} \to \boldsymbol{R}_{\geqslant 0}$. Оно служит измерением системы величин $\mathfrak{U}$. В этом смысле группа преобразований $F(G)$ определяет не только систему величин $\mathfrak{U}$ (включая структуру $+$, $\leqslant$, $0_{\mathfrak{U}}$ в ней), но и само измерение этих величин. По T' легко определяется и измерение T самих отрезков в X.

б) Рассмотрим двумерный аналог примера а). А именно, пусть U — множество всех треугольников в арифметической плоскости $\boldsymbol{R}^2$. Напомним, что треугольником называется множество, состоящее из трех точек, не лежащих на одной прямой. Эти три точки называются вершинами треугольника. Другое определение треугольника состоит в том, что треугольником называется любое множество вида $\{\lambda_1 \cdot a + \lambda_2 \cdot b + \lambda_3 \cdot c \mid \lambda_1, \lambda_2, \lambda_3 \geqslant 0, \lambda_1 + \lambda_2 + \lambda_3 = 1\}$, где $\{a, b, c\}$ — произвольное множество из трех точек, не лежащих на одной прямой. Как и в случае отрезка, мы будем пользоваться любым из этих двух эквивалентных определений без специальных оговорок.

Упражнение 2*. *Докажите, что множество $\{\lambda_1 \cdot a + \lambda_2 \cdot b + \lambda_3 \cdot c \mid \lambda_1, \lambda_2, \lambda_3 \geqslant 0 \wedge \lambda_1 + \lambda_2 + \lambda_3 = 1\}$ совпадает с пересечением всех полуплоскостей в $\boldsymbol{R}^2$, содержащих три точки a, b, c.*

Последняя характеристика треугольника иногда также принимается за его определение.

Пусть $\triangle = \{a, b, c\}$ — треугольник. Положим, $T(\triangle) \rightleftarrows \frac{1}{2} |\det$

$\begin{pmatrix}1&1&1\\a&b&c\end{pmatrix}$|, где пары чисел, образующие координаты точек a, b, c, располагаются в столбцы и в круглых скобках находится матрица размера 3×3. Такая функция T называется измерением площади треугольника. При этом $R_T=R_{>0}$ и величина треугольника $\triangle$ является классом эквивалентности $[\triangle]$, состоящим из всех треугольников $\triangle_1$, у которых площадь равна площади треугольника $\triangle$. Возникает система величин $\mathfrak{U}$. Как и в примере 1, измерение T опускается на это семейство и возникает измерение величин T', которое является биекцией вида $T':\mathfrak{U}\leftrightarrow\boldsymbol{R}_{>0}$. С помощью этой биекции и системы $\boldsymbol{R}_{\geqslant0}$ определим операцию сложения $+$ и порядок $\leqslant$ в множество $\mathfrak{U}$.

Добавим к $\mathfrak{U}$ новый элемент, который обозначим $0_{\mathfrak{U}}$. Его можно представлять себе как множество всех отрезков в плоскости $\boldsymbol{R}^2$. Иными словами, к множеству U добавляются все отрезки в $\boldsymbol{R}^2$. Отрезком в $\boldsymbol{R}^2$, естественно, называется пара точек $\langle a, b\rangle$, где $a, b\in\boldsymbol{R}^2$. В частности, точки a и b могут совпадать, и в этом случае говорят о вырожденном отрезке; но в данном контексте удобнее иметь в виду только невырожденные отрезки. Невырожденные отрезки можно рассматривать как вырожденные треугольники. Итак, к U добавляются все вырожденные треугольники, измерение T по определению продолжается на них значением нуль, и они образуют новый класс эквивалентности $0_{\mathfrak{U}}$, входящий в $\mathfrak{U}$. В результате T' на классе $0_{\mathfrak{U}}$ имеет значение нуль и получается биекция вида $T':\mathfrak{U}\leftrightarrow R_{\geqslant0}$. По ней число 0 соответствует величине $0_{\mathfrak{U}}$.

Получилась алгебраическая система вида $\langle\mathfrak{U}, +, \leqslant, 0\rangle$, которая, очевидно, изоморфна системе $\langle\boldsymbol{R}_{\geqslant0}, +, \leqslant, 0\rangle$. Изоморфизм определяется биекцией $T':\mathfrak{U}\leftrightarrow\boldsymbol{R}_{\geqslant0}$.

Теперь мы хотим, как и в примере а), определить систему величин $\mathfrak{U}$, структуру в ней и измерение величин T', не используя самого измерения T. Как и в примере а), вместо измерения T мы используем группу преобразований плоскости $\boldsymbol{R}^2$ (затем $\boldsymbol{R}^2$ можно будет заменить «евклидовой плоскостью»). Сначала рассмотрим одно вспомогательное

Предложение 1. ***Для любой невырожденной матрицы A (т. е. любой матрицы A из GL_2) и любого треугольника $\triangle=\{a, b, c\}$ множество $A\cdot\triangle\rightleftarrows\{A\cdot a, A\cdot b, A\cdot c\}$ является треугольником.***

▷ Нужно доказать, что три точки $A\cdot a$, $A\cdot b$ и $A\cdot c$ не лежат на одной прямой. Допустим, что они лежат на одной прямой. Из курса алгебры известно, что отображение плоскости на себя, задаваемое невырожденной матрицей A, переводит различные ее точки в различные. Поэтому векторы $Aa-Ac=A\cdot(a-c)$ и $Ab-Ac=A\cdot(b-c)$ отличны от нуля, и так как они коллинеарны (рис. 47), то при некотором $\lambda\in\boldsymbol{R}$, $\lambda\neq0$ выполняется $Aa-Ac=\lambda(Ab-Ac)$. Следовательно, $A(a-c)=A(\lambda(b-c))$ и по той же теореме из курса алгебры получаем, что $a-c=\lambda(b-c)$. Поэтому векторы $a-c$ и $b-c$ коллинеарны, т. е. точки a, b, c лежат на одной прямой. Это противоречит условию предложения 1. ■

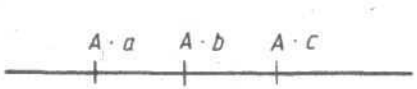


Рис. 47

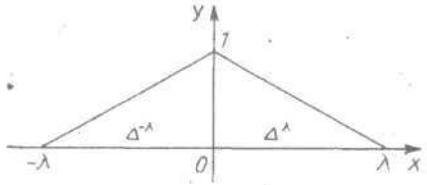


Рис. 48

Напомним, что множество L_2 состоит из всех матриц размера 2×2 с определителем, равным $+1$ или -1. Ясно, что функция $f_{\langle A,\xi\rangle}(x)\rightleftarrows A\cdot x+\xi$, где $A\in L_2$ и $\xi\in(\boldsymbol{R}^2)^+$, является биекцией вида $\boldsymbol{R}^2\leftrightarrow\boldsymbol{R}^2$ (иными словами — является преобразованием плоскости). Такие функции будем называть эквиаффинными. Множество всех эквиаффинных функций относительно операции композиции образует группу преобразований плоскости (проверьте). Обозначим ее G_s.

Упражнение 3. *Пусть Δ произвольный треугольник. Докажите, что:*

а) для любой матрицы A из L_2 выполняется $T(\Delta)=T(A\cdot\Delta)$, аналогично для любой пары $\langle A, \xi\rangle$ из $L_2\times\boldsymbol{R}^2$ выполняется $T(\Delta)=T(A\cdot\Delta+\xi)$, т. е. $T(\Delta)=T(f_{\langle A,\xi\rangle}(\Delta))$.

б) Для любой матрицы A из $(GL_2\setminus L_2)$ выполняется $T(\Delta)\neq T(A\cdot\Delta)$. Поэтому L_2 — наибольшее подмножество в GL_2 и $L_2\times\boldsymbol{R}^2$ наибольшее подмножество в $GL_2\times\boldsymbol{R}^2$, для которых соответствующие множества преобразования плоскости $\boldsymbol{R}^2$ сохраняют измерение T.

Теперь рассмотрим основное

Предложение 2. ***Треугольники Δ и Δ_1 имеют одинаковую площадь (т. е. $T(\Delta)=T(\Delta_1)$) в том и только в том случае, если $\exists A\in L_2\exists\xi\in\boldsymbol{R}^2(\Delta_1=A\cdot\Delta+\xi)$. Поэтому величина $[\Delta]$ равна $\{\Delta\exists\langle A, \xi\rangle\in(L_2\times\boldsymbol{R}^2)\ (\Delta_1=A\cdot\Delta+\xi)\}$.***

Обозначим Δ^λ треугольник с вершинами в точках

$$\left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}\lambda\\0\end{pmatrix}\right\}$$

Лемма 1. ***Для любого треугольника Δ существует пара $\langle A, \xi\rangle\in(L^2\times\boldsymbol{R}^2)$, для которой треугольник $A\cdot\Delta+\xi$ является треугольником вида Δ^λ и $T(\Delta)=T(\Delta^\lambda)$.***

Последнее равенство, очевидно, эквивалентно тому, что $\lambda=2\cdot T(\Delta)$, и тому, что $[\Delta]=[\Delta^\lambda]$. Треугольник Δ^λ назовем каноническим представителем величины $[\Delta]$ (рис. 48).

Доказательство предложения 2 легко вытекает из леммы 1: пусть Δ и Δ_1 два треугольника и треугольники $\Delta^\lambda=A_1\cdot\Delta+\xi_1$ и $\Delta^{\lambda_1}=A_2\cdot\Delta_1+\xi_2$ получаются по ней. Тогда $\Delta_1=A_2^{-1}\cdot(\Delta^{\lambda_1}-\xi_1)$. Если $T(\Delta)=T(\Delta_1)$, то $T(\Delta^\lambda)=T(\Delta^{\lambda_1})$ и $\Delta^\lambda=\Delta^{\lambda_1}$, и $\lambda=\lambda_1$. Поэтому $\Delta_1=A_2^{-1}(A_1\cdot\Delta+\xi_1-\xi_2)=(A_2^{-1}\cdot A_1)\cdot\Delta+(A_2^{-1}(\xi_1-\xi_2))$ и

$$A\rightleftarrows(A_2^{-1}\cdot A_1)\in L_2,\ \xi\rightleftarrows(A_2^{-1}\cdot(\xi_1-\xi_2))\in\boldsymbol{R}^2.$$

Наоборот, если существует такое преобразование $f_{\langle A,\,\xi\rangle}$: $\boldsymbol{R}^2 \to \boldsymbol{R}_2$, что $\Delta_1 = A\cdot\Delta + \xi$, то согласно упражнения 3а) выполняется $T(\Delta_1) = T(\Delta)$. ■

▷ Теперь докажем лемму 1. Положим $\xi \rightleftarrows -a$. Тогда треугольник $\Delta' \rightleftarrows \Delta + \xi$ имеет одну из вершин в точке $\overline{0}$. Две другие его вершины $b+\xi$ и $c+\xi$ обозначим соответственно b' и c'. По упражнению 3а) легко проверяется $T(\Delta') = T(\Delta)$. Рассмотрим треугольник Δ^λ, в котором $\lambda = 2\cdot T(\Delta') = 2\cdot T(\Delta)$. Очевидно, $T(\Delta) = T(\Delta^\lambda)$. Найдем матрицу B из L_2, для которой $B\cdot\Delta^\lambda = \Delta'$. Тогда $B^{-1}\cdot\Delta' = \Delta^\lambda$ и $B^{-1}\cdot(\Delta+\xi) = B^{-1}\cdot\Delta + B^{-1}\cdot\xi = \Delta^\lambda$, причем $B^{-1}\in L_2$. Это и требуется доказать в лемме.

Из условия $B\cdot\binom{0}{1} = b'$ вытекает, что первый столбец в матрице B равен b'. Из условия $B\binom{\lambda}{0} = B\lambda\binom{1}{0} = \lambda B\binom{1}{0} = c'$ вытекает $B\binom{1}{0} = \frac{1}{\lambda}c'$, т. е. второй столбец в B равен $\frac{1}{\lambda}c'$. Итак, $B = \left(b'\,\frac{1}{\lambda}c'\right)$. Осталось проверить, что $B\in L_2$.

Действительно, $\det B = \frac{1}{\lambda}\det(b', c') = \frac{1}{\lambda}\det\begin{pmatrix}1 & 1\\ 0 & b'\end{pmatrix}$ и $|\det B| = \frac{1}{\lambda}|\det\begin{pmatrix}1 & 1\\ 0 & b'\end{pmatrix}| = \frac{1}{\lambda}\cdot 2\cdot T(\Delta') = \frac{2T(\Delta')}{2T(\Delta)} = 1$. ■

Упражнение 4. *Докажите, что для любого треугольника Δ существует преобразование плоскости вида $f_{\langle A,\,\xi\rangle}$, где $A\in L_2$ и $\xi\in\boldsymbol{R}^2$, сохраняющее площадь треугольника, при котором он перейдет в треугольник вида $\Delta^{-\lambda}$, показанный на рисунке 48. Условие сохранения площади треугольника равносильно тому, что $[\Delta] = [\Delta^{-\lambda}] = [\Delta^\lambda]$.*

Итак, мы построили группу G_s преобразований плоскости $\boldsymbol{R}^2$, с помощью которой в сущности уже определили все то, о чем говорилось перед предложением 1. Действительно, величина определяется как множество треугольников вида

$$\{\Delta_1 \mid \exists\langle A, \xi\rangle \in (L_2\times\boldsymbol{R}^2)(\Delta_1 = A\cdot\Delta + \xi)\},$$

где Δ какой-то фиксированный треугольник. Множество всех таких величин обозначим $\mathfrak{U}$. По предложению 2 это то же самое множество $\mathfrak{U}$, которое определялось в начале этого примера с помощью измерения T. Сумму двух величин $[\Delta_1]$ и $[\Delta_2]$ определим как величину с представителем $\Delta \rightleftarrows \left\{\binom{0}{1}, \binom{\lambda_1}{0}, \binom{-\lambda_2}{0}\right\}$, где вершина $\binom{\lambda_1}{0}$ возникает из канонического представителя Δ^{λ_1} первой величины, а вершина $\binom{-\lambda_2}{0}$ возникает по упражнению 4 из представителя $\Delta^{-\lambda_2}$ второй величины.

По определению выполняется $[\Delta_1] \leqslant [\Delta_2]$, если $\lambda_1 \leqslant \lambda_2$, где точки λ_1 и λ_2 на прямой $\boldsymbol{R}$ возникают из канонических представителей Δ^{λ_1} и Δ^{λ_2} этих двух величин. Может быть, более эстетично не использовать здесь прямую $\boldsymbol{R}$, а исходить из второго определения треугольника. Тогда по определению выполняется $[\Delta_1] \leqslant [\Delta_2]$, если $\Delta^{\lambda_1} \subseteq \Delta^{\lambda_2}$, где Δ^{λ_1} и Δ^{λ_2} канонические представители этих величин. Ясно, что эти два определения порядка в $\mathfrak{U}$ эквивалентны.

Упражнение 5. *Докажите, что операции сложения и порядки в $\mathfrak{U}$, определяемые с помощью биекции T' и с помощью группы G_s, соответственно совпадают. (Сравните это упражнение с упражнением 1).*

Добавим к множеству $\mathfrak{U}$ элемент $0_{\mathfrak{U}}$ (его описание не зависит ни от измерения T, ни от группы G_s) и получим ту же, что и раньше, алгебраическую систему $\langle\mathfrak{U}, +, \leqslant, 0_{\mathfrak{U}}\rangle$, определенную с помощью группы G_s без использования самого измерения T. Конечно, и второе определение этой системы величин использует понятие вещественного числа тем косвенным образом, что треугольники рассматриваются в арифметической плоскости $\boldsymbol{R}^2$ и определение группы G_s использует представление преобразований плоскости матрицами. (Последнее предполагает фиксирование базиса в плоскости, а векторная плоскость с фиксированным базисом канонически изоморфна $\boldsymbol{R}^2$ и в этом смысле неотличима от $\boldsymbol{R}^2$; впрочем, понятия векторной и даже аффинной плоскости используют $\boldsymbol{R}$ в качестве поля скаляров.)

Если мы хотим определить эту систему величин, не используя понятия вещественного числа, то нужно идти по пути, указанному в конце примера а). А именно, пусть теперь X — теоретико-множественная модель аксиоматики Евклида — Гильберта для случая плоскости, т. е. X — множество, в котором выполняются все соответствующие аксиомы. Пусть фиксирована группа G, которая действует в множестве X, т. е. $F: G\to \mathrm{Is}\,X$ и $F(h) \rightleftarrows f_h: X\to X$, где $h\in G$. Точно так же, как и для случая $G = G_s$ и $X = \boldsymbol{R}^2$, можно определить фактор-множество множества всех треугольников в X (т. е. определить семейство величин $\mathfrak{U}$). Чтобы определить в $\mathfrak{U}$ операцию сложения и порядок, уже нужны условия на группу G такого типа: для любых треугольников Δ_1 и Δ_2 существуют эквивалентные им (относительно G) треугольники Δ^{λ_1} и $\Delta^{-\lambda_2}$, которые имеют одну общую сторону и по одной стороне на общей прямой по разные стороны от общей вершины этих сторон. Иными словами, нужно, чтобы любые две величины $[\Delta_1]$ и $[\Delta_2]$ содержали треугольники, расположенные друг относительно друга так, как треугольники Δ^{λ_1} и $\Delta^{-\lambda_2}$, а также, как треугольники Δ^{λ_1} и Δ^{λ_2}. Тогда в $\mathfrak{U}$ определяется, как и раньше, алгебраическая система вида $\langle\mathfrak{U}, +, \leqslant, 0_{\mathfrak{U}}\rangle$. Некоторые дополнительные условия на G обеспечивают, что эта система удовлетворяет определению 1, и тогда по теореме 1 возникает биекция вида $T': \mathfrak{U}\leftrightarrow R_{\geqslant 0}$. Она является измерением величин, входящих в $\mathfrak{U}$, и определяет измерение T треугольников в X.

в) Рассмотрим только такие отрезки из примера а), которые

содержатся в одном фиксированном отрезке $\langle a_0, b_0\rangle$. Измерение T из примера а) сузим на множество U_1 всех таких отрезков. Это сужение также является измерением, которое мы обозначим T_1. Система $\mathfrak{U}_1$ величин относительно нового измерения T_1 является частью старой системы $\mathfrak{U}$ величин относительно старого измерения T. Такой переход от системы величин $\mathfrak{U}$ к системе величин $\mathfrak{U}_1$ можно пояснить, например, следующим образом: величина длины отрезка связана с реальным измерением расстояния «от пункта a до пункта b». Если $b-a$ больше 10 километров, то нужно учитывать кривизну поверхности Земли и использовать особые методы измерения длины, отличные от случая, когда $b-a$ меньше 10 километров.

Отношение порядка в $\mathfrak{U}_1$ определяется так же, как и в $\mathfrak{U}$, т. е. порядок в $\mathfrak{U}_1$ является сужением порядка в $\mathfrak{U}$. Однако операцию сложения в $\mathfrak{U}_1$ нельзя определить как сужение операции сложения в $\mathfrak{U}$, так как сумма величин $[\langle a, b\rangle]+[\langle a, c\rangle]$ может не принадлежать $\mathfrak{U}_1$. Поэтому мы вынуждены определить операцию сложения только для части величин: для величин, которые содержат представитель вида $\langle a_0, b\rangle$ и $2\cdot(b-a_0)\leqslant(b_0-a_0)$. Иными словами, обозначим

$$I \rightleftarrows \{w\in\mathfrak{U}_1 \mid \exists\langle a, b\rangle\in w([a, b]\subseteq\left[a_0, \frac{a_0+b_0}{2}\right])\}.$$

Для величин из I операция сложения определяется точно так же, как и для величин из $\mathfrak{U}$, т. е. как сужение операции $+$ из $\mathfrak{U}$ на множество I. При этом она не выводит за пределы $\mathfrak{U}_1$, т. е. $+ : I^2\to\mathfrak{U}_1$. В частности, может случиться, что $w_1, w_2\in I$ и w_1+w_2 не принадлежит I. В результате получилась система вида $\langle\mathfrak{U}_3, I, +, \leqslant, 0_{\mathfrak{U}}\rangle$. Она изоморфна алгебраической системе в $\boldsymbol{R}_{\geqslant 0}$ вида $\langle\boldsymbol{R}_{\geqslant 0}, [0, \lambda], +, \leqslant, 0\rangle$, где $\lambda=T_1\left(\left[a_0, \frac{a_0+b_0}{2}\right]\right)$. Изоморфизм задается биекцией $T_1' : \mathfrak{U}_1\leftrightarrow[0, \lambda]$, являющейся измерением величин из $\mathfrak{U}_1$.

В данном случае понятие «изоморфизм» означает, что биекция $f\rightleftarrows T_1'$ обладает свойствами:

1) $\forall w, w'\in\mathfrak{U}_1(w\leqslant w'\Leftrightarrow f(w)\leqslant f(w'))$;
2) $\forall w, w'\in I(w+w'\in I\Rightarrow f(w+w')=f(w)+f(w'))$;
3) $f(0_{\mathfrak{U}})=0$.

Такое, связанное с дополнительным носителем I понимание изоморфизма отличается от того, как мы понимали изоморфизм в примерах а) — б).

г) Рассмотрим только такие треугольники из примера б), которые являются подмножеством какого-либо фиксированного квадрата u_0. Для них определим обычное измерение T площади треугольника. Получается система величин $\mathfrak{U}_1$, которая содержится в системе $\mathfrak{U}$ величин из примера б). Структура в $\mathfrak{U}$ индуцируется в $\mathfrak{U}_1$ с тем единственным отличием, что операция сложения сужается не на все $\mathfrak{U}_1$, а на его часть $I\rightleftarrows\{w \mid \exists\Delta\in w(\Delta\subseteq u_1)\}$, где

u_1 — «половина» квадрата u_0. Последнее означает, например, что существует параллельный перенос ξ, для которого $u_0=u_1+(u_1+\xi)$ (см. рис. 49).

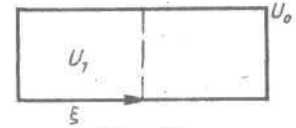


Рис. 49

Полученная система в $\mathfrak{U}_1$ имеет вид $\langle\mathfrak{U}_1, I, +, \leqslant, 0_{\mathfrak{U}}\rangle$. Она изоморфна системе в $\boldsymbol{R}_{\geqslant 0}$ вида $\langle\boldsymbol{R}_{\geqslant 0}, [0, \lambda], +, \leqslant, 0\rangle$, где $\lambda=\sup\{T(\Delta) \mid \Delta\subseteq u_1\}$. При этом изоморфизм задается биекцией T', суженной с $\mathfrak{U}$ на $\mathfrak{U}_1$.

д) Обозначим U множество всех многоугольников в арифметической плоскости $\boldsymbol{R}^2$ и T измерение, которое каждому многоугольнику сопоставляет его площадь. Понятия многоугольника и его площади рассматриваются в следующем § 3; здесь их можно понимать на интуитивном уровне.

В соответствии с общим определением величины величиной площади многоугольника u называется множество всех многоугольников u', которые имеют площадь, равную площади многоугольника u, т. е. множество $[u]\rightleftarrows\{u'\in U \mid T(u')=T(u)\}$. Возникает соответствующая система величин $\mathfrak{U}$. Для нее по T определяется измерение величин T'. Функция T' является биекцией вида $T:\mathfrak{U}\leftrightarrow R_{\geqslant 0}$. Точно как в примере б), добавим к системе величин $\mathfrak{U}$ новую величину $0_{\mathfrak{U}}$ и продолжим на нее измерение T' значением нуль; в результате получим биекцию вида $T':\mathfrak{U}\leftrightarrow R_{\geqslant 0}$. С помощью нее перенесем в $\mathfrak{U}$ операцию сложения и порядок в $\boldsymbol{R}_{\geqslant 0}$. Получим алгебраическую систему вида $\langle\mathfrak{U}, +, \leqslant, 0_{\mathfrak{U}}\rangle$, которая, очевидно, изоморфна системе $\langle R_{\geqslant 0}, +, \leqslant, 0\rangle$. Изоморфизм задается биекцией T'.

Как и в примерах а), б), опишем эту систему величин и структуру в ней без использования измерения T.

Предложение 3. ***Многоугольники u и u' имеют одинаковую площадь в том и только в том случае, когда существуют треугольники $\Delta_1, \ldots, \Delta_n$, попарно не имеющие общих внутренних точек, и треугольники $\Delta_1', \ldots, \Delta_n'$, также попарно не имеющие общих внутренних точек, для которых***

$$u=\bigcup_{i=1}^{n}\Delta_i, \quad u'=\bigcup_{i=1}^{n}\Delta_i'$$

и найдутся такие пары $\langle\boldsymbol{A}_1, \xi_1\rangle, \ldots, \langle\boldsymbol{A}_n, \xi_n\rangle$ из множества $L_2\times\boldsymbol{R}^2$, что $\Delta_1'=\boldsymbol{A}_1\cdot\Delta_1+\xi_1, \ldots, \Delta_n'=\boldsymbol{A}_n\cdot\Delta_n+\xi_n$.

Запись $\Delta_1+\ldots+\Delta_n$ обозначает множество $\Delta_1\cup\Delta_2\cup\ldots\cup\Delta_n$ в том случае, когда любые два треугольника Δ_i и Δ_j, где $i\neq j$ и $1\leqslant i, j\leqslant n$, не имеют общих внутренних точек.

▷ Пусть $u=\Delta_1+\ldots+\Delta_n$ и $u'=\Delta_1'+\ldots+\Delta_n'$ и $\Delta_i'=A_i\Delta_i+\xi_i$, где $\langle A_i, \xi_i\rangle\in(L_2\times\boldsymbol{R}^2)$ для всех $1\leqslant i, j\leqslant n$. Тогда по предложению 2 площади треугольников Δ_i и Δ_i' равны. Так как площадь многоугольника u равна сумме площадей составляющих его треугольников Δ_i, то получаем $T(u)=T(u')$.

Для доказательства утверждения в обратную сторону используем

следующую лемму, доказательство которой почти совпадает с доказательством аналогичной леммы 1 из предложения 2.

Лемма 2. ***Пусть Δ — произвольный треугольник и λ — произвольное положительное число (рис. 50). Тогда для треугольника $\Delta^{\lambda,\lambda_1}$ с вершинами в точках $\left\{\binom{0}{1}, \binom{\lambda}{0}, \binom{\lambda_1}{0}\right\}$, где $\lambda_1=\lambda+2\cdot T(\Delta)$, выполняется:***

а) $T(\Delta)=T(\Delta^{\lambda,\lambda_1})$,

б) найдется такая пара $\langle A, \xi\rangle \in (L_2\times R^2)$, что $\Delta^{\lambda,\lambda_1}=A\cdot\Delta+\xi$.

Продолжим доказательство предложения 3 и допустим, что $T(u)=T(u')$. Выберем произвольное разложение многоугольника u на треугольники, т. е. $u=\Delta_1+\ldots+\Delta_k$. Используя лемму 2, построим треугольники Δ^{λ_1}, $\Delta^{\lambda_1,\lambda_2}$, ..., $\Delta^{\lambda_{k-1},\lambda_k}$, у которых $\lambda_1=2\cdot T(\Delta_1)$ и $\lambda_i=\lambda_{i-1}+2\cdot T(\Delta_{i-1})$ при $1<i\leqslant k$ (см. рис. 51). Объединение этих треугольников — треугольник Δ^u с вершинами в точках $\left\{\binom{0}{0}, \binom{0}{1}, \binom{\lambda_k}{0}\right\}$. По построению $\lambda_k=2\cdot T(u)$ и $T(\Delta^u)=T(u)$.

Если провести аналогичное построение для многоугольника u' и его произвольного разложения на треугольники Δ'_1, ... Δ'_m, то получим треугольник $\Delta^{u'}$, для которого $T(\Delta^{u'})=T(u')$. Поэтому $T(\Delta^u)=T(\Delta^{u'})$ и, следовательно, треугольник Δ^u совпадает с треугольником $\Delta^{u'}$.

Пусть μ_1, μ_2, ... μ_n нумерация точек из множества $\{\lambda_1, \ldots, \lambda_k, \lambda'_1, \ldots \lambda'_m\}$ в порядке их возрастания. Рассмотрим разложение отрезка $[0, 2\cdot T(u)]$ на отрезки $[\mu_i, \mu_{i+1}]$. При этом каждый из треугольников Δ^{λ_1}, $\Delta^{\lambda_1\lambda_2}$, ..., $\Delta^{\lambda_{k+1},\lambda_k}$ разлагается точками μ_1, ..., μ_n на конечное число треугольников.

Тогда и каждый из треугольников Δ_1, Δ_2, ..., Δ_k разложится на конечное число треугольников. Например, если $\Delta^{\lambda_1}=\Delta^{\mu_1}+\Delta^{\mu_1,\mu_2}+\Delta^{\mu_2,\mu_3}$ и $\Delta^{\lambda_1}=A\cdot\Delta_1+\xi_1$ (рис. 52), то при эквиаффинном преобразовании, обратном к эквиаффинному преобразованию $\xi\mapsto A\cdot\xi+\xi_1$, треугольники Δ^{μ_1}, Δ^{μ_1,μ_2}, Δ^{μ_2,μ_3} перейдут в треугольники ∇_1, ∇_2, ∇_3, попарно не имеющие общих внутренних точек и в объединении дающие весь треугольник Δ_1. Тем самым и весь многоугольник u будет разложен на треугольники: $u=\nabla_1+\ldots+\nabla_n$. Аналогичным образом и весь многоугольник u' будет разложен на треугольники $u'=\nabla'_1+\ldots+\nabla'_n$. При этом по предложению 2

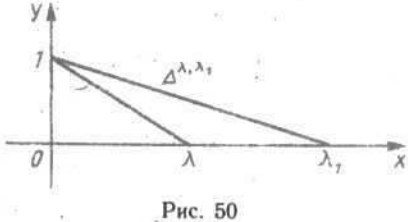


Рис. 50

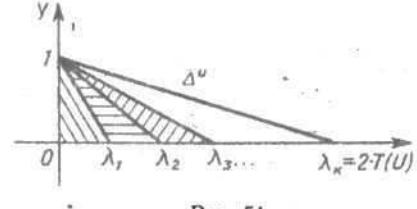


Рис. 51

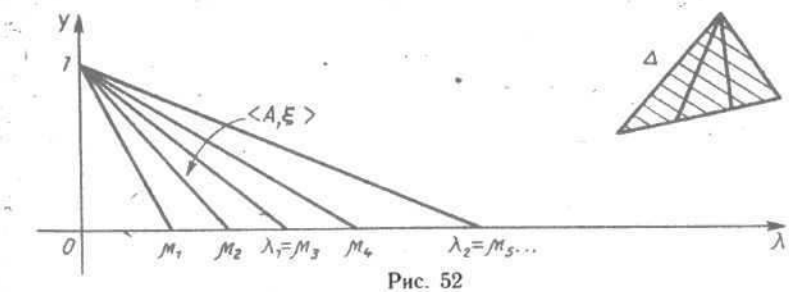


Рис. 52

треугольники ∇_i и ∇'_i будут иметь одинаковую площадь, так как оба они получаются в результате эквиаффинного преобразования одного и того же треугольника. А именно, i-го треугольника в разложении треугольника $\Delta^u=\Delta^{u'}$ на треугольники, определяемые точками μ_1, ..., μ_n. ■

Предложение 3 позволяет сформулировать то обстоятельство, что многоугольники u и u' имеют одинаковую площадь, не используя самого измерения площади. А именно, многоугольники u и u' назовем равносоставленными, если существуют их разложения на треугольники $u'=\Delta'_1+\ldots+\Delta'_n$ и $u=\Delta_1+\ldots+\Delta_n$, для которых каждый треугольник Δ_i эквиаффинно преобразуется в треугольник Δ'_i, где $i=1, \ldots n$. По предложению 3 многоугольники u и u' равновелики (т. е. имеют одинаковую величину площади) в том и только в том случае, если они равносоставлены. Итак, получено описание семейства величин $\mathfrak{U}$, не использующее измерение T.

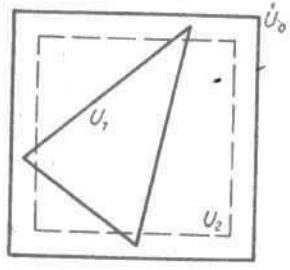
Определим операцию сложения и порядок в $\mathfrak{U}$ также без использования измерения T. Выберем такие представители u_1 и u_2 величин w_1 и w_2, что $u_1\cap u_2=\varnothing$. Это всегда можно сделать с помощью параллельного переноса. Тогда суммой w величин w_1 и w_2 из $\mathfrak{U}$ назовем величину, содержащую $u_1\cup u_2$. (В подразумеваемое здесь определение многоугольника не входит условие связности.)

Чтобы определить отношение порядка $w_1\leqslant w_2$, нужно, как и в доказательстве предложения 3, заменить представителей u_1 и u_2 величин w_1 и w_2 на эквивалентные им треугольники Δ^{u_1} и Δ^{u_2} и положить $(w_1\leqslant w_2)\rightleftarrows\Delta^{u_1}\subseteq\Delta^{u_2}$.

Аналогичное определение этой системы величин можно сформулировать и в «евклидовой плоскости».

Упражнение 6. *Докажите, что определенные выше операции сложения и порядки в $\mathfrak{U}$, совпадают.*

е) Рассмотрим только такие величины из предыдущего семейства величин $\mathfrak{U}$, которые содержат многоугольник, являющийся подмножеством какого-то фиксированного квадрата u_0. Семейство всех таких величин обозначим $\mathfrak{U}_1$. Поскольку $\mathfrak{U}_1\subset\mathfrak{U}$, то в $\mathfrak{U}_1$ определены операция сложения и отношение порядка, индуцированные структурой в $\mathfrak{U}$. Однако значение операции сложения может



$U_1 \in W_1, U_2 \in W_2$, $W_1 + W_2$ не определено

Рис. 53

не принадлежать $\mathfrak{U}_1$ (рис. 53). Поэтому ее приходится ограничить на множество I, состоящее из всех таких величин w из $\mathfrak{U}_1$, которые содержат многоугольник, являющийся частью «половины» квадрата u_0. Таким же образом мы поступали в примере 2.

Получается система вида $\langle \mathfrak{U}_1, I, +, \leqslant, 0_{\mathfrak{U}} \rangle$. Легко доказать изоморфность этой системы с алгебраической системой $\langle R_{\geqslant 0}, [0, \lambda], +, \leqslant, 0 \rangle$, где λ — площадь квадрата u_0. Этот изоморфизм задается биекцией T' из предыдущего примера, суженной на $\mathfrak{U}_1$.

ж) В качестве последнего примера рассмотрим группу углов An (см. гл. I). Операция сложения углов была определена и относительно нее имеется нейтральный элемент 0. Отношение порядка определяется следующим образом: если точка $A(\alpha)$ из верхней полуокружности, т. е. $P_2(A(\alpha)) > 0$, то положим (рис. 54) $R(\alpha) \rightleftarrows \boldsymbol{R}_{\geqslant 0} + c$, где $\langle \boldsymbol{R}_{\geqslant 0}, c \rangle$ — канонический представитель угла α и $A(\cdot)$ функция вида $A: \mathrm{An} \to S$, определенная в § 5, п. 2 главы I. Если $P_2(A(\alpha)) = 0$, то

$$R(\alpha) = \begin{cases} \boldsymbol{R}_{\geqslant 0}, & \text{если } P_1(A(\alpha)) > 0, \\ {}^{+}\boldsymbol{R}^2, & \text{если } P_1(A(\alpha)) < 0, \end{cases}$$

где ${}^{+}\boldsymbol{R}^2$ — верхняя полуплоскость (т. е. ${}^{+}R^2 \rightleftarrows \{\langle x, y \rangle \in \boldsymbol{R}^2 \mid y \geqslant 0\}$). Если $P_2(A(\alpha)) < 0$, то $R(\alpha) \rightleftarrows {}^{+}\boldsymbol{R}^2 \cup (\boldsymbol{R}_{\leqslant 0} + c)$, где $\langle \boldsymbol{R}_{\geqslant 0}, c \rangle \in \alpha$. Такое $R(a)$ — подмножество $\boldsymbol{R}^2$. Как раз такое подмножество обычно для наглядности связывают с углом α. Теперь положим: $(\alpha \leqslant \beta) \rightleftarrows R(\alpha) \subseteq R(\beta)$. Итак, в An возникла система уже привычного вида $\langle \mathrm{An}, +, \leqslant, 0 \rangle$. Однако для этой системы не выполняется свойство, имевшее место во всех предыдущих примерах: $\forall \alpha, \beta, \gamma (\alpha \leqslant \beta \Rightarrow \alpha + \gamma \leqslant \beta + \gamma)$.

Например, если $\alpha = d$ и $\beta = \gamma = 2d$. Тогда $\alpha \leqslant \beta$ и $\alpha + \gamma = 3d$, $\beta + \gamma = 0$. Чтобы сохранить это принципиально важное свойство, приходится, как и в некоторых предыдущих примерах, сузить операцию сложения на множество $I \rightleftarrows \{\alpha \in \mathrm{An} \mid \alpha < 2d\}$.

Как легко проверить, теперь это свойство выполняется.

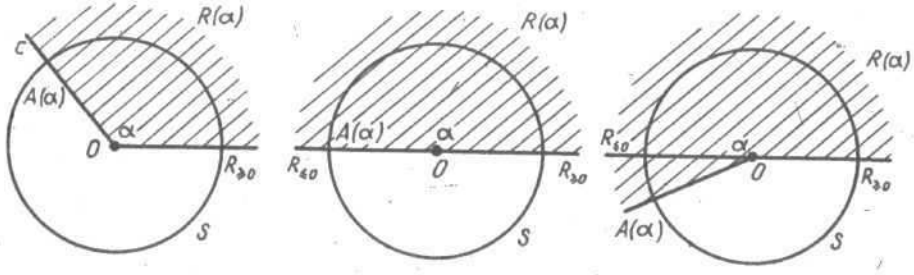


Рис. 54.

Главная мера угла по какому-нибудь фиксированному основанию $a > 0$ определяет изоморфизм этой системы величин с соответствующей алгебраической системой в $\boldsymbol{R}_{\geqslant 0}$ вида $\langle [0, a[, [0, {}^{a}/_{2}[, +, \leqslant, 0 \rangle$. Изоморфизм задается биекцией $f_a^{-1}: \mathrm{An} \to [0, a[$. Здесь понятие изоморфизма определяется также, как в примере в).

Такой же анализ величин скалярной скорости, массы, объема, температуры, заряда и т. п. приводит к удивительному экспериментальному открытию: самые разные (хотя и не все) системы величин изоморфны алгебраическим системам в $\boldsymbol{R}_{\geqslant 0}$ одного и того же вида и соответствующими изоморфизмами являются измерения величин.

Изученные примеры мотивируют следующее аксиоматическое определение.

Определение 1. **Системой положительных скалярных величин называется пятерка объектов $\langle V, I, +, \leqslant, w_0 \rangle$, где V — произвольное множество, $I \subseteq V$ и $\leqslant$ — бинарное отношение на V, причем $w_0 \in I$ и $+$ — бинарная операция, определенная на множестве I со значениями в V (но не обязательно в I).** При этом должны выполняться следующие аксиомы.

1) Операция $+$ ассоциативна, т. е. $\forall w_1, w_2, w_3 \in I$

$$(w_1 + w_2,\ w_2 + w_3 \in I) \Rightarrow w_1 + (w_2 + w_3) = (w_1 + w_2) + w_3,$$

и w_0 — ее нейтральный элемент, т. е. $\forall w \in I (w_0 + w = w + w_0 = w)$.

2) Отношение $\leqslant$ является линейным порядком в множестве V и w_0 — наименьший относительно этого порядка элемент в V; кроме того, подмножество I множества V обладает свойством:

$$\forall w \in I \forall w_1 \in V (w_1 \leqslant w \Rightarrow w_1 \in I).$$

3) Операция $+$ и порядок $\leqslant$ согласованы на I, т. е.

$$\forall w_1, w_2, w_3 \in I (w_1 < w_2 \Rightarrow (w_1 + w_3 \leqslant w_2 + w_3 \wedge w_3 + w_1 \leqslant w_3 + w_2)).$$

4) Из большего элемента множества I можно «вычитать» меньший элемент множества I, т. е.

$$\forall w_1, w_2 \in I (w_1 \leqslant w_2 \Rightarrow \exists w \in I (w_1 + w = w_2)).$$

5) Множество $I_{>w_0} = \{w \in I \mid w_0 \leqslant w \wedge w \neq w_0\}$ не пусто и не содержит наименьшего элемента.

6) Всякая последовательность элементов из множества I, возрастающая и ограниченная сверху элементом из I, имеет точную верхнюю грань, принадлежащую множеству I.

Запись $w_1 < w_2$, как всегда, означает $(w_1 \leqslant w_2) \wedge (w_1 \neq w_2)$.

По аксиоме 4 можно определить частично определенную бинарную операцию вычитания вида $I^2 \to I$. А именно, по аксиоме 4 для любых элементов w_1 и w_2 из I, удовлетворяющих условию $w_1 \leqslant w_2$, существует w из I, для которого $w_1 + w = w_2$. Причем такое w единственно: если $w_1 + w = w_2$ и $w_1 + w' = w_2$, то по аксиоме 3 $w < w' \Rightarrow w_1 + w < w_1 + w'$, $w_2 < w_2$, что невозможно. (Ради этого рассуждения в аксиоме 3 вместо отношения $\leqslant$ написано отноше-

ние $<$.) Итак, элемент w существует, принадлежит I и единствен; поэтому можно определить бинарную операцию, которая паре $\langle w_1, w_2\rangle$ сопоставляет элемент w.

Элементы множества V называют величинами.

Вместо аксиомы 6 иногда рассматривают более слабую аксиому — аксиому Архимеда:

6′) $\forall w_1,\ w \in I(w_0 < w \Rightarrow \exists n \in N(n \cdot w > w_1))$.

Здесь запись $n \cdot w$ означает, что все суммы $\underbrace{w+w}_{2}$, $\underbrace{w+w+w}_{3}$, ..., $\underbrace{w+\ldots+w}_{n-1}$ принадлежат I, и поэтому определена сумма

$$n \cdot w \rightleftarrows ((n-1)w + w).$$

В примерах а) — ж) были указаны системы, очевидно удовлетворяющие аксиомам 1—6; следовательно, каждая из них является примером системы положительных скалярных величин.

Бывают и другого рода системы величин. В физике — это системы, состоящие из величин импульса $m \cdot v$, или момента импульса $[x, p]$, или спина. Здесь m — масса частицы, v — ее скорость и x — положение. В математике — это, например, системы величин ориентированной площади или параллельного переноса. Такие системы величин не изоморфны алгебраическим системам в $\boldsymbol{R}$ — они имеют векторный характер. Однако эти системы величин изоморфны алгебраическим системам в $\boldsymbol{R}^n$. Они могут быть аксиоматически определены подобно определению положительной скалярной величины.

Кажется удивительным, что любая система, удовлетворяющая аксиомам 1—6, т. е. любая система положительных скалярных величин, изоморфна алгебраической системе в $\boldsymbol{R}_{\geqslant 0}$ вида $\langle \boldsymbol{R}_{\geqslant 0}, [0, \lambda], +, \leqslant, 0\rangle$. Правда, системы из предыдущих примеров а) — ж) допускали такой изоморфизм, но почему невозможны другие примеры, не допускающие такого изоморфизма.

Следующая теорема является итогом длительного развития математики.

Теорема 1. *Для любой системы положительных скалярных величин* $\langle V, I, +, \leqslant, w_0\rangle$ *существует инъективная функция* $f: I \to R_{\geqslant 0}$, *обладающая свойствами:*

а) для всех величин $w_1, w_2 \in I$ *выполняется*

$$w_1 + w_2 \in I \Rightarrow f(w_1 + w_2) = f(w_1) + f(w_2);$$

б) $\forall w_1, w_2 \in V(w_1 < w_2 \Rightarrow f(w_1) < f(w_2))$;

в) $f(w_0) = 0$;

г) *функция* f *биективно отображает множество* I *на отрезок вида* $[0, \lambda]$ *или* $[0, \lambda[$, *где* $0 < \lambda \leqslant \infty$ *(если* $\lambda = \infty$, *то речь идет об отрезке второго вида).*

Полезно сначала объяснить идею доказательства теоремы на конкретном примере.

Пусть на евклидовой прямой X (которая вовсе не совпадает с $\boldsymbol{R}$) дан отрезок $[a, b]$ и мы хотим «измерить его длину» относитель-

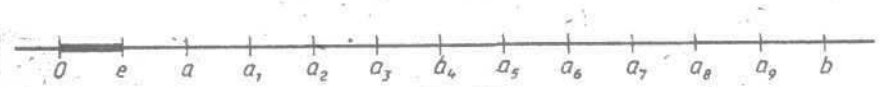


Рис. 55

но некоторого фиксированного отрезка $[o, e]$, который совершенно условно будем называть единичным отрезком. Начнем с того, что будем последовательно откладывать отрезок $[o, e]$ от точки a вдоль отрезка $[a, b]$, и остановимся в тот момент, когда увидим, что при следующем откладывание» состоит в действии на множество $[o, e]$ некоторой группы преобразований множества X. Если $X = \boldsymbol{R}$, то это группа параллельных переносов R^+; а именно, $[a, a_1] = [o, e] + a$, $[a_1, a_2] = [o, e] + a + e$, $[a_2, a_3] = [o, e] + a + 2e$, ... и т. д. (рис. 55).

Например, пусть мы сделали 198 откладываний, начиная с точки a, отрезок $[o, e]$ 198 раз отложился внутри отрезка $[a, b]$ и 199 откладывание вывело нас за пределы отрезка $[a, b]$. Будем записывать это обстоятельство следующим образом: $198/(ab : oe)$, и говорить, что длина отрезка $[a, b]$ относительно $[o, e]$ равна числу 198 с точностью до целых (по недостатку). Если мы хотим увеличить точность измерения длины отрезка $[a, b]$, то единичный отрезок $[o, e]$ нужно разделить на 10 равных частей, затем взять отрезок $[o, e_1]$, где e_1 — первая точка этого деления, и найти длину отрезка $[a, b]$ относительно отрезка $[o, e_1]$ с точностью до целых. Иными словами, найти число $(ab : oe_1)$. Ясно, что это число будет лежать в пределах от 1980 до 1989. Пусть, например, $(ab : oe_1)$ равно числу 1983. Какое же число следует назвать приближением длины отрезка $[a, b]$ относительно отрезка $[o, e]$ с точностью до десятых? Разумеется, не число 1983, а число $\frac{1983}{10} = 198{,}3$. По построению $10 = (oe : oe_1)$, и длиной отрезка $[a, b]$ относительно отрезка $[o, e]$ с точностью до десятых (по недостатку) назовем число $\frac{(ab : oe_1)}{(oe : oe_1)}$.

Отметим, что понятие «разделить отрезок на 10 частей» опять-таки апеллирует к действию группы преобразований множества X. Если $X = \boldsymbol{R}$, то это группа $\boldsymbol{R}^+$ и отрезок $[o, e_1]$ определяется как такой отрезок, для которого $([o, e_1] + o \cdot e_1) \cup ([o, e_1] + 1 \cdot e_1) \cup \ldots \cup ([o, e_1] + 9 \cdot e_1) = [o, e]$.

Дальнейшее очевидно: нужно последовательно вычислять следующие отношения натуральных чисел (рациональные числа):

$$\frac{(ab : oe_2)}{(oe : oe_2)},\ \frac{(ab : oe_3)}{(oe : oe_3)},\ \ldots,\ \frac{(ab : oe_n)}{(oe : oe_n)},\ \ldots,$$

где $[o, e_n]$ — десятая часть отрезка $[o, e_{n-1}]$.

По построению такая последовательность монотонно возрастает и ограничена сверху.

В нашем конкретном примере такой верхней границей служит число 1990. Следовательно, у этой последовательности существует предел, который по определению и называется длиной отрезка $[a, b]$ относительно единичного отрезка $[o, e]$.

В предыдущем построении каждый последующий отрезок $[o, e_n]$ совсем не обязательно полагать равным именно одной десятой части предыдущего отрезка $[o, e_{n-1}]$. С таким же успехом можно производить деление отрезка $[o, e_{n-1}]$, например, на семь равных отрезков. (Не следует думать, что таким образом получится длина отрезка $[a, b]$ в семеричной системе счисления.) Более того, не обязательно предполагать, что отрезок $[o, e_n]$ является целой частью отрезка $[o, e_{n-1}]$: существенно только «стремление» отрезка $[o, e_n]$ к нулю. Это условие опять-таки апеллирует к действию группы преобразований множества X: оно означает, что число итераций (повторений) преобразования отрезка $[o, e_n]$ (в частности, число параллельных переносов отрезка $[o, e_n]$), которое необходимо, для того чтобы объединение всех перенесенных отрезков равнялось отрезку $[o, e]$, стремилось к бесконечности.

Итак, длину отрезка $[a, b]$ относительно отрезка $[o, e]$ можно определить таким образом: положим

$$f(e') = \frac{(ab : oe')}{(oe : oe')},$$

где $e' \in [o, e]$, тогда искомая длина равна

$$T([a, b]) \rightleftarrows \lim_{e' \to o} f(e').$$

Наиболее тонким моментом в этом определении является вопрос о точном смысле выражения «e' стремится к o».

Мы не будем рассматривать этот вопрос в данном конкретном случае, а решим его в общем виде для любой положительной скалярной величины $\langle V, I, +, \leqslant, w_0 \rangle$, и таким образом получим доказательство теоремы 1. Этот материал содержится в приложении 2.

Отметим, что длину отрезка $[a, b]$ относительно отрезка $[o, e]$ можно определить и через приближения с избытком как $\lim_{e' \to o} \frac{(ab : oe') + 1}{(oe : oe') + 1}$.

§ 3. ИЗМЕРЕНИЕ ПЛОЩАДИ МНОГОУГОЛЬНИКА

1. Конструктивное определение площади многоугольника. Свойство конечной аддитивности. В этом параграфе удобнее пользоваться вторым определением треугольника как множества вида

$$\Delta \rightleftarrows \{\lambda_1 \cdot a + \lambda_2 \cdot b + \lambda_3 \cdot c \mid \lambda_1, \lambda_2, \lambda_3 \geqslant 0, \lambda_1 + \lambda_2 + \lambda_3 = 1\},$$

где тройка точек a, b, c из плоскости $\boldsymbol{R}^2$ не лежит на одной прямой. Стороной треугольника Δ называется отрезок, соединяющий какие-то из точек a, b, c; ясно, что любая сторона треугольника состоит из всех точек вида $\lambda_1 \cdot a + \lambda_2 \cdot b + \lambda_3 \cdot c$, у которых один из

параметров λ_1, λ_2, λ_3 тождественно равен нулю. Сами точки a, b, c называются вершинами треугольника.

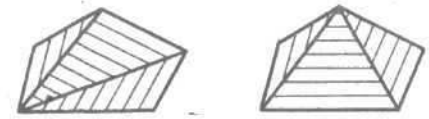


Рис. 56

Многоугольником мы называем любое подмножество арифметической плоскости $\boldsymbol{R}^2$, которое представимо в виде объединения конечного числа треугольников, не имеющих общих внутренних точек. Иными словами, подмножество u плоскости $\boldsymbol{R}^2$ называется многоугольником, если существуют такие треугольники $\Delta_1, \Delta_2, \ldots, \Delta_n$, что:

1) $u = \bigcup_{i=1}^{n} \Delta_i$,

2) $i \neq j \Rightarrow \mathring{\Delta}_i \cap \mathring{\Delta}_j = \varnothing$, где $\mathring{\Delta}$ — множество внутренних точек Δ (внутренность Δ).

Это определение многоугольника существенно шире обычного. Множество треугольников $\{\Delta_1, \ldots, \Delta_n\}$, для которого выполняются условия 1) и 2), называется разложением многоугольника u (рис. 56). Любой многоугольник имеет бесконечное число различных разложений на треугольники. Два из них показаны на рисунке 57.

Отметим, что определение многоугольника зависит от выбора топологии в $\boldsymbol{R}^2$, так как в нем используется понятие внутренности треугольника. Правда, внутренность треугольника можно определить, используя вместо топологии векторную структуру в $\boldsymbol{R}^2$. А именно, $\Delta = \{\lambda_1 \cdot a + \lambda_2 \cdot b + \lambda_3 \cdot c \mid \lambda_1, \lambda_2, \lambda_3 > 0, \lambda_1 + \lambda_2 + \lambda_3 = 1\}$.

Понятие разложения не следует путать с понятием триангуляции. Триангуляцией называется разложение, удовлетворяющее еще одному дополнительному условию: любые два треугольника Δ_i и Δ_j при $i \neq j$ пересекаются или по общей вершине, или по общей стороне, или вообще не пересекаются. Однако для наших целей это условие излишне.

Если w и $w_1, \ldots w_n$ — любые подмножества $\boldsymbol{R}^2$ и для них выполняются условия: $w = \bigcup_{i=1}^{n} w_i$ и $i \neq j \Rightarrow \mathring{w}_i \cap \mathring{w}_j = \varnothing$, то семейство $\{w_1, \ldots, w_n\}$ будем называть разложением множества w и писать $w = w_1 + \ldots + w_n$. Запись $w_1 + \ldots + w_n$ означает, что семейство $\{w_1,$



Рис. 57

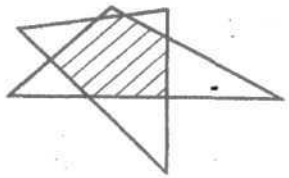


Рис. 58

..., $w_n\}$ является разложением множества $w_1 \cup \ldots \cup w_n$. Напомним, что произвольное подмножество w плоскости $\boldsymbol{R}^2$ называется выпуклым, если $\forall x, y \in w([x, y] \subseteq w)$, где $[x, y] \rightleftarrows \{\lambda \cdot x + (1-\lambda) \cdot y \mid 0 \leqslant \lambda \leqslant 1\}$.

Упражнение 7. *Докажите, что: а) пересечение любых двух треугольников — пустое множество или точка, или отрезок, или выпуклый многоугольник (рис. 58); б) для любого выпуклого многоугольника можно получить его разложение, соединив отрезками одну из его вершин со всеми остальными вершинами.*

Множество всех многоугольников в арифметической плоскости $\boldsymbol{R}^2$ обозначим U.

Пусть $u \in U$ и $u = \Delta_1 + \ldots + \Delta_n$. Обозначим $T(u)$ число $T(\Delta_1) + \ldots + T(\Delta_n)$, где число $T(\Delta_i)$ определяется в примере б) из предыдущего параграфа. Интуитивно $T(u)$ — площадь многоугольника u.

Здесь мы сталкиваемся с типичным для темы многоугольников вопросом: зависит ли данное определение числа $T(u)$ от выбора разложения $\{\Delta_1, \ldots, \Delta_n\}$ многоугольника u на треугольники? Иными словами, верно ли следующее $((u = \Delta_1 + \ldots + \Delta_n) \wedge (u = \Delta'_1 + \ldots + \Delta'_m)) \Rightarrow T(\Delta_1) + \ldots + T(\Delta_n) = T(\Delta'_1) + \ldots + T(\Delta'_m)$. Если это не так, то определение числа $T(u)$, где $u \in U$, задает функцию от двух аргументов вида $T(u, \{\Delta_i\})$, но не функцию от одного аргумента вида $T(u)$. Нам нужна именно вторая функция, так как измерение площади многоугольника должно зависеть только от самого многоугольника.

Предложение 4. ***а) Допустим, что данное определение числа $T(u)$, где $u \in U$, не зависит от выбора разложения многоугольника u на треугольники, т. е. оно определяет функцию вида $T: U \to \boldsymbol{R}_{>0}$. Тогда эта функция конечно-аддитивна, т. е.***

$$\forall u, v \in U\,(T(u+v) = T(u) + T(v)).$$

б) Если существует хотя бы одна конечно-аддитивная функция вида $f: U \to \boldsymbol{R}_{>0}$, значение которой на любом треугольнике $\Delta = \{a, b, c\}$ равно $f(\Delta) = \frac{1}{2}\left|\det\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \end{pmatrix}\right|$, то определение числа $T(u)$ не зависит от выбора разложения для произвольного многоугольника u, т. е. оно определяет функцию вида $T: U \to \boldsymbol{R}_{>0}$.

в) Если треугольник Δ разложен на два треугольника Δ_1 и Δ_2 отрезком, соединяющим одну из его вершин с точкой на противоположной стороне, то $T(\Delta) = T(\Delta_1) + T(\Delta_2)$.

▷ а) Свойство конечной аддитивности, очевидно, эквивалентно свойству $\forall n \in \boldsymbol{N} \forall u_1, \ldots, u_n \in U(T(u_1 + \ldots + u_n) = T(u_1) + \ldots + T(u_n))$, откуда и происходит его название.

Если T является функцией от u, $u = \Delta_1 + \ldots + \Delta_n$ и $v = \Delta'_1 + \ldots + \Delta'_m$, то, как легко проверить, $u + v = \Delta_1 + \ldots + \Delta_n + \Delta'_1 + \ldots + \Delta'_m$, т. е. $\{\Delta_1, \ldots, \Delta_n, \Delta'_1, \ldots, \Delta'_m\}$ — одно из разложений многоугольника $u + v$.

Поэтому $T(u+v) = T(\Delta_1) + \ldots + T(\Delta'_m) = T(u) + T(v)$.

б) Если $u = \Delta_1 + \ldots + \Delta_n$ и $u = \Delta'_1 + \ldots + \Delta'_m$, то $T(\Delta_1) + \ldots + T(\Delta_n) = f(\Delta_1) + \ldots + f(\Delta_n) = f(u) = f(\Delta'_1) + \ldots + f(\Delta'_m) = T(\Delta'_1) + \ldots + T(\Delta'_m)$.

в) По лемме 1 из § 2 существует эквиаффинное преобразование $f_{A, \xi}$, сохраняющее площадь треугольника, при котором треугольник Δ перейдет в треугольник вида Δ^λ. При этом преобразовании треугольники Δ_1 и Δ_2 перейдут в треугольники Δ'_1 и Δ'_2 той же площади (по предложению 2) и треугольники Δ'_1 и Δ'_2 разлагают треугольник Δ^λ отрезком, соединяющим одну из его вершин с точкой на противоположной стороне (почему?). Осталось проверить, что $T(\Delta^\lambda) = T(\Delta'_1) + T(\Delta'_2)$, что следует из легкого прямого подсчета. ■

Замечание. 1) Казалось бы, объединяя рассуждение из доказательства предложения 4 а) с предложением 4 в), можно доказать, что число $T(u)$ не зависит от разложения u на треугольники. А именно, если $u = \Delta_1 + \ldots + \Delta_n = \Delta'_1 + \ldots + \Delta'_m$ — два разложения одного многоугольника u, то образуем третье разложение $\{\Delta_i \cap \Delta_j\}$, где $1 \leqslant i \leqslant n$ и $1 \leqslant j \leqslant m$. Возможно для какой-то пары i, j множество $\Delta_i \cap \Delta_j$ не треугольник, а выпуклый многоугольник. Тогда нужно разбить его на треугольники по упражнению 7 б). Для простоты записи допустим, что такой случай не встречается. Тогда внутренности треугольников $\Delta_{ij} \rightleftarrows (\Delta_i \cap \Delta_j)$ попарно не пересекаются и $\Delta_i = \Delta_{i1} + \ldots + \Delta_{im}$, $\Delta'_j = \Delta_{1j} + \ldots + \Delta_{nj}$. Если эти разложения получаются последовательно так, что применимо предложение 4 в), то по индукции $T(\Delta_i) = T(\Delta_{i1}) + \ldots + T(\Delta_{im})$ и $T(\Delta'_j) = T(\Delta_{1j}) + \ldots + T(\Delta_{nj})$. Тогда $T(\Delta_1) + \ldots + T(\Delta_n) = \sum\limits_{i,\,j} T(\Delta_{ij}) = T(\Delta'_1) + \ldots + T(\Delta'_m)$, что и нужно доказать для определения функции $T: U \to \boldsymbol{R}_{>0}$. Однако это рассуждение не годится, так как не всегда применимо предложение 4 в).

2) Верно, что определение числа $T(u)$ не зависит от разложения u, и существует несколько весьма разных доказательств этого факта. Каждое из них не просто и с методической точки зрения имеет свои трудности. Мы докажем этот факт следующим образом. Множество всех треугольников обозначим U_0 и для $\Delta = \{a, b, c\}$ из U_0 положим $T_0(\Delta) = \frac{1}{2}\left|\det\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \end{pmatrix}\right|$. Определение функции $T_0: U_0 \to \boldsymbol{R}_{>0}$ не требует никаких обоснований, так как оно носит совершенно явный характер. В § 5 мы продолжим функцию T_0 с множества U_0 на очень широкое множество $\mathscr{L}$, включающее все многоугольники, т. е. $U \subset \mathscr{L}$, и это продолжение M (называемое мерой Лебега) будет обладать свойством конечной аддитивности (и даже более сильным свойством счетной аддитивности) не только на U, но и на всем множестве $\mathscr{L}$. Используя предложение 4 б) и функцию $M \mid U$, сразу получим независимость числа $T(u)$ от разложения u. Методическим недостатком такого подхода является необходимость принять в качестве гипотезы независимость числа $T(u)$ от разложения u (вплоть до теоремы 7 в) из § 5). Конечно, можно переставить местами параграфы 3 и 5, но тогда нарушится историческая (и, вероятно, логическая) последовательность в развитии темы измерения плоских фигур: понятие меры Лебега возникло поз-

же, чем понятие площади, и является его развитием, а также было мотивировано свойствами функции T. Кроме того, определение функции M и проверка ее конечной аддитивности требуют кропотливых рассуждений, далеко уводящих от темы многоугольников.

Итак, в этом параграфе мы установим характерные свойства функции T, принимая в качестве гипотезы, что она существует. В следующем параграфе эти свойства будут положены в основу аксиоматического определения площади многоугольника.

В заключение напомним два определения, которые часто используются в дальнейшем.

Как обычно, фигурой называется произвольное подмножество в $\boldsymbol{R}^2$. Поскольку в $\boldsymbol{R}^2$ фиксирована топология, то для любой фигуры определены понятия внутренней и граничной точки. В частности, они определены для многоугольника. Если точка c при каждом разложении многоугольника является вершиной некоторого треугольника из этого разложения, то назовем c вершиной многоугольника. Легко доказать, что вершина является граничной точкой многоугольника u, но обратное неверно.

2. Инвариантность функции площади относительно эквиаффинной группы. Напомним, что линейной биекцией векторной плоскости X, в частности $\boldsymbol{R}^2$, называется биекция $f: X \to X$, обладающая свойством

$$\forall \lambda_1, \lambda_2 \in \boldsymbol{R}\, \forall x, y \in X (f(\lambda_1 \cdot x + \lambda_2 \cdot y) = \lambda_1 \cdot f(x) + \lambda_2 \cdot f(y)).$$

Аффинной биекцией векторной плоскости X называется биекция $f: X \to X$ вида $f(x) \rightleftarrows (h(x) + \xi)$, где $h: X \to X$ линейная биекция и ξ фиксированный вектор из X.

Упражнение 8. *Докажите, что: а) для любого многоугольника и любой линейной биекции $A: \boldsymbol{R}^2 \to \boldsymbol{R}^2$ множество $A(u)$ также является многоугольником;*

б) для любого линейного отображения $A: \boldsymbol{R}^2 \to \boldsymbol{R}^2$ выполняется $\forall x \in \boldsymbol{R}^2 (A(x) = A \cdot x)$, где матрица $A \rightleftarrows \left(A\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) A\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)\right)$; при этом отображение A является биекцией в том и только том случае, если $\det A \neq 0$, *т. е.* $A \in GL_2$. Обычно отображение A и матрицу A обозначают одной буквой и не различают. Так же будем поступать и мы.

В этом пункте мы снова рассмотрим эквиаффинные преобразования $f_{A, \xi}$, с которыми имели дело в примере б) из § 2, и группу G_s всех эквиаффинных преобразований опишем более подробно и в более абстрактном духе. При первом чтении можно пропустить этот материал и перейти непосредственно к теореме 2.

Как и раньше, обозначим L_2 множество всех (2×2) — матриц из вещественных чисел с определителем $+1$ или -1. Рассмотрим декартово произведение множеств L_2 и $\boldsymbol{R}^2$. Каждому $g = \langle A, \xi \rangle$ из $L_2 \times \boldsymbol{R}^2$ поставим в соответствие отображение $f_g: \boldsymbol{R}^2 \to \boldsymbol{R}^2$, определенное следующим образом: $f_g(x) \rightleftarrows A \cdot x + \xi$, где $A \cdot x$ — результат умножения матрицы A на вектор x. Иногда вместо f_g пишут $f_{A, \xi}$.

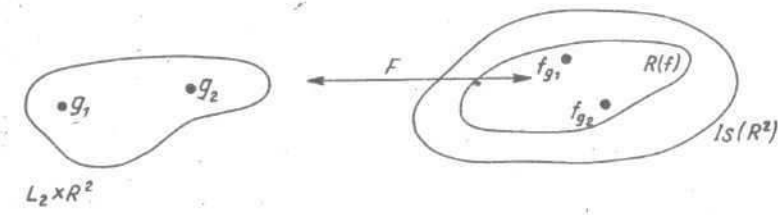


Рис. 59

Если $f_g(x_1) = f_g(x_2)$, то $A \cdot x_1 + \xi = A \cdot x_2 + \xi$ и $A(x_1 - x_2) = 0$. Учитывая, что $\det A \neq 0$, получаем $x_1 - x_2 = 0$ и $x_1 = x_2$. Следовательно, f_g — инъективное отображение $\boldsymbol{R}^2$ в себя. Поскольку $f_g(A^{-1}(\xi_0 - \xi)) = A(A^{-1}(\xi_0 - \xi)) + \xi = \xi_0$, то отображение f_g сюръективно. Поэтому оно биективно. Итак, f_g принадлежит множеству всех биекций $\boldsymbol{R}^2$ на себя, которое мы обозначаем $\mathrm{Is}(\boldsymbol{R}^2)$. Соответствие $F: g \mapsto f_g$ задает отображение множества $L_2 \times \boldsymbol{R}^2$ в множество $\mathrm{Is}(\boldsymbol{R}^2)$. Это соответствие инъективно: если $F(g_1) = F(g_2)$, то $f_{g_1}(x) = f_{g_2}(x)$ для всех $x \in \boldsymbol{R}^2$, т. е. для всех $x \in \boldsymbol{R}^2$ выполняется $A_1 \cdot x + \xi_1 = A_2 \cdot x + \xi_2$. Полагая $x = 0$, получим $\xi_1 = \xi_2$. Поэтому $A_1 \cdot x = A_2 \cdot x$ для всех $x \in \boldsymbol{R}^2$. Значит, $A_1 = A_2$ и, следовательно, $g_1 = g_2$.

Будем отождествлять множество $L_2 \times \boldsymbol{R}^2$ с множеством значений $F(L_2 \times \boldsymbol{R}^2)$ описанного инъективного отображения $F: L_2 \times \boldsymbol{R}^2 \to \mathrm{Is}(\boldsymbol{R}^2)$ (рис. 59).

Докажем, что для любых пар g_1 и g_2 из $L_2 \times \boldsymbol{R}^2$ выполняется $f_{g_1} \circ f_{g_2} \in R(F)$.

Рассмотрим биекцию $h \rightleftarrows f_{g_1} \circ f_{g_2}: \boldsymbol{R}^2 \to \boldsymbol{R}^2$ и найдем такое $g = \langle A, \xi \rangle$, что $f_g = h$. Для этого вычислим $h(x) = f_{g_1}(f_{g_2}(x)) = f_{g_1}(A_2 \cdot x + \xi_2) = A_1 \cdot (A_2 \cdot x + \xi_2) + \xi_1$. Положим $A \rightleftarrows A_1 \cdot A_2 \in L_2$ и $\xi \rightleftarrows A_1 \cdot \xi_2 + \xi_1$. Ясно, что $A \in L_2$ и $\xi \in \boldsymbol{R}^2$. Если теперь $g \rightleftarrows \langle A, \xi \rangle \in L_2 \times \boldsymbol{R}^2$, то $\forall x \in \boldsymbol{R}^2 (h(x) = f_g(x))$.

Иными словами, композиция двух биекций, являющихся образами произвольных пар g_1 и g_2 из $L_2 \times \boldsymbol{R}^2$ при отображении F, будет образом некоторой пары g при этом отображении. Этот факт мотивирует определение следующей бинарной операции в множестве $L_2 \times \boldsymbol{R}^2$:

$$\langle A_1, \xi_1 \rangle \cdot \langle A_2, \xi_2 \rangle \rightleftarrows \langle A_1 \cdot A_2, A_1 \xi_2 + \xi_1 \rangle.$$

Упражнение 9. *Докажите, что: а) относительно этой операции в множестве $L_2 \times \boldsymbol{R}^2$ найдется нейтральный элемент;*

б) относительно этой операции всякий элемент g из $L_2 \times \boldsymbol{R}^2$ имеет обратный элемент (в том же множестве);

в) относительно этой операции множество $L_2 \times \boldsymbol{R}^2$ образует группу (коммутативна ли она?);

г) отображение F — инъективный гомоморфизм группы $L_2 \times \boldsymbol{R}^2$ в группу $\mathrm{Is}(\boldsymbol{R}^2)$ и изоморфизм группы $L_2 \times \boldsymbol{R}^2$ и подгруппы $R(F)$ группы $\mathrm{Is}(\boldsymbol{R}^2)$.

Пусть даны две группы G_1 и G_2, групповую операцию в G_1 обозначим $\cdot$ и групповую операцию в G_2 обозначим $+$. Кроме того, пусть задан гомоморфизм группы G_1 в группу $\mathrm{Is}(G_2)$, который мы обозначим $g \mapsto g(\cdot)$, где $g(\cdot): G_2 \to G_2$. Если в декартовом произведении групп G_1 и G_2 групповую операцию определить по формуле

$$\langle g_1, h_1\rangle \cdot \langle g_2, h_2\rangle \rightleftharpoons \langle g_1 \cdot g_2,\ g_1(h_2) + h_1\rangle,$$

то получится группа (проверьте); она называется полупрямым произведением групп G_1 и G_2. Таким образом, выше мы определили полупрямое произведение групп L_2 и $(\boldsymbol{R}^2)^+$. Напомним, что в примере б) из § 2 мы назвали эквиаффинной группой G_s множество всех биекций плоскости $\boldsymbol{R}^2$ на себя вида $f_{\langle A,\ \xi\rangle}(x) = A \cdot x + \xi$, где $A \in L_2$ и $\xi \in \boldsymbol{R}^2$, а $x \in \boldsymbol{R}^2$, вместе с операцией композиций таких биекций. Согласно упражнению 9 г) полупрямое произведение $L^2 \times (\boldsymbol{R}^2)^+$ изоморфно эквиаффинной группе G_s плоскости $\boldsymbol{R}^2$.

Роль этой группы видна из следствия 2 к следующей теореме.

Теорема 2. ***Функция T обладает свойствами:***

а) ***конечной аддитивности, т. е.***

$$\forall u, v \in U\,(T(u+v) = T(u) + T(v));$$

б) ***G_s — инвариантности, т. е.*** $\forall u \in U\ \forall \langle A, \xi\rangle \in L_2 \times R^2$

$$(T(u) = T(A \cdot u + \xi)).$$

▷ Пункт а) этой теоремы обсуждался в замечании 2 после предложения 4.

Докажем пункт б). Для этого сначала проверим:

$$\forall u \in U\ \ \forall \langle A, \xi\rangle \in L_2 \times \boldsymbol{R}^2 (A \cdot u + \xi \in U).$$

Если u — многоугольник и $\Delta_1, \ldots, \Delta_n$ — его разложение, то треугольники $f_g(\Delta_1), \ldots, f_g(\Delta_n)$ не имеют общих внутренних точек. Действительно, f_g — гомеоморфизм $\boldsymbol{R}^2$ на себя, а при гомеоморфизме внутренние точки фигуры переходят во внутренние точки образа этой фигуры. Поэтому $f_g(u)$ — многоугольник и $f_g(\Delta_1), \ldots, f_g(\Delta_n)$ — его разложение.

По предложению 2 выполняется: $\forall \Delta \forall g = \langle A, \xi\rangle \in L_2 \times \boldsymbol{R}^2$

$$T(\Delta) = T(f_g(\Delta)).$$

Следовательно, $T(f_g(u)) = T(f_g(\Delta_1) + \ldots + f_g(\Delta_n)) = T(f_g(\Delta_1)) + \ldots + T(f_g(\Delta_n)) = T(\Delta_1) + \ldots + T(\Delta_n) = T(u)$. ■

Следствие 1. ***Если $g = \langle A, \xi\rangle$, где $\xi \in \boldsymbol{R}^2$ и A — линейная биекция плоскости $\boldsymbol{R}^2$, то $\forall u \in U$***

$$T(f_g(u)) = |\det A| \cdot T(u).$$

▷ Для треугольников такая формула проверяется прямым подсчетом, который, по существу, был выполнен в примере б) из § 2. Затем используется пункт а) теоремы 2. ■

Итак, биекция f_g изменяет площадь $T(u)$ в $|\det A|$ раз. Поэтому

необходимым и достаточным условием сохранения числа $T(u)$ под действием биекции f_g является равенство $\det A = \pm 1$. Отсюда получаем

Следствие 2. ***Группа L_2 — наибольшая среди всех групп, содержащихся в GL_2, относительно которых инвариантна функция T, а группа G_s — наибольшая среди всех групп, содержащихся в полупрямом произведении $GL_2 \times \boldsymbol{R}^2$, относительно которых инвариантна та же функция T.***

Замечание. Свойство конечной аддитивности включено в формулировку теоремы 2, для того чтобы эти два свойства (конечной аддитивности и G_s-инвариантности) прозвучали вместе. Мы увидим, что как раз они однозначно характеризуют функцию T. Основанное на этих свойствах аксиоматическое определение по существу функции T содержится в следующем § 4.

Следствие 3. Пусть H_k — гомотетия с коэффициентом $k \neq 0$ плоскости $\boldsymbol{R}^2$, т. е. $H_k(x) \rightleftharpoons kx$. Тогда

$$\forall u \in U (T(H_k(u)) = k^2 T(u)).$$

▷ Гомотетия является частным случаем преобразования из следствия 1. ■

Свойство функции T, содержащееся в этом следствии, называют гомотетичностью площади.

§ 4. СРАВНЕНИЕ КОНСТРУКТИВНОГО И АКСИОМАТИЧЕСКОГО ОПРЕДЕЛЕНИЙ ПЛОЩАДИ МНОГОУГОЛЬНИКА. СРАВНЕНИЕ РАЗЛИЧНЫХ СПОСОБОВ ИЗМЕРЕНИЯ ПЛОЩАДИ МНОГОУГОЛЬНИКА

1. Аксиоматическое определение площади многоугольника и его сравнение с конструктивным определением. Часто фигурой называют произвольное подмножество $\boldsymbol{R}^2$. В таком случае совокупность всех фигур в $\boldsymbol{R}^2$ совпадает с $\mathfrak{P}(\boldsymbol{R}^2)$. Измерение площади фигуры должно задавать, может быть, частично определенную функцию вида $f: \mathfrak{P}(\boldsymbol{R}^2) \to R_{\geqslant 0}$, обладающую свойствами, похожими на свойства функции T; а именно, свойствами аддитивности и инвариантности, так как кажется, что любой процесс измерения приводит к функции, обладающей этими двумя свойствами. Поскольку конкретное измерение — функция T — определено на множестве всех многоугольников U, то представляется естественным, предположение, что измерение f применимо, по крайней мере, ко всем многоугольникам, т. е. $U \subseteq D(f)$. Что еще можно включить в $D(f)$?

Практика указывает на необходимость и возможность измерения площади у таких фигур, как круг, внутренность эллипса, сектор, сегмент, кольцо, и других элементарных криволинейных фигур с сохранением при этом основных свойств аддитивности и инвариантности. Поэтому область определения $D(f)$ шире множества U и включает большое число разнообразных криволинейных фигур. Однако

нельзя определить измерение, обладающее этими основными свойствами и пригодное для измерения любых фигур, т. е. $D(f) \neq \mathfrak{P}(\boldsymbol{R}^2)$.

Традиционным является деление фигур на «прямолинейные» (многоугольники) и «криволинейные». Имеется много оснований, для того чтобы различать эти два случая. Поэтому отдельно рассмотрим измерение площади многоугольника и измерение площади криволинейной фигуры. Соответственно возникают две функции s и μ, причем $D(s)=U$, а $D(\mu)$ — более широкое множество, включающее все те фигуры, площадь которых можно измерить; в частности, $U \subset D(\mu)$. Поскольку функции s и μ имеют общий интуитивный источник в виде измерения T, то естественно считать, что s — сужение μ на U и μ — продолжение s на $D(\mu)$. Первую из этих функций будем называть функцией площади, а вторую — мерой. В соответствии с этой терминологией $\mu(w)$ называют не площадью, а мерой фигуры w.

В этом параграфе аксиоматически определяется функция s, при этом за аксиомы принимаются свойства аддитивности и инвариантности, точнее — свойства функции T из теоремы 2. В следующем § 5 аксиоматически определяется мера — за аксиомы опять-таки принимаются свойства аддитивности (в ее сильном виде) и инвариантности.

В теореме 3 доказывается, что любая аксиоматически определенная функция площади s отличается от функции T на постоянный множитель: $s(\cdot) \equiv c \cdot T(\cdot)$. Поэтому свойства аддитивности и инвариантности по существу однозначно определяют измерение T, т. е. возможно, по существу, только одно измерение площади многоугольника. Это удивительное обстоятельство используется затем для установления того, что на первый взгляд весьма различные способы измерения площади многоугольника приводят к одному и тому же численному результату.

Определение 2. **Функцией площади называется любая функция $s: U \to \boldsymbol{R}_{>0}$, обладающая свойствами:**

1) **конечной аддитивности**, т. е. $\forall u, v \in U$

$$s(u+v) = s(u) + s(v);$$

2) **G_s-инвариантности**, т. е. $\forall u \in U \forall g \in (L_2 \times \boldsymbol{R}^2)$

$$s(u) = s(f_g(u)).$$

Теорема 2 из предыдущего параграфа утверждает, что функция T удовлетворяет этому определению и, следовательно, хотя бы одна функция площади существует. Очевидно, что любая функция вида $c \cdot T(\cdot)$, где $c \in \boldsymbol{R}_{>0}$, также является функцией площади. Таким образом, существует бесконечное число различных функций площади. Существует ли функция площади, которая не имеет вида $c \cdot T(\cdot)$? Отрицательный ответ на этот вопрос содержится в следствии 1 к приводимой ниже теореме 3.

В теореме 3 потребуются две леммы.

Лемма 3. ***Множество точек разрыва монотонно возрастающей функции $h: \boldsymbol{R} \to \boldsymbol{R}$ не более чем счетно.***

▷ Пусть K — множество точек разрыва функции h, $x \in K$. Множество $\{h(t) \mid x < t\}$ ограничено снизу, например числом $h(x)$. Поэтому существует число $b_x \rightleftarrows \inf\limits_{t>x} h(t)$. Аналогичным образом существует число $a_x \rightleftarrows \sup\limits_{t<x} h(t)$. Очевидно, $a_x \leqslant b_x$. Точке x поставим в соответствие интервал $]a_x, b_x[$. Из монотонности h следует, что интервалы $]a_x, b_x[$ и $]a_{x'}, b_{x'}[$ дизъюнктивны при $x \neq x'$. Так как x — точка разрыва функции h, то $a_x \neq b_x$, а именно, $a_x < b_x$. Поэтому $]a_x, b_x[$ — непустой интервал. Итак, множеству K биективно соответствует некоторое семейство попарно не пересекающихся интервалов на оси Oy. Такое семейство не более чем счетно. Поэтому и множество K не более чем счетно. ■

Упражнение 10. *Докажите, что: а) если u_1 и u_2 прямоугольники, стороны которых соответственно равны по длине, то u_2 получается из u_1 с помощью композиции параллельного переноса и поворота;*

б) если s — функция площади и u_1, u_2 — прямоугольники из пункта а), то $s(u_1) = s(u_2)$.

Лемма 4. ***Пусть $s: U \to \boldsymbol{R}_{>0}$ — некоторая функция площади. Тогда найдется такая постоянная $c > 0$, для которой $s(u_{x,y}) = c \cdot x \cdot y$, где $u_{x,y}$ любой прямоугольник, у которого стороны имеют длину соответственно x и y.***

▷ По упражнению 10 б) можно считать, что $u_{x,y}$ — прямоугольник вида (рис. 60).

Обозначим $s(u_{x,y})$ в виде $s(x, y)$. Фиксируем произвольное значение $y_0 > 0$ и рассмотрим функцию $h(x) \rightleftarrows s(x, y_0): \boldsymbol{R}_{>0} \to \boldsymbol{R}_{>0}$. Из конечной аддитивности площади s и ее инвариантности относительно параллельных переносов следует, что $h(x_1+x_2) = h(x_1) + h(x_2)$. Поскольку $h(x_2) > 0$, то получаем, что функция h строго возрастает. Так как множество $\boldsymbol{R}_{>0}$ несчетно, то по лемме 3 функция h имеет хотя бы одну точку непрерывности, которую обозначим x_0. Докажем, что функция h непрерывна в любой точке вида x_0+t, где t — произвольное число. Действительно, $h(x_0+t+\Delta x) - h(x_0+t) = h(x_0+\Delta x) - h(x_0) \to 0$ при $\Delta x \to 0$. Следовательно, функция h непрерывна на всем множестве $\boldsymbol{R}_{>0}$ (рис. 61).

Допустим, что $h(0) \rightleftarrows \inf\limits_{0<x} h(x) = \varepsilon > 0$. Тогда $h(1) = nh\left(\frac{1}{n}\right) > n\varepsilon \to \infty$ при $n \to \infty$. Поэтому $h(0) = 0$.

Продолжим h на множество $\boldsymbol{R}_{<0}$ по нечетности. Получим непрерывный гомоморфизм вида $h: \boldsymbol{R}^+ \to \boldsymbol{R}^+$. По теореме 1 из главы I

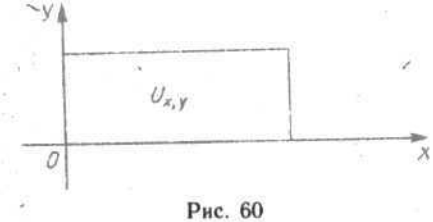


Рис. 60

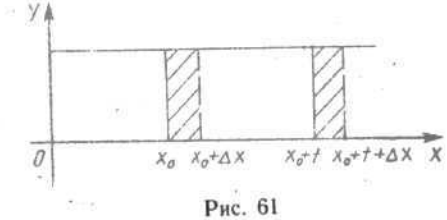


Рис. 61

получаем, что h — линейная функция, т. е. $h(x)=ax$, где $a \rightleftarrows h(1)=s(1, y_0)$.

Число a, вообще говоря, зависит от фиксированного выше числа y_0. Поэтому $a \doteq a(y_0)$. Положим $c=s(1, 1)$. Тогда $s(x, y)=a(y)\cdot x$ и по упражнению 4г) $a(y)=s(1, y)=s(y, 1)$. Поэтому $a(y)=a(1)\cdot y=s(1, 1)\cdot y=c\cdot y$ и $s(x, y)=c\cdot x\cdot y$. ■

Теорема 3. ***Фиксируем многоугольник $u_0 \in U$ и число $\lambda_0>0$. Существует и притом ровно одна функция площади $s: U \to \boldsymbol{R}_{>0}$, для которой $s(u_0)=\lambda_0$.***

▷ По теореме 2 из предыдущего параграфа функция $T(\cdot)$ является площадью. Если $T(u_0) \neq \lambda_0$, то рассмотрим функцию $s(u) \rightleftarrows \frac{\lambda_0}{T(u_0)} T(u)$. Тогда $s(\cdot)$ — функция площади, для которой $s(u_0)=\lambda_0$. Это доказывает существование искомой функции.

Пусть функции $s_1(\cdot)$ и $s_2(\cdot)$ являются функциями площади. По лемме 4 $s_1(u_{x, y})=c_1\cdot x\cdot y$ и $s_2(u_{x, y})=c_2\cdot x\cdot y$. Разбивая $u_{x, y}$ диагональю на два прямоугольных треугольника и используя конечную аддитивность и инвариантность функции площади относительно поворота на угол 180°, получаем, что значение функции $s_i(\cdot)$ для любого прямоугольного треугольника равно половине произведения постоянной c_i на произведение длин катетов. Здесь $i=1, 2$. Разбивая любой треугольник Δ высотой на два прямоугольных треугольника, получаем, что значение функции $s_i(\cdot)$ равно половине произведения постоянной c_i на произведение длины высоты и длины стороны треугольника, на которую опущена высота (рис. 62). Следовательно, $s_1(\Delta)=\frac{c_1}{c_2}\cdot s_2(\Delta) \rightleftarrows c\cdot s_2(\Delta)$, где $c \rightleftarrows \frac{c_1}{c_2}$. Пусть теперь $s_1(u_0)=s_2(u_0)=\lambda_0>0$ и $\Delta_1, \ldots, \Delta_n$ — некоторое разложение u_0. По конечной аддитивности функций s_1 и s_2 получаем $s_1(u_0)=s_1(\Delta_1+\ldots+\Delta_n)=s_1(\Delta_1)+\ldots+s_1(\Delta_n)=c\cdot s_2(\Delta_1)+c\cdot s_2(\Delta_2)+\ldots+c\cdot s_2(\Delta_n)=c\cdot(s_2(\Delta_1)+\ldots+s_2(\Delta_n))=c\cdot s_2(u_0)=c\cdot\lambda_0$.
Отсюда следует, что $c=1$ и тем самым функции s_1 и s_2 тождественно совпадают. ■

Замечание. Часто в качестве многоугольника u_0 выбирают квадрат со стороной, длина которой равна заранее выбранной единице измерения длин отрезков в $\boldsymbol{R}$, и полагают λ_0 равным 1. В этом смысле выбор некоторого отрезка в $\boldsymbol{R}$ в качестве масштаба длины однозначно определяет функцию площади в $\boldsymbol{R}^2$. В этом смысле измерение длин отрезков определяет измерение площадей многоугольников.

Следствие 1. ***Если две функции площади совпадают хотя бы***

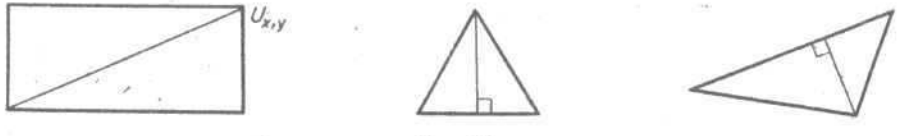


Рис. 62

на одном многоугольнике, то они совпадают на всех многоугольниках, т. е. тождественно равны. Любая функция площади имеет вид $c\cdot T(\cdot)$.

▷ Первое утверждение непосредственно вытекает из теоремы 3. Если s функция площади и u_0 любой фиксированный многоугольник, то выберем постоянную $c>0$ таким образом, что $s(u_0)=c\cdot T(u_0)$. Тогда по первому утверждению функции s и $c\cdot T$ тождественно совпадают. ■

Следствие 2. ***Если s_1 и s_2 две произвольные функции площади, то существует постоянная c, для которой $s_1(\cdot)\equiv c\cdot s_2(\cdot)$.***

Следствие 3. ***Пусть H_k — гомотетия с коэффициентом $k\neq 0$. Всякая площадь обладает свойством гомотетичности: $\forall u \in U$***

$$s(H_k(u))=k^2\cdot s(u).$$

▷ По следствию 1 $s(\cdot)\equiv c\cdot T(\cdot)$. Так как функция T обладает свойством гомотетичности (следствие 3 к теореме 2), то и s обладает этим свойством. ■

2. Определение площади многоугольника с помощью движений. В этом пункте рассмотрим вопрос о возможных переформулировках свойства G_s-инвариантности. А именно, вопрос о возможности заменить группу G_s на ее «маленькую» подгруппу, состоящую из всех движений в плоскости $\boldsymbol{R}^2$.

Множество движений $J(\boldsymbol{R}^2)$ определяется как совокупность всех пар вида $\langle A, \xi\rangle$, где $A \in O_2$ и $\xi \in \boldsymbol{R}^2$, т. е.

$$J(\boldsymbol{R}^2) \rightleftarrows O_2\times \boldsymbol{R}^2.$$

Напомним, что O_2 — это множество всех ортогональных (2×2) — матриц над полем $\boldsymbol{R}$. Групповая операция в O_2 определяется обычным образом как произведение двух матриц. В множестве $J(\boldsymbol{R}^2)=O_2\times\boldsymbol{R}^2$ определим групповую операцию по формуле

$$\langle A_1, \xi_1\rangle\cdot\langle A_2, \xi_2\rangle \rightleftarrows \langle A_1\cdot A_2, A_1\cdot\xi_2+\xi_1\rangle.$$

Следовательно, $J(\boldsymbol{R}^2)$ — полупрямое произведение групп O_2 и $(\boldsymbol{R}^2)^+$ и, в частности, подгруппа группы G_s.

Пусть теперь $g=\langle A, \xi\rangle \in O_2\times\boldsymbol{R}^2$. Как и в случае группы G_s, определим функцию $f_g: \boldsymbol{R}^2\to\boldsymbol{R}^2$ формулой $f_g(x) \rightleftarrows A\cdot x+\xi$ и определим отображение $F(g) \rightleftarrows f_g$; $F: O_2\times\boldsymbol{R}^2\to \mathrm{Is}(\boldsymbol{R}^2)$. Легко проверить, что отображение F инъективно и, следовательно, множество $O_2\times\boldsymbol{R}^2$ биективно со множеством функций $\{f_g \mid g\in O_2\times\boldsymbol{R}^2\}$. Каждая функция f_g также является биекцией, но другого вида, чем F.

Отображение $F: O_2\times\boldsymbol{R}^2\to \mathrm{Is}(\boldsymbol{R}^2)$ является, очевидно, инъективным гомоморфизмом, т. е. $F(g_1)\circ F(g_2)=F(g_1\cdot g_2)$. Следовательно, отображение $F: O_2\times\boldsymbol{R}^2 \leftrightarrow R(F)$ является изоморфизмом. Обычно отождествляют группы $O_2\times\boldsymbol{R}^2$ и $R(F)$.

Упражнение 11. *Докажите, что: а) O_2 и $J(\boldsymbol{R}^2)$ — группы и $J(\boldsymbol{R}^2)$ — подгруппа группы G_s;*

б) $J(\boldsymbol{R}^2)$ содержит все параллельные переносы, осевые симметрии и собственные повороты (все элементы группы SO_2, см. § 7 главы I).

Каждая биекция f_g переводит любую прямую снова в некоторую прямую и сохраняет расстояние (метрику) в $\boldsymbol{R}^2$. Действительно, $\|f_g(x)-f_g(y)\|^2=\|A\cdot x+\xi-A\cdot y-\xi\|^2=\|A(x-y)\|^2=\|x-y\|^2$.

Можно доказать, что любая биекция $\psi:\boldsymbol{R}^2\leftrightarrow\boldsymbol{R}^2$, которая сохраняет расстояние (метрику) в $\boldsymbol{R}^2$ (т. е. $\forall x,\ y\in\boldsymbol{R}^2\,(\rho(x,\ y)=\rho(\psi(x),\ \psi(y)))$, где $\rho(x,\ y)=\sqrt{(x-y,\ x-y)}=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2}$) имеет вид

$$\psi(x)=A\cdot x+\xi,$$

где $\xi\in\boldsymbol{R}^2$ и $A\in O_2$ и, в частности, она переводит любую прямую снова в прямую. Поэтому любая такая биекция ψ вида $\psi=f_g$, где $g\in J(\boldsymbol{R}^2)$, и, следовательно, $R(F)$ состоит из всех преобразований плоскости, сохраняющих метрику. Отсюда происходит название группы $J(\boldsymbol{R}^2)$. Иногда такая характеристика множества $R(F)$ принимается за определение группы движений; поскольку при этом используется только метрическая структура в $\boldsymbol{R}^2$, то это определение годится для любого метрического пространства.

Как мы увидим ниже, свойства конечной аддитивности и G_s-инвариантности эквивалентны свойствам конечной аддитивности и $J(\boldsymbol{R}^2)$-инвариантности. Мы предпочитаем говорить именно о G_s-инвариантности функции площади. Во-первых, потому что G_s — наибольшая из групп, сохраняющих функцию площади (см. следствие 2 к теореме 2). Во-вторых, определение группы G_s не использует понятие скалярного произведения в $\boldsymbol{R}^2$ и, следовательно, в определенном смысле основано на меньшем числе исходных понятий по сравнению с определением группы $J(\boldsymbol{R}^2)$.

Можно описать некоторую «количественную» разницу между группами O_2 и L_2 и, следовательно, между группами $J(\boldsymbol{R}^2)$ и G_s. Всякий элемент A из O_2 имеет определитель, равный $+1$ или -1 (см. § 7 главы I). В первом случае $A\in\mathrm{SO}_2$, а во втором случае $A=A'\circ E'$, где $A'\in\mathrm{SO}_2$ и $E'=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}$. Следовательно, множество O_2 является дизъюнктным объединением вида $\mathrm{SO}_2\cup\mathrm{SO}_2'$, где множество $\mathrm{SO}_2'=\{A\cdot E'\mid A\in\mathrm{SO}_2\}$.

Последнее множество в топологии, индуцированной O_2, гомеоморфно SO_2. Отсюда легко проверить, что O_2 гомеоморфно множеству S^2 с топологией произведения в нем.

Как мы видели в § 7 главы I, всякий элемент $A\in\mathrm{SO}_2$ представим в виде $\begin{pmatrix}\cos t & -\sin t\\ \sin t & \cos t\end{pmatrix}$ и, следовательно, зависит от одного произвольно выбираемого числового параметра $t\in[0,\ 2\pi[$. Поэтому всякий элемент A из O_2 однозначно соответствует паре $\langle\det A,\ t\rangle$, где $t\in[0,\ 2\pi[$. Первый элемент этой пары дискретен (принимает только два значения). Пренебрегая дискретным параметром, естественно сказать, что группа O_2 «одномерна». В то же время группа L_2 «трехмерна», так как на четыре числа, составляющих матрицу A из L_2, имеется одно ограничение вида $\det A=\pm1$. Иными словами, можно произвольно задать три числовых параметра и тогда оставшийся четвертый параметр определится по этим трем из уравнения $\det A=\pm1$. Правда, это уравнение имеет два решения и, следовательно, четвертый параметр определится по трем другим опять-таки с точностью до дискретного параметра. Пренебрегая дискретным параметром, мы говорим, что группа L_2 трехмерна. В этом смысле O_2 — маленькая подгруппа группы L_2.

Группы $J(\boldsymbol{R}^2)$ и G_s получаются из O_2 и L_2 одной и той же добавкой: умножением на $(\boldsymbol{R}^2)^+$. Последняя группа двумерна. Поэтому группа $J(\boldsymbol{R}^2)$ трехмерна (задается тремя произвольными параметрами), а группа G_s пятимерна (задается пятью произвольными параметрами). В этом смысле $J(\boldsymbol{R}^2)$ — маленькая подгруппа группы G_s.

Отметим, что группа $(\boldsymbol{R}^2)^+$ изоморфна подгруппе группы $J(\boldsymbol{R}^2)$. А именно: $(\boldsymbol{R}^2)^+$ изоморфна подгруппе $\{\langle E,\ \xi\rangle\mid\xi\in\boldsymbol{R}^2\}$ с помощью функции $\xi\leftrightarrow\langle E,\ \xi\rangle$. Эта подгруппа двумерна и в этом смысле является маленькой подгруппой в группах $J(\boldsymbol{R}^2)$ и G_s.

В доказательстве теоремы 3 не используется в полном объеме инвариантность функции площади s. Действительно, в лемме 4 используется инвариантность s относительно параллельных переносов и поворотов для доказательства того, что $s(u_{x,\ y})$ не зависит от расположения прямоугольника $u_{x,\ y}$ со сторонами x и y. В самой же теореме 3 используется инвариантность s относительно центральной симметрии и некоторых движений плоскости (при разбиении произвольного треугольника на прямоугольные треугольники и рассмотрении произвольных прямоугольных треугольников). Поэтому, если $f:U\to\boldsymbol{R}_{>0}$ конечно-аддитивная и $J(\boldsymbol{R}^2)$ — инвариантная функция, то для нее остаются верными доказательства теоремы 3 и ее следствия 2, т. е. $f(\cdot)\equiv c\cdot T(\cdot)$. Так как функция T G_s-инвариантна, то и функция f G_s-инвариантна и, следовательно, является функцией площади. Отсюда получаем

Следствие 4 к теореме 3. ***Всякая функция вида $s:U\to\boldsymbol{R}_{>0}$ со свойствами конечной аддитивности и $J(\boldsymbol{R}^2)$-инвариантности обладает свойством G_s-инвариантности, следовательно, является функцией площади.***

3. Способы измерения площади многоугольника.

В этом пункте мы рассмотрим четыре способа измерения площади многоугольника, которые соответственно определяют четыре функции $T_1(\cdot)-T_4(\cdot)$, и докажем, что все они приводят к одинаковому численному результату, несмотря на явные различия в их определениях (инструкциях для измерения). А именно, каждая из этих четырех функций совпадает с функцией T. В случае функции T_1 и T_2 такое совпадение легко получить непосредственно на основе их определения. В случае функций T_3 и T_4 необходимо использовать теорему 3 о единственности функции площади: мы докажем, что функции T_3 и T_4 являются функциями площади и совпадают с T на некотором многоугольнике (а именно, на треугольнике для

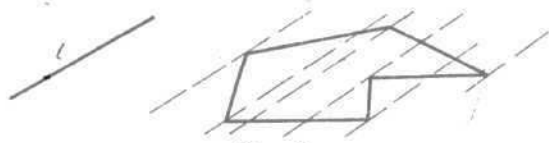


Рис. 63

случая T_3 и на параллелограмме для случая T_4). Отсюда по следствию 1 вытекает совпадение функций T_3 и T_4 с функцией T.

Определение 3. **Фиксируем произвольную прямую l в плоскости R^2** (рис. 63). **Через каждую вершину многоугольника u проведем прямую, параллельную l. В результате u разобьется на фигуры, которые разлагаются на конечное число трапеций. Обозначим b_i и B_i — длины оснований и h_i — длину высоты i-ой трапеции. Положим**

$$T_1(u) \rightleftarrows \frac{1}{2} \sum_i (b_i + B_i) h_i.$$

В этом определении мы рассматриваем в качестве трапеции и треугольник, считая, что длина одного из оснований равна нулю. Разобьем i-ую трапецию диагональю на два треугольника Δ_{i_1} и Δ_{i_2}.

Упражнение 12. *Докажите, что площадь T треугольника равна половине произведения длины основания на длину высоты, опущенной на это основание.*

По этому упражнению $\frac{1}{2}(b_i + B_i)h_i = T(\Delta_{i_1}) + T(\Delta_{i_2})$ и $T_1(u)$ совпадает с $\sum_i (T(\Delta_{i_1}) + T(\Delta_{i_2}))$. Поскольку треугольники Δ_{11}, Δ_{12}, ... в совокупности образуют некоторое разложение многоугольника u, то $T_1(u) = T(u)$. Тем самым доказано

Предложение 5. ***Функция T_1 является функцией площади, которая тождественно совпадает с T.***

Как следствие предложения 5, получаем независимость значений функции T_1 от выбора прямой l, что довольно сложно доказать непосредственно.

Идея следующего способа измерения площади многоугольника состоит в следующем: в плоскости фиксируется некоторый луч l и некоторый отрезок t, выходящий из вершины луча и не лежащий на нем (например, в случае арифметической плоскости R^2 — это луч $R_{\geqslant 0}$ и отрезок $\left[\binom{0}{0}, \binom{0}{1}\right]$); для произвольного многоугольника u и какого-нибудь его разложения на треугольники по каждому из этих треугольников образуем равновеликий параллелограмм, у которого одна сторона лежит на луче l, а другая совпадает с отрезком t; затем сдвинем эти параллелограммы вдоль луча l таким образом, чтобы они пересекались не более чем по граничному отрезку и тем самым образуем общий параллелограмм $\square(u)$, у

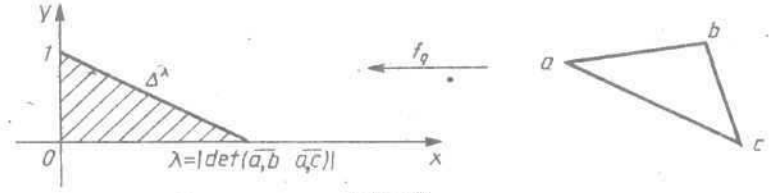


Рис. 64

которого одна сторона t_1 лежит на луче l, а другая совпадает с отрезком t. Теперь, чтобы измерить площадь исходного многоугольника u, нужно измерить только длину отрезка t_1. Кажется, что задача измерения длины более простая, чем задача измерения площади. При этом важно, что переход от каждого треугольника к соответствующему параллелограмму и от исходного многоугольника u — к параллелограмму $\square(u)$ происходит «чисто геометрическим» способом. Мы не будем уточнять слова, взятые в кавычки, а проведем все построение в арифметической плоскости R^2. Начнем, как всегда, с треугольника.

Пусть Δ — треугольник с вершинами в точках a, b, c. Обозначим Δ^λ — треугольник с вершинами в точках (0, 0), (0, 1) и (λ, 0), где $\lambda \rightleftarrows 2 \cdot T(\Delta) = |\det(\overrightarrow{a,\ b}\quad \overrightarrow{a,\ c})|$. По предложению 2 существует элемент g группы G_s, для которого $f_g(\Delta) = \Delta^\lambda$ (рис. 64).

Отметим, что такой элемент $g = \langle A, \xi \rangle$ можно выбирать многими способами. Например, можно положить $\xi \rightleftarrows -b$ и $A \rightleftarrows \left(\frac{1}{\lambda} \cdot \overrightarrow{b,\ a}\quad \overrightarrow{b,\ c}\right)$. Следовательно, и преобразование f_g, переводящее Δ в Δ^λ, может быть выбрано многими способами.

Упражнение 13. *Укажите вид всех таких элементов $g = \langle A, \xi \rangle \in G_s$, для которых $f_g(\Delta) = \Delta^\lambda$.*

Теперь замени треугольник Δ^λ прямоугольником $u_{\lambda/2}$ с вершинами в точках (0, 0), (0, 1), $\left(\frac{\lambda}{2}, 0\right)$, $\left(\frac{\lambda}{2}, 1\right)$ (рис. 65). Ясно, что $T(u_{\frac{\lambda}{2}}) = T(\Delta^\lambda) = T(\Delta)$.

Пусть Δ_1, ..., Δ_n — некоторое разложение исходного многоугольника u. Каждый из треугольников Δ_i заменим на прямоугольник $u_{\lambda_i/2} \rightleftarrows u_i$. Полученные прямоугольники сдвинем по оси абсцисс

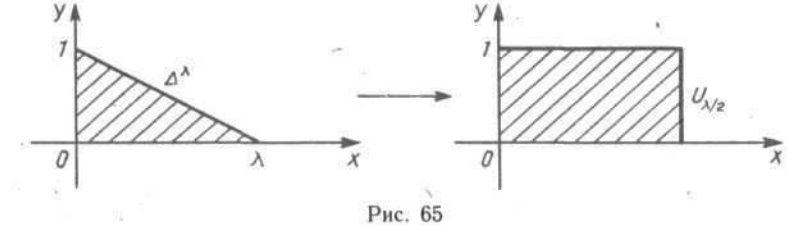


Рис. 65

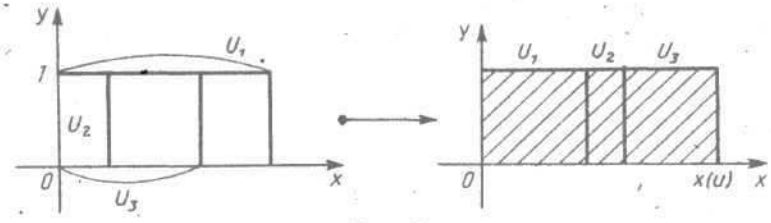


Рис. 66

друг за другом и получим один прямоугольник $\square(u)$ с вершинами в точках (0, 0), (0, 1), $(x(u), 0)$, $(x(u), 1)$ (рис. 66).

Определение 4. **Пусть u — многоугольник. Определим $T_2(u)$ *как число* $x(u)$, построенное вышеуказанным способом.**

Из построения $x(u)$ непосредственно следует, что функция T_2 совпадает с функцией T, поэтому верно следующее

Предложение 6. ***Функция T_2 является функцией площади, которая тождественно совпадает с T.***

Чтобы перейти к очередному способу измерения площади многоугольника, нам потребуются некоторые, наглядно очевидные свойства многоугольников. Тем не менее их доказательства довольно сложны и у нас нет места, чтобы их привести.

Предложение 7. ***а) Граница любого многоугольника является объединением конечного числа отрезков, которые (без их концов) попарно не пересекаются и не содержат вершин многоугольника. (Эти отрезки условно назовем сторонами многоугольника*** (рис. 67). Все отрезки здесь и далее подразумеваются невырожденными.)

б) Если c внутренняя точка любой стороны t многоугольника u, то найдется такой открытый круг $\mathcal{O}_c$ с центром в точке c, что один из полукругов $\mathcal{O}_c \setminus t$ состоит только из внутренних точек u, а другой не пересекается с u (рис. 68).

в) Множество всех сторон многоугольника можно разбить на конечное число подмножеств, каждое из которых обладает следующим свойством: концы всех входящих в него отрезков, кроме одного, можно пронумеровать таким образом, что концы каждого отрезка имеют соседние номера.

(Объединение отрезков, входящих в любое из этих подмножеств, называется простой ломаной. Пункт в) говорит, что граница любого многоугольника является объединением конечного числа простых ломаных).

г) Пусть u и v многоугольники без общих внутренних точек. Все стороны многоугольников u и v можно разложить на отрезки, которые в совокупности попарно не пересекаются (здесь и далее без учета их концов); любой из этих отрезков либо содержится в u и не пересекается с v, либо содержится в v и не пересекается с u, либо содержится в $u \cap v$; отрезки первых двух типов разлагают любую сторону многоугольника $u+v$.

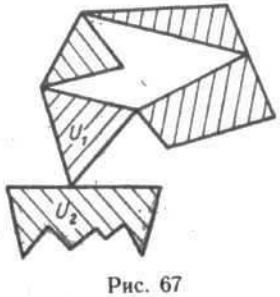


Рис. 67

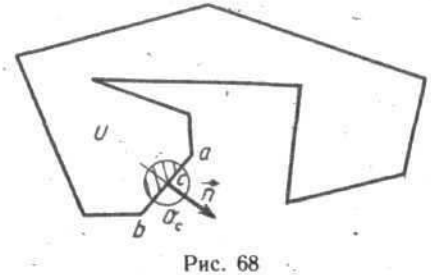


Рис. 68

Границу многоугольника u будем обозначать $\partial(u)$.

Пусть c любая внутренняя точка стороны t многоугольника u. Через точку c проходит перпендикуляр к стороне t. Точка c разбивает этот перпендикуляр на два открытых луча. По предложению 7 б) пересечение ровно одного из этих двух лучей с открытым кругом $\mathcal{O}_\varepsilon$ состоит из точек, не лежащих в u. Вектор n единичной длины, лежащий на этом луче, с началом в точке c называется нормальным вектором к стороне t многоугольника u (рис. 68).

Упражнение 14. *Докажите, что нормальный вектор $\bar{n}$ к стороне многоугольника не зависит от выбора внутренней точки c на этой стороне.*

Определение 5. **Фиксируем какую-то точку c в R^2. Выберем на каждой из прямых, проходящих через стороны многоугольника u, по произвольной точке a_i. Пусть m число сторон многоугольника u, $1 \leqslant i \leqslant m$. Положим $T_3(u) = \frac{1}{2}\sum_{i=1}^{m} (\overline{c, a_i}, \bar{n}_i) \cdot l_i$, где $\bar{n}_i$ — нормальный вектор к i-ой стороне, l_i — длина i-ой стороны и $(\cdot, \cdot)$ — скалярное произведение в R^2.**

Предложение 8. ***Функция T_3 обладает следующими свойствами:***

а) ее значения не зависят от выбора точек c и a_i;

б) она является функцией площади;

в) она совпадает с функцией T.

▷ а) Если a_i' — какие-то другие точки на прямых, определяемых сторонами многоугольника u, то при фиксированном c получаем $\sum_i (\overline{c, a_i'}, \bar{n}_i) \cdot l_i = \sum_i (\overline{c, a_i}, \bar{n}_i) l_i + \sum_i (\overline{a_i, a_i'}, \bar{n}_i) \cdot l_i = \sum_i (\overline{c, a_i}, \bar{n}_i) \cdot l_i$, так как по определению $\bar{n}_i$ перпендикулярно a_i, a_i'.

Фиксируем теперь все точки a_i и заменим точку c на какую-то точку c'. Тогда $\sum_i (\overline{c', a_i}, \bar{n}_i) \cdot l_i = \sum_i (\overline{c, a_i}, \bar{n}_i) \cdot l_i + \sum_i (\overline{c, c'}, \bar{n}_i) \cdot l_i$. Докажем, что второе слагаемое равно нулю. Оно, очевидно, равно $(\overline{c', c}, \sum_i l_i \cdot \bar{n}_i)$. Поэтому достаточно показать, что $\sum_i l_i \cdot \bar{n}_i = 0$.

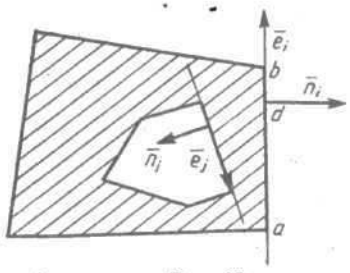


Рис. 69

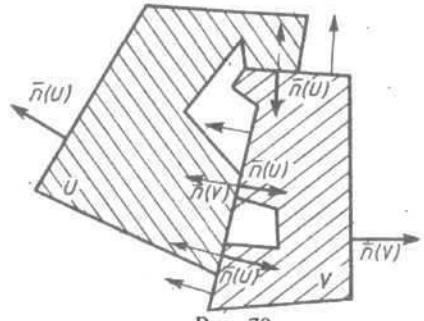


Рис. 70

Вместе с нормальными к i-ой стороне многоугольника u векторами $\bar{n}_i$ рассмотрим вектора $\bar{e}_i$, которые получаются из $\bar{n}_i$ поворотом на угол $\frac{\pi}{2}$, т. е. $\bar{e}_i \rightleftarrows W \cdot \bar{n}_i$, где $W \rightleftarrows \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Выберем любую внутреннюю точку d на i-ой стороне многоугольника u и отложим от нее вектор e_i. Конец вектора e_i лежит на прямой, проходящей через i-ую сторону многоугольника, а точка d разбивает эту прямую на два луча, в каждом из которых лежит ровно по одной граничной точке i-ой стороны многоугольника u (рис. 69). Концом i-ой стороны назовем ее граничную точку, лежащую на том из этих двух лучей, который содержит конец вектора $\bar{e}_i$. Началом i-ой стороны назовем вторую из ее граничных точек.

Легко видеть, что если a — начало, а b — конец i-ой стороны многоугольника u, то векторы $\overline{a, b}$ и $\bar{e}_i$ коллинеарны. Так как l_i — длина i-ой стороны, то $a, b = l_i \cdot \bar{e}_i$.

По предложению 7 в) разобьем границу исходного многоугольника на простые ломаные и пусть $a_1, \ldots a_n$ соответствующая нумерация вершин некоторой простой ломаной.

Каждый из векторов $\overline{a_1, a_2}$, $\overline{a_2, a_3}$, $\overline{a_3, a_4}$, ..., $\overline{a_n, a_1}$ совпадает с одним из векторов $l_i \cdot \bar{e}_i$. Поскольку $\overline{a_1, a_2} + \ldots + \overline{a_n, a_1} = 0$, то часть всей суммы $\sum_i l_i \bar{e}_i$, относящаяся к этой простой ломаной, равна нулю. Так как каждая из сторон многоугольника u входит ровно в одну из этих простых ломаных, то получаем

$$\sum_i l_i \bar{e}_i = 0.$$

В силу линейности функции W и равенства $\bar{e}_i = W \cdot \bar{n}_i$ получаем:

$$W \cdot (\sum_i l_i \cdot \bar{n}_i) = \sum_i l_i W \cdot \bar{n}_i = \sum_i l_i \cdot \bar{e}_i = 0.$$

Так как функция W является биекцией, то $\Sigma l_i \cdot \bar{n}_i = 0$.

б) Докажем конечную аддитивность функции T_3. Пусть многоугольники u и v не имеют общих внутренних точек (рис. 70). Обозначим $t_1, \ldots, t_n$ последовательность отрезков, о которой говорится в предложении 7 г). Рассмотрим сумму

$$\frac{1}{2}(\sum_i \overline{(c, a_i, \bar{n}_i)} \cdot l_i + \sum_j \overline{(c, b_j, \bar{n}_j)} \cdot l_j). \qquad (*)$$

Каждое слагаемое в этой сумме соответствует стороне k одного из многоугольников u и v, а сторона k разлагается некоторыми отрезками $t_{i_1}, \ldots, t_{i_p}$; обозначим их длины соответственно $l_{i_1}, \ldots, l_{i_p}$. Стороне k соответствует слагаемое вида $\overline{(c, a, \bar{n})} \cdot l$, равное $\overline{(c, a, \bar{n})} \cdot l_{i_1} + \ldots + \overline{(c, a, \bar{n})} \cdot l_{i_p}$, так как по упражнению 14 нормальный вектор $\bar{n}$ к стороне k является и нормальным вектором к каждой из сторон $t_{i_1}, \ldots, t_{i_p}$. Подставим в выражение $(*)$ все такие суммы. Если отрезок t_i принадлежит пересечению границ $\partial(u) \cap \partial(v)$, то в первой и второй сумме в выражении $(*)$ по одному разу используется сторона t_i. При этом многоугольники u и v расположены по разные стороны от нее, т. е. внешние нормали к этой стороне противоположны. Поэтому соответствующие слагаемые в первой и второй сумме сокращаются. Остается сумма по всем отрезкам из числа $t_1, \ldots, t_n$, которые входят только в $\partial(u)$ или только в $\partial(v)$. По предложению 7 г) эти отрезки разлагают границу $\partial(u+v)$. Слагаемые, относящиеся к одной стороне многоугольника $u+v$, имеют одинаковые нормали, и отсюда получаем, что оставшаяся сумма равна $T_3(u+v)$. (Рассуждение, которое мы привели в этом пункте, апеллирует к интуиции. Его строгое изложение не просто).

Так как ортогональная матрица A из O_2 сохраняет скалярное произведение, то число $T_3(f_A(u))$, если выбрать в качестве c точку $f_A(c)$, совпадает с числом $T_3(u)$, подсчитанным относительно точки c. Поэтому T_3 инвариантна относительно всех преобразований плоскости вида f_A, где $A \in O_2$.

Инвариантность функции T_3 относительно параллельных переносов по существу доказана в пункте а). Действительно, пусть $c \in \boldsymbol{R}^2$ и $\xi \in \boldsymbol{R}^2$ и $c' \rightleftarrows c + \xi$. Выберем произвольно точки a_i на сторонах многоугольника u и положим $a_i' \rightleftarrows a_i + \xi$. Тогда точки a_i будут лежать на сторонах многоугольника $u+\xi$ и $\overline{c', a_i'} = \overline{c, a_i}$. Так как вектор $\bar{n}_i$ при параллельном переносе с помощью вектора ξ не меняется, то значение $T_3(u)$ относительно точки c и точек a_i совпадает со значением $T_3(u+\xi)$ относительно точки c' и точек a_i'. Поэтому $T_3(u) = T_3(u+\xi)$. Отсюда по следствию 4 к теореме 3 получаем, что T_3 — функция площади.

в) Поскольку доказано, что функция T_3 является функцией площади, то по теореме 3 достаточно проверить совпадение функций T_3 и T на каком-нибудь одном треугольнике Δ. Например, рассмотрим треугольник Δ с вершинами в точках (0, 0), (0, 1), (1, 0). Точку c положим равной (0, 1) и положим $a_1 \rightleftarrows (1, 0)$, $a_2 \rightleftarrows a_3 \rightleftarrows c$. Тогда $\overline{c, a_2} = \overline{c, a_3} = 0$, $\bar{n}_1 = \overline{c, a_1}$, $\overline{(c, a_1, \bar{n}_1)} = 1$, $l_1 = 1$. Поэтому $T_3(\Delta) = \frac{1}{2}(\overline{(c, a_1, \bar{n}_1)} \cdot l_1 + \overline{(c, a_2, n_2)} \cdot l_2 + \overline{(c, a_3, n_3)} \cdot l_3) =$

$= \frac{1}{2} = T(\Delta)$. ■

Следующее определение измерения площади важно по

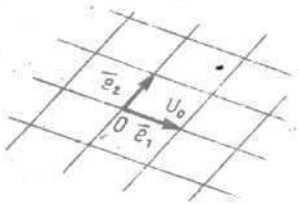


Рис. 71

двум причинам. Во-первых, в нем множество $\boldsymbol{R}^2$ рассматривается только как векторное пространство: не используются ни длина (метрика), ни скалярное произведение в $\boldsymbol{R}^2$; например, оно без всяких изменений применимо к векторной и после незначительной переработки — к аффинной плоскости. Во-вторых, возникающая функция T_4 может быть продолжена на более широкую область определения, чем U, с сохранением ее основных свойств — конечной аддитивности и G_s-инвариантности. В то же время для функций T_1—T_3 совсем не видно, как осуществить такое продолжение.

Перейдем к определению этого способа измерения и соответствующей ему функции T_4. Для этого выберем в $\boldsymbol{R}^2$ два произвольных линейно независимых вектора $\bar{e}_1$ и $\bar{e}_2$, т. е. выберем базис $\{\bar{e}_1, \bar{e}_2\}$. Базисным назовем параллелограмм u_0 с вершинами в точках o, $\bar{e}_1$, $\bar{e}_2$, $\bar{e}_1+\bar{e}_2$ (рис. 71). Разложим $\boldsymbol{R}^2$ на параллелограммы, являющиеся параллельными переносами базисного параллелограмма u_0 с помощью всевозможных векторов вида $n\cdot\bar{e}_1+m\cdot\bar{e}_2$, где $n,\ m\in\boldsymbol{Z}$. Все полученные параллелограммы назовем ячейками нулевого ранга.

Далее, разложим $\boldsymbol{R}^2$ на параллелограммы, являющиеся параллельными переносами параллелограмма u_1 с вершинами в точках $\bar{o}$, $\frac{1}{10}\bar{e}_1$, $\frac{1}{10}\bar{e}_2$, $\frac{1}{10}\bar{e}_1+\frac{1}{10}\bar{e}_2$ с помощью всевозможных векторов вида $n\cdot\frac{1}{10}\bar{e}_1+m\cdot\frac{1}{10}\bar{e}_2$, где $n,\ m\in\boldsymbol{Z}$. Все полученные параллелограммы назовем ячейками первого ранга. Параллелограмм u_2 является в геометрическом смысле сотой частью базисного параллелограмма u_0. Продолжая таким образом, для каждого числа i получим разложение $\boldsymbol{R}^2$ на ячейки i-го ранга, являющиеся параллельными переносами параллелограмма u_i с вершинами в точках $\bar{o}$, $\frac{1}{10^i}\bar{e}_1$, $\frac{1}{10^i}\bar{e}_2$, $\frac{1}{10^i}\bar{e}_1+\frac{1}{10^i}\bar{e}_2$ с помощью всевозможных векторов все того же вида $n\cdot\frac{1}{10^i}\bar{e}_1+m\cdot\frac{1}{10^i}\bar{e}_2$, где $n,\ m\in\boldsymbol{Z}$. (Переносы выполняются все время целочисленными векторами вида $n\cdot\xi+m\cdot\eta$, где ξ, η два фиксированных вектора.)

Пусть $X\subseteq\boldsymbol{R}^2$ и n_i — число ячеек i-го ранга, целиком содержащихся в множестве X, а N_i — число ячеек i-го ранга, пересекающихся с множеством X. Каждая из ячеек i-го ранга, целиком содержащаяся в X, состоит из ста ячеек $(i+1)$-го ранга, целиком содержащихся в X. Поэтому $n_{i+1}\geqslant 100\cdot n_i$ и, следовательно, $n_0\leqslant\frac{n_1}{100^1}\leqslant\frac{n_2}{100^2}\leqslant\frac{n_3}{100^3}\leqslant\dots$. Аналогично $N_0\geqslant\frac{N_1}{100^1}\geqslant\frac{N_2}{100^2}\geqslant\frac{N_3}{100^3}\geqslant\dots$. Так как $N_i\geqslant n_i$, то $\frac{N_i}{100^i}\geqslant\frac{n_j}{100^j}$ при всех $i,\ j\in N$.

Действительно, если $i\geqslant j$, то $\frac{N_i}{100^i}\geqslant\frac{n_i}{100^i}\geqslant\frac{n_j}{100^j}$, а если $i<j$, то $\frac{N_i}{100^i}\geqslant\frac{N_j}{100^j}\geqslant\frac{n_j}{100^j}$.

Определение 6. **Фиксируем положительное число s_0. Верхней площадью фигуры X назовем число или знак $+\infty$, равные $s^*(X)\rightleftarrows s_0\cdot\lim\limits_{i\to\infty}\frac{N_i}{100^i}$. Нижней площадью фигуры X назовем число или знак $+\infty$, равные $s_*(X)\rightleftarrows s_0\cdot\lim\limits_{i\to\infty}\frac{n_i}{100^i}$. Фигуру $X\subseteq R^2$ назовем измеримой по Жордану, если ее верхняя и нижняя площадь совпадают. В этом случае их общее значение обозначим $T_4(X)$.**

Замечания. 1) Часто функцию T_4 определяют с помощью ортонормированного базиса $\{\bar{e}_1, \bar{e}_2\}$, что уже влечет наличие евклидовой структуры в плоскости.

2) Иногда функцию T_4 строят, используя понятие площади многоугольника. Тогда нижнюю площадь фигуры X определяют как точную верхнюю грань площадей любых, содержащихся в X многоугольников, а верхнюю площадь — как точную нижнюю грань площадей любых, содержащих X многоугольников. Легко доказать, что для ограниченных фигур таким образом получается та же самая функция T_4.

3) Выбор числа s_0 является по существу выбором масштаба, так как значение функции T_4 на базисном параллелограмме u_0, как мы увидим ниже, равно s_0; иными словами, базисному параллелограмму произвольно приписывается площадь s_0.

4) Для любой фигуры X верхняя и нижняя площади существуют, так как числовая последовательность $\left\{\frac{n_i}{100^i}\right\}$ возрастает, а числовая последовательность $\left\{\frac{N_i}{100^i}\right\}$ убывает и ограничена снизу числом 0.

5) Приведем пример ограниченной и не измеримой по Жордану фигуры X. Пусть $u_0=[0,\ 1]\times[0,\ 1]$. Положим

$$X\rightleftarrows\{r_1\cdot\bar{e}_1+r_2\cdot\bar{e}_2\mid r_1,\ r_2\in(\boldsymbol{Q}\cap(0,\ 1))\}.$$

Тогда нижняя площадь фигуры X равна нулю, так как никакая ячейка целиком не содержится в X, т. е. все $n_i=0$. Любая из ячеек i-го ранга, содержащаяся в u_0, пересекается с X, и других ячеек i-го ранга, пересекающихся с X, нет. Поэтому $N_i=100^i$. Следовательно, верхняя площадь X равна $s_0>0$.

Явное вычисление функции T_4 для многих конкретных фигур довольно сложно. Приведем несколько примеров такого вычисления.

а) Докажем, что значение функции T_4 на параллелограмме u с вершинами в точках $a\cdot\bar{e}_1+b\cdot\bar{e}_2$, $a\cdot\bar{e}_1+B\cdot\bar{e}_2$, $A\cdot\bar{e}_1+B\cdot\bar{e}_2$, $A\cdot\bar{e}_1+b\cdot\bar{e}_2$, где $B\geqslant b$ и где $A\geqslant a$, равна $s_0\cdot(A-a)(B-b)$.

▷ Рассмотрим «сетку», состоящую из ячеек i-го ранга (рис. 72).

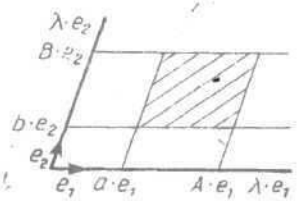


Рис. 72

В отрезке $[a\cdot\bar{e}_1, A\cdot\bar{e}_1]$ целиком содержится a_i каких-то отрезков, полученных в результате пересечения этой сетки с прямой $\{\lambda\cdot\bar{e}_1 \mid \lambda\in\boldsymbol{R}\}$. С тем же отрезком пересекается A_i таких отрезков. Аналогичным образом определим натуральные числа b_i и B_i для отрезка $[b\cdot\bar{e}_2, B\cdot\bar{e}_2]$. Тогда $n_i=a_i\cdot b_i$ и $N_i=A_i\cdot B_i$ и $A_i\leqslant a_i+2$, $B_i\leqslant b_i+2$. Для краткости записи предположим, что длина векторов $\bar{e}_1$ и $\bar{e}_2$ равна 1. Тогда

$$\frac{a_i}{10^i}\leqslant A-a\leqslant\frac{A_i}{10^i}\ \text{и}\ \frac{b_i}{10^i}\leqslant B-b\leqslant\frac{B_i}{10^i}.$$

Следовательно, $\frac{n_i}{100^i}=\frac{a_i}{10^i}\cdot\frac{b_i}{10^i}\leqslant(A-a)\cdot(B-b)\leqslant\frac{A_i}{10^i}\cdot\frac{B_i}{10^i}=\frac{N_i}{100^i}$ и

$$\frac{N_i-n_i}{100^i}=\frac{A_i\cdot B_i-a_i\cdot b_i}{100^i}\leqslant\frac{(a_i+2)\cdot(b_i+2)-a_i\cdot b_i}{100^i}=\frac{2}{10^i}\left(\frac{a_i}{10^i}+\frac{b_i}{10^i}+\frac{2}{10^i}\right)\leqslant$$
$$\leqslant\frac{2}{10^i}((A-a)+(B-b)+2)\to0\ \text{при}\ i\to\infty.$$

Отсюда получаем $\lim\limits_{i\to\infty}s_0\cdot\frac{n_i}{100^i}=\lim\limits_{i\to\infty}s_0\cdot\frac{N_i}{100^i}=s_0\cdot(A-a)\cdot(B-b)$. ■

б) Значение функции T_4 на любом одноточечном множестве равно нулю.

▷ В этом случае все $n_i=0$ и $\frac{N_i}{100^i}\leqslant\frac{4}{100^i}$. Следовательно, верхняя и нижняя площади одноточечного множества равны нулю. ■

в) Значение функции T_4 на любом отрезке равно нулю.

▷ Проведем через концы отрезка прямые, параллельные векторам $\bar{e}_1$ и $\bar{e}_2$. Тем самым отрезок будет диагональю параллелограмма u с вершинами в точках $a\cdot\bar{e}_1+b\cdot\bar{e}_2$, $a\cdot\bar{e}_1+B\cdot\bar{e}_2$, $A\cdot\bar{e}_1+B\cdot\bar{e}_2$, $A\cdot\bar{e}_1+b\cdot\bar{e}_2$.

Если $X_1\subseteq X_2$, то ясно, что верхняя площадь множества X_1 меньше или равна верхней площади множества X_2. Поэтому верхняя площадь отрезка не больше $s_0\cdot(A-a)\cdot(B-b)$. Разобьем средними линиями параллелограмм u на четыре параллелограмма. Тогда отрезок содержится в объединении двух параллелограммов и его верхняя площадь не превосходит числа $2\cdot\frac{A-a}{2}\cdot\frac{B-b}{2}s_0=\frac{s_0}{2}\cdot(A-a)\cdot(B-b)$. Продолжая такие деления средними линиями, получим, что верхняя площадь отрезка не превосходит числа $2^i\cdot\frac{A-a}{2^i}\cdot\frac{B-b}{2^i}\cdot s_0=\frac{s_0}{2^i}\cdot(A-a)\cdot(B-b)$. Устремляя i к ∞, получим, что верхняя площадь отрезка равна 0. Для отрезка, параллельного $\bar{e}_1$ или $\bar{e}_2$, доказательство аналогично. ■

Упражнение 15. *Докажите, что:* а) *Значение функции T_4 на множестве, состоящем из конечного числа любых точек, равно нулю.*

б) *Если множество $X_1\cap X_2$ состоит из конечного числа точек и $T_4(X_1)=T_4(X_2)=0$, то $T_4(X_1\cup X_2)=0$.*

в) *Значение $T_4(\cdot)$ на объединении конечного числа отрезков равно нулю.* Указание: *по индукции с учетом пункта б).*

г) *Если $X_1\subseteq X_2$, то $s^*(X_1)\leqslant s^*(X_2)$ и $s_*(X_1)\leqslant s_*(X_2)$.*

Предложение 9. ***а) Область определения функции T_4 содержит множество всех многоугольников U и ее сужение на U является функцией площади. При некотором выборе числа s_0 выполняется равенство***

$$T_4\restriction U=T.$$

б) Ограниченная фигура X измерима по Жордану тогда и только тогда, когда

$$\forall\varepsilon>0\exists u_1,\ u_2\in U(u_1\subseteq X\subseteq u_2)\wedge(T_4(u_2)-T_4(u_1)<\varepsilon).$$

▷ а) Очевидно, ограниченная фигура X измерима по Жордану тогда и только тогда, когда $\frac{N_i(X)-n_i(X)}{100^i}\to0$ при $i\to\infty$.

Ячейки i-го ранга, входящие в число N_i, но не входящие в число n_i, пересекаются с границей многоугольника, т. е. число N_i-n_i меньше, чем число ячеек i-го ранга, пересекающихся с конечным числом отрезков, составляющих границу многоугольника. Поэтому $\frac{N_i-n_i}{100^i}$ стремится к верхней мере этого объединения конечного числа отрезков, равной, согласно упражнения 15, в) нулю.

Проверим конечную аддитивность функции T_4. Пусть u и v многоугольники без общих внутренних точек. Ясно, что множество ячеек i-го ранга, пересекающихся с $u+v$, равно объединению двух множеств: ячеек i-го ранга, пересекающихся с u и пересекающихся с v. Поэтому $N_i(u+v)=N_i(u)+N_i(v)-r_i$, где r_i — число ячеек i-го ранга, пересекающихся и с u, и с v. При достаточно большом i число r_i совпадает с числом ячеек, пересекающихся с границей u или v, т. е. $\frac{r_i}{100^i}\to0$ при $i\to\infty$. Отсюда $\lim\limits_{i\to\infty}\frac{N_i(u+v)}{100^i}=$

$\lim\limits_{i\to\infty}\frac{N_i(u)}{100^i}+\lim\limits_{i\to\infty}\frac{N_i(v)}{100^i}$, т. е. $T_4(u+v)=T_4(u)+T_4(v)$.

Пусть $\xi\in\boldsymbol{R}^2$ и $u\in U$. Рассмотрим сетку, полученную в результате параллельного переноса с помощью вектора ξ сетки i-го ранга. Обозначим N_i' и n_i' — число ячеек этой сетки, соответственно пересекающихся и содержащихся в множестве $u+\xi$. Тогда $N_i=N_i'$ и $n_i=n_i'$. Согласно примеру а) функция T_4 конечно-аддитивна на прямоугольниках со сторонами, параллельными векторам $\bar{e}_1$ и $\bar{e}_2$, и при параллельном переносе ячеек любого ранга значения на них этой функции не меняются; поэтому значение T_4 на многоугольнике, образованном ячейками исходной сети, содержащимися в u, равно значению T_4 на многоугольнике, образованном ячейками новой сетки, содержащимися в $u+\xi$. С учетом предложения 9в получаем $T_4(u)=T_4(u+\xi)$.

Свойство G_s-инвариантности вытекает из инвариантности относительно преобразований вида $f_{\langle A,\,\xi\rangle}(x) \rightleftarrows A\cdot x+\xi$, где $A\in O_2$. Однако проверить инвариантность T_4 относительно ортогональных матриц — не простая задача. В случае сетки, у которой $e_1 \rightleftarrows \binom{1}{0}$, $e_2 \rightleftarrows \binom{0}{1}$ и $s_0 \rightleftarrows 1$, такая O_2-инвариантность будет доказана (см. § 5, следствие 2 к предложению 14) с использованием принципиально нового понятия меры. А именно, мы докажем, что мера является O_2-инвариантной функцией, которая на множестве многоугольников совпадает с функцией T_4.

б) Пусть фигура X измерима по Жордану и пусть i таково, что $T_4(X) - s_0\frac{n_i}{100^i} < \frac{\varepsilon}{2}$, и u_1 — многоугольник, который является объединением ячеек i-го ранга, целиком содержащихся в X. Пусть j таково, что $s_0\cdot\frac{N_j}{100^j} - T_4(X) < \frac{\varepsilon}{2}$, и многоугольник u_2 является объединением ячеек j-го ранга, пересекающихся с X. Тогда выполняется

$$u_1\subseteq X\subseteq u_2 \text{ и } T_4(u_2)-T_4(u_1)=s_0\frac{N_j}{100^j}-s_0\frac{n_i}{100^i}<\varepsilon.$$

Наоборот, пусть многоугольники u_1 и u_2 таковы, что $u_1\subseteq X\subseteq u_2$ и $T_4(u_2)-T_4(u_1)<\varepsilon$. Так как по пункту а) все многоугольники измеримы по Жордану, то $T_4(u_2)=s^*(u_2)$ и $T_4(u_1)=s_*(u_1)$. Следовательно, $T_4(u_1)=s_*(u_1)\leqslant s_*(X)\leqslant s^*(u_2)=T_4(u_2)$ и поэтому $s^*(X)-s_*(X)<\varepsilon$. Так как последнее неравенство верно при всех $\varepsilon>0$, то получаем: $s^*(X)=s_*(X)$. Поэтому фигура X измерима по Жордану. ■

Замечание. Функцию T_4 можно было бы использовать в качестве конечно-аддитивной функции из предложения 4 а) и тем самым доказать независимость определения числа $T(u)$ от выбора разложения многоугольника на треугольники.

Аналогичным образом можно использовать функцию T_3.

§ 5. СРАВНЕНИЕ КОНСТРУКТИВНОГО И АКСИОМАТИЧЕСКОГО ОПРЕДЕЛЕНИЙ МЕРЫ ПЛОСКОЙ ФИГУРЫ. ВЫЧИСЛЕНИЕ МЕРЫ ПРОСТЕЙШИХ КРИВОЛИНЕЙНЫХ ФИГУР

1. Измерение плоских криволинейных фигур. Теперь мы хотим определить такую процедуру измерения плоских фигур, которая была бы применима не только для многоугольников, но и для криволинейных фигур, по крайней мере таких, как круг, сектор, сегмент и т. п.

Необходимость и возможность такого измерения были осознаны очень давно.

Аналогично функции площади T конструктивно определим функцию M; на множестве всех многоугольников она будет совпадать с

T, т. е. T — сужение M на U и M — продолжение T с U на некоторое широкое семейство $\mathscr{L}$ плоских фигур. По основным свойствам функция M будет похожа на функцию T.

В этом параграфе условимся называть прямоугольником множество в $\boldsymbol{R}^2$, определяемое одним из неравенств вида

$$\begin{array}{l} a\leqslant x\leqslant b,\\ a\leqslant x<b,\\ a<x\leqslant b,\\ a<x<b \end{array}$$

и одним из неравенств вида

$$\begin{array}{l} c\leqslant y\leqslant d,\\ c\leqslant y<d,\\ c<y\leqslant d,\\ c<y<d. \end{array}$$

Ясно, что стороны всякого прямоугольника параллельны координатным осям. Прямоугольник может быть открытым или замкнутым или не открытым и не замкнутым множеством относительно обычной топологии в $\boldsymbol{R}^2$. Комбинируя одно из неравенств первого типа с одним из неравенств второго типа, получаем 16 типов прямоугольников. Любой замкнутый прямоугольник является многоугольником.

Назовем элементарной фигурой (элементарным множеством) любую фигуру в $\boldsymbol{R}^2$, которая представима в виде объединения конечного числа прямоугольников. Обозначим U_1 семейство всех элементарных фигур. Ясно, что замыкание $\bar{u}$ любой непустой элементарной фигуры u является многоугольником и множество $(\bar{u}\setminus u)$ является объединением конечного числа отрезков. В этом смысле любая элементарная фигура — почти многоугольник.

Упражнение 16. *Докажите, что объединение, пересечение и разность двух элементарных фигур является элементарной фигурой. В частности, любая элементарная фигура — дизъюнктное объединение прямоугольников.*

Обозначим $\square_n$ фиксированный квадрат в плоскости $\boldsymbol{R}^2$ с вершинами в точках $\binom{-n}{-n}, \binom{-n}{n}, \binom{n}{-n}, \binom{n}{n}$, где $n\in\boldsymbol{N}$.

Определение 7. а) **Обозначим f функцию, которая прямоугольнику с вершинами $\binom{a}{c}, \binom{a}{d}, \binom{b}{c}, \binom{b}{d}$ сопоставляет число $(b-a)\cdot(d-c)$.**

Внешней мерой в плоскости $\boldsymbol{R}^2$ назовем отображение вида

$$\mu^*:\mathscr{P}(\boldsymbol{R}^2)\to\boldsymbol{R}\cup\{+\infty\},$$

определяемое формулой:

$$\mu^*(X)\rightleftarrows\inf\Big\{\sum_{n=1}^{\infty}f(u_n)\mid X\subseteq\bigcup_{n=1}^{\infty}u_n,\ u_n\text{ — прямоугольники}\Big\}.$$

Семейство прямоугольников $\{u_n\}$ и соответствующее множество $\bigcup\limits_n u_n$ называется покрытием исходной фигуры X. Если в определении μ^* использовать какой-то один вид прямоугольников, например только замкнутые прямоугольники, то получится такая же функция μ^* (проверьте).

б) Фигуру (множество) X, содержащуюся в $\square_n$ при некотором значении n, назовем измеримой по Лебегу (измеримой), если для любого положительного числа ε найдется такая элементарная фигура u, что множество $X \triangle u \rightleftarrows (X \setminus u) \cup (u \setminus X)$ имеет внешнюю меру, меньшую ε, т. е. $\mu^*(X \triangle u) < \varepsilon$.

Здесь $X \triangle Y \rightleftarrows (X \setminus Y) \cup (Y \setminus X)$ — симметрическая разность.

в) Фигуру (множество) X в R^2 назовем измеримой, если для любого числа $n \in N$ фигура $X \cap \square_n$ измерима. Семейство всех измеримых фигур (измеримых множеств) обозначим $\mathscr{L}$.

Объединение семейства $\{X_\alpha\}$ попарно не пересекающихся множеств будем обозначать $\bigsqcup\limits_\alpha X_\alpha$. Такое объединение называется дизъюнктным.

Упражнение 17. *Докажите, что: а) Если X — любая прямая, то $\mu^*(X)=0$; если X — любая полоса ширины ε, где $\varepsilon \in R_{>0}$, и бесконечной длины, то $\mu^*(X)=+\infty$.*

б) Любая элементарная фигура и любое объединение счетного числа прямоугольников являются измеримым множеством.

в) Если фигура X лежит в квадратах $\square_n$ и $\square_m$, то измеримость X, как части $\square_n$, эквивалентна измеримости X, как части $\square_m$. Поэтому в определении 7б) можно использовать любой квадрат $\square_n$, содержащий X. Для любой ограниченной фигуры найдется такой квадрат, поэтому для нее определение 7б) не зависит от выбора содержащего ее квадрата.

г) Если фигура X, содержащаяся в квадрате $\square_n$, измерима, то и пересечение $X \cap \square_m$ также измеримо, где n, m — любые натуральные числа.

д) Ограниченная фигура измерима в смысле определения 7б) в том и только том случае, если она измерима в смысле определения 7в); поэтому для ограниченных фигур определения 7б) и 7в) совпадают; следовательно, определение 7в) является обобщением определения 7б) со случая ограниченных фигур на случай произвольных фигур.

Теорема 4. ***Внешняя мера любого прямоугольника равна произведению длин его сторон.***

▷ Пусть u — прямоугольник. Он покрывает сам себя и поэтому $\mu^*(u) \leqslant f(u)$. Допустим, что $\mu^*(u) < f(u)$. Тогда существует покрытие $\{u_n\}$ прямоугольника u, для которого $\sum\limits_n f(u_n) < f(u)$.

Немного увеличив прямоугольники u_n, можно заменить их открытыми прямоугольниками u_n', для которых $\sum\limits_n f(u_n') < f(u)$, а немного уменьшив прямоугольник u, можно заменить его замкнутым прямоугольником u', для которого $\sum\limits_n f(u_n') < f(u')$. Из открытого покрытия $\{u_n'\}$, компакта u', выберем конечное подпокрытие. Еще уменьшив, заменим его на конечное дизъюнктное покрытие $\{v_k\}$, содержащееся в u'. Итак,

$$\bigsqcup_{k=1}^{m} v_k = u' \text{ и } \sum_{k=1}^{m} f(v_k) < f(u').$$

Через все стороны всех прямоугольников v_k проведем прямые. Получится сетка в u.

Семейство всех прямоугольников, определяемое сеткой, обозначим $\{v_e'\}$. Тогда

$$\bigsqcup_e v_e' = u' \text{ и } \sum_e f(v_e') < f(u'),$$

так как каждое v_k разобьется на конечное число прямоугольников, для которых сумма значений f, как проверяется прямым элементарным подсчетом, равна $f(v_k)$. С другой стороны, тем же элементарным подсчетом получим $\sum\limits_e f(v_e') = f(u')$ — противоречие. ■

Теорема 5. ***а) Если $X \subseteq \bigcup\limits_n X_n$, где $X \subseteq R^2$ и X_n — произвольное конечное или счетное семейство множеств в R^2, то***

$$\mu^*(X) \leqslant \sum_{n=1}^{\infty} \mu^*(X_n). \qquad (*)$$

Если сумма в правой части имеет хотя бы одно слагаемое, равное $+\infty$, то она по определению равна $+\infty$. Свойство $(*)$ внешней меры μ^* называется счетной полуаддитивностью.

б) Если измеримые множества X_1 и X_1 содержатся в квадрате $\square_n$ и не пересекаются, то $\mu^*(X_1 \cup X_2) = \mu^*(X_1) + \mu^*(X_2)$. Это свойство μ^* называется конечной аддитивностью.

в) Для любых множеств X и Y, если $X \subseteq Y$, то $\mu^*(X) \leqslant \mu^*(Y)$. Это свойство μ^* называется монотонностью.

г) Для любого множества X

$$\mu^*(X) = \lim_{k \to \infty} \mu^*(X \cap \square_k).$$

▷ а) По определению μ^* для каждого множества X_n найдется такое его покрытие прямоугольниками $\{u_{n_k}\}$, что

$$X_n \subseteq \bigcup_k u_{n_k} \text{ и } \sum_k \mu^*(u_{n_k}) \leqslant \mu^*(X_n) + \frac{\varepsilon}{2^n},$$

где ε — любое положительное число. Тогда $X \subseteq \bigcup\limits_{n,\,k} u_{n_k}$ и $\mu^*(X) \leqslant \sum\limits_n \sum\limits_k \mu^*(u_{n_k}) \leqslant \sum\limits_n \mu^*(X_n) + \varepsilon$. В силу произвольности ε получаем требуемое неравенство.

б) По пункту а) достаточно доказать неравенство:

$$\mu^*(X_1)+\mu^*(X_2)\leqslant\mu^*(X_1\cup X_2).$$

Для любого $\varepsilon>0$ выберем такие элементарные фигуры Y_1 и Y_2, что $\mu^*(X_1\triangle Y_1)<\varepsilon$ и $\mu^*(X_2\triangle Y_2)<\varepsilon$. Обозначим $X=X_1\cup X_2$ и $Y=Y_1\cup Y_2$. Здесь Y — элементарная фигура, так как Y_1 и Y_2 таковы. Так как $X_1\cap X_2=\varnothing$ и $Y_1\cap Y_2\subseteq(X_1\triangle Y_1)\cup(X_2\triangle Y_2)$, то $\mu^*(Y_1\cap Y_2)\leqslant 2\varepsilon$. Для дальнейшего доказательства нужна следующая

Лемма 5. ***Для любых фигур X и Y выполняется***

$$|\mu^*(X)-\mu^*(Y)|\leqslant\mu^*(X\triangle Y).$$

▷ Пусть $\mu^*(X)\geqslant\mu^*(Y)$ (в противном случае поменяем местами X и Y). Так как $X\subseteq(Y\cup(X\triangle Y))$, то по пункту а) получаем $\mu^*(X)\leqslant\leqslant\mu^*(Y)+\mu^*(X\triangle Y)$, что дает искомое неравенство. ■

Продолжим доказательство и по лемме 5 получаем $|\mu^*(X)-\mu^*(Y)|\leqslant\mu^*(X\triangle Y)\leqslant 2\varepsilon$, где последнее неравенство следует из включения

$$X\triangle Y\subseteq(X_1\triangle Y_1)\cup(X_2\triangle Y_2).$$

Аналогичным образом при $i=1, 2$ получаем

$$|\mu^*(Y_i)-\mu^*(X_i)|\leqslant\mu^*(X_i\triangle Y_i)\leqslant\varepsilon$$

и, следовательно,

$$|\mu^*(Y_1)+\mu^*(Y_2)\geqslant(\mu^*(X_1)+\mu^*(X_2))|\leqslant 2\varepsilon.$$

Для продолжения доказательства нужна еще одна

Лемма 6. ***Функция μ^* на множестве всех элементарных фигур U_1 обладает свойством аддитивности:***

$$\forall\, u,\ v\in U_1\,(\mu^*(u\cup v)+\mu^*(u\cap v)=\mu^*(u)+\mu^*(v)).$$

▷ Доказательство непосредственно вытекает из того, что значение μ^* на элементарной фигуре равно сумме значений функции f, т. е. сумме обычных площадей дизъюнктных прямоугольников, составляющих эту фигуру (проверьте это подробно). ■

Теперь по лемме 6 получаем

$$\mu^*(Y)=\mu^*(Y_1\cup Y_2)=\mu^*(Y_1)+\mu^*(Y_2)-\mu^*(Y_1\cap Y_2)$$

и окончательно получаем, что

$$|\mu^*(X)-(\mu^*(X_1)+\mu^*(X_2))|\leqslant 6\varepsilon.$$

Следовательно, $\mu^*(X)\geqslant\mu^*(X_1)+\mu^*(X_2)$.

в) Так как любое покрытие множества Y является и покрытием множества X, то свойство монотонности непосредственно следует из определения верхней меры.

г) Так как $X\cap\square_k\subseteq X\cap\square_{k+1}$ и $X\cap\square_k\subseteq X$ для любого $k\in\boldsymbol{N}$, то по свойству монотонности получаем: числовая последовательность $\mu^*(X\cap\square_k)$ возрастает и ограничена сверху $\mu^*(X)$.

Поэтому предел, входящий в проверяемое равенство, существует; обозначим его t, и следует только показать, что $\mu^*(X)\leqslant t$. Выберем такие покрытия Y_k, что $X\cap\square_k\subseteq Y_k$ и $\mu^*(Y_k)\leqslant\mu^*(X\cap\square_k)+\varepsilon$,

где ε — положительное число. Можно считать, что $Y_k\subseteq\square_k$ (иначе заменим Y_k на $Y_k\cap\square_k$), и последовательность $\{Y_k\}$ возрастает. Тогда

$$X\subseteq\bigcup_k Y_k=\bigcup_k Y_k',\ \text{где } Y_k'\rightleftarrows(Y_k\setminus Y_{k-1}),\ Y_1'\rightleftarrows Y_1.$$

По пункту а) теоремы

$$\mu^*(X)\leqslant\sum_{k=1}^{\infty}\mu^*(Y_k')=\lim_{m\to\infty}\sum_{k=1}^{m}\mu^*(Y_k')=\lim_{m\to\infty}(\mu^*(Y_1)+(\mu^*(Y_2)-$$
$$-\mu^*(Y_1))+\ldots+(\mu^*(Y_m)-\mu^*(Y_{m-1})))=$$
$$=\lim_{m\to\infty}\mu^*(Y_m)\leqslant\lim_{m\to\infty}(\mu^*(X\cap\square_m)+\varepsilon),$$

где мы использовали пункт б) этой теоремы. Поэтому $\mu^*(X)\leqslant t-\varepsilon$ и $\mu^*(X)\leqslant t$. ■

Замечание. Термины «фигура» и «подмножество плоскости $\boldsymbol{R}^2$» означают (формально говоря) одно и то же. В геометрическом контексте, в частности в связи с темой многоугольников, чаще употребляют первый термин; в связи с мерой чаще употребляют второй термин. Поэтому в дальнейшем будем говорить о множествах, имея в виду подмножества в $\boldsymbol{R}^2$ (фигуры).

Семейство множеств называется кольцом множеств (кольцом), если оно замкнуто относительно операций объединения и разности. Любое кольцо множеств замкнуто относительно операций пересечения и симметрической разности.

По упражнению 16 семейство всех элементарных фигур является кольцом множеств.

Семейство множеств называется σ-кольцом, если оно — кольцо множеств, которое замкнуто относительно операции счетного объединения. Из соотношения $\bigcap_n X_n=\bigcup_n X_n\setminus(\bigcup_n(\bigcup_n X_n)\setminus X_n)$ вытекает, что любое σ-кольцо замкнуто относительно операции счетного пересечения.

Следующее определение является основным в этом параграфе.

Определение 8. **Сужение функции μ^* на семейство $\mathscr{L}$ всех измеримых множеств называется мерой Лебега и обозначается** M, т. е. **для всех** X из $\mathscr{L}$

$$M(X)\rightleftarrows\mu^*(X).$$

Ясно, что $M:\mathscr{L}\to\boldsymbol{R}_{\geqslant 0}\cup\{+\infty\}$.

Теорема 6. ***а) Семейство $\mathscr{L}$ всех измеримых множеств является σ-кольцом.***

б) Мера Лебега M обладает свойством счетной аддитивности:

$$\forall X\in\mathscr{L}\,\forall\{X_n\}\subseteq\mathscr{L}\,(X=\bigsqcup_{n=1}^{\infty}X_n\Rightarrow M(X)=\sum_{n=1}^{\infty}M(X_n)).$$

Последнее равенство понимается так: если одна из его частей равна $+\infty$, то и другая часть равна $+\infty$; если одна из его частей равна числу, то и другая часть равна тому же числу.

▷ Отметим следующее свойство операции симметрической разности

$X \triangle Y$ двух множеств X и Y: каждое из трех множеств $(X_1 \cup X_2) \triangle (Y_1 \cup Y_2)$, $(X_1 \cap X_2) \triangle (Y_1 \cap Y_2)$, $(X_1 \setminus X_2) \triangle (Y_1 \setminus Y_2)$ содержится в множестве $(X_1 \triangle Y_1) \cup (X_2 \triangle Y_2)$. Доказательство этих включений проводится обычным теоретико-множественным рассуждением.

а) Сначала проверим, что $\mathscr{L}$ кольцо. Пусть X_1 и X_2 — измеримые множества, содержащиеся в каком-то квадрате $\square_n$. При любом фиксированном $\varepsilon > 0$ найдутся элементарные множества Y_1 и Y_2, для которых

$$\mu^*(X_i \triangle Y_i) < \frac{\varepsilon}{2}.$$

Тогда $Y_1 \setminus Y_2$ также элементарное множество, и по свойству полуаддитивности функции μ^* получаем

$$\mu^*((X_1 \setminus X_2) \triangle (Y_1 \setminus Y_2)) \leqslant \mu^*(X_1 \triangle Y_1) + \mu^*(X_2 \triangle Y_2) < \varepsilon.$$

Следовательно, множество $X_1 \setminus X_2$ измеримо.

Если X_1 и X_2 произвольные измеримые множества, то, используя равенство $(X_1 \setminus X_2) \cap \square_n = (X_1 \cap \square_n) \setminus (X_2 \cap \square_n)$, получим, что множество $X_1 \setminus X_2$ измеримо. Аналогично проверяется замкнутость двух множеств семейства $\mathscr{L}$ относительно операции объединения. Итак, семейство $\mathscr{L}$ является кольцом.

Докажем часть пункта б) данной теоремы, а именно — конечную аддитивность функции M, т. е. для множеств X_1, $X_2 \in \mathscr{L}$ проверим соотношение $M(X_1 \sqcup X_2) = M(X_1) + M(X_2)$. При этом мы используем, что для ограниченных множеств X_1 и X_2 это соотношение уже доказано в предыдущей теореме 5 б). По теореме 5 получаем

$$M(X \sqcup Y) = \lim_n M((X \sqcup Y) \cap \square_n) = \lim_n M((X \cap \square_n) \cup (Y \cap \square_n)) =$$
$$= \lim_n (M(X \cap \square_n) + M(Y \cap \square_n)) = \lim_n M(X \cap \square_n) + \lim_n M(Y \cap \square_n) =$$
$$= M(X) + M(Y).$$

Теперь мы можем завершить доказательство пункта а). Пусть даны произвольные измеримые множества $X_1, X_2, \ldots$. Докажем измеримость множества $X \rightleftarrows \bigcup_{n=1}^{\infty} X_n$. Для этого достаточно из измеримости множеств $(X_n \cap \square_m)$ вывести измеримость множества $X \cap \square_m = \bigcup (X_n \cap \square_m)$. Поэтому будем исходить сразу из предположения, что все множества X_n лежат в некотором $\square_m$. Представим X в виде объединения $\bigcup_{n=1}^{\infty} X'_n$ попарно не пересекающихся измеримых множеств $X'_n \rightleftarrows X_n \setminus \bigcup_{k=1}^{n-1} X_k$. Тогда $\sum_{k=1}^{n} M(X'_k) = M\left(\bigcup_{k=1}^{n} X'_k\right) = M\left(\bigcup_{k=1}^{n} X_k\right) =$ $= \mu^*\left(\bigcup_{k=1}^{n} X_k\right) \leqslant \mu^*(X) \leqslant \mu^*(\square_m) < +\infty$ и, следовательно, ряд $\sum_{k=1}^{\infty} M(X'_k)$ сходится. Для любого $\varepsilon > 0$ выберем натуральное число m, для которого $\sum_{n=m+1}^{\infty} M(X'_n) < \frac{\varepsilon}{2}$, т. е. $\mu^*\left(\bigcup_{n=m+1}^{\infty} X'_n\right) < \frac{\varepsilon}{2}$. Затем выберем

элементарное множество Y, для которого $\mu^*\left(\left(\bigcup_{k=1}^{m} X'_n\right) \triangle Y\right) < \frac{\varepsilon}{2}$.

Так как $\left(\bigcup_{n=1}^{\infty} X'_n\right) \triangle Y \subseteq \left(\left(\bigcup_{n=1}^{m} X'_n\right) \triangle Y\right) \cup \left(\bigcup_{n=m+1}^{\infty} X'_n\right)$,

то $\mu^*(X \triangle Y) < \varepsilon$. Поэтому множество X измеримо.

б) Наконец, докажем счетную аддитивность меры M.

Пусть измеримые множества X_n, где $n \in \boldsymbol{N}$, попарно не пересекаются. По пункту а) этой теоремы множество $\bigcup_n X_n$ измеримо. Тогда по конечной аддитивности и монотонности функции M получаем

$$M\left(\bigcup_{n=1}^{m} X_n\right) = \sum_{n=1}^{m} M(X_n) \leqslant M\left(\bigcup_{n=1}^{\infty} X_n\right).$$

Переходя к пределу при $m \to \infty$, получаем

$$\sum_{n=1}^{\infty} M(X_n) \leqslant M\left(\bigcup_{n=1}^{\infty} X_n\right).$$

С другой стороны, по доказанной счетной полуаддитивности функции μ^* получаем

$$M\left(\bigcup_{n=1}^{\infty} X_n\right) = \mu^*\left(\bigcup_{n=1}^{\infty} X_n\right) \leqslant \sum_{n=1}^{\infty} \mu^*(X_n) = \sum_{n=1}^{\infty} M(X_n).$$

Следовательно, $M\left(\bigcup_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} M(X_n)$. ■

Теорема 7. *а) Любой многоугольник является измеримым множеством, т. е. $U \subseteq \mathscr{L}$.*

б) Для любого треугольника $\Delta = \{a, b, c\}$ выполняется

$$M(\Delta) = \frac{1}{2} \left| \det \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \end{pmatrix} \right|.$$

в) Пусть и — многоугольник и $\{\Delta_1, \ldots, \Delta_n\}$ — его разложение на треугольники. Тогда число $T(u) \rightleftarrows T(\Delta_1) + \ldots + T(\Delta_n)$ не зависит от выбора разложения и на треугольники и, следовательно, определяется конечно-аддитивная функция $T: U \to R_{>0}$.

г) Функция T — сужение меры M на семейство U, т. е. $T = M \restriction U$ и M — продолжение T с семейства U на семейство $\mathscr{L}$.

▷ а) Сначала рассмотрим прямоугольный треугольник с катетами длины a_0 и b_0, параллельными координатным осям; для определенности допустим, что он расположен, как показано на рисунке 73 а). Разделим один из его катетов, например параллельный оси ординат, на k равных частей, где $k \in \boldsymbol{N}$. Проведем через точки деления прямые, параллельные оси абсцисс, и образуем семейство из k прямоугольников, покрывающее Δ. Обозначим это покрытие Y_k. Тогда $Y_k \triangle \Delta$ содержится в объединении Y'_k заштрихованных



Рис. 73

на этом рисунке прямоугольников, сумма их верхних мер меньше, чем $\underbrace{\left(\frac{b_0}{k}\cdot\frac{a_0}{k}\right)+\ldots+\left(\frac{b_0}{k}\cdot\frac{a_0}{k}\right)}_{k\text{ слагаемых}}$, т. е. она меньше, чем $\frac{a_0\cdot b_0}{k^2}\cdot k$, которое стремится к 0 при $k\to\infty$. Последнее означает, что треугольник Δ измерим.

Если Δ произвольный треугольник, то разложим его на n прямоугольных треугольников со сторонами, параллельными координатным осям, и еще один треугольник Δ_{n+1} (на рисунке 73, *б* показано такое разложение для $n=4$). Пусть переменная n пробегает только четные натуральные числа. Образуем прямоугольник u_n с диагональю $[a, b]$. Ясно, что $\Delta_{n+1}\subseteq u_n$ и произведение длин сторон прямоугольника u_n стремится к нулю при $n\to\infty$. Для каждого из первых n прямоугольных треугольников Δ_i выберем такое покрытие Y_i прямоугольниками, которое описано выше. Как и выше, $Y_i\setminus\Delta_i$ содержится в конечном объединении Y_i' соответствующих прямоугольников. Тогда $\Delta\subseteq\left(\bigcup_{i=1}^{n}Y_i\right)\cup\Delta_{n+1}$ и $\left(\bigcup_{i=1}^{n}Y_i\right)\triangle\Delta\subseteq\left(\bigcup_{i=1}^{n}(Y_i\setminus\Delta_i)\right)\cup u_n\subseteq\left(\bigcup_{i=1}^{n}Y_i'\right)\cup u_n$.

Сумма внешних мер всех прямоугольников из последнего объединения сколь угодно мала. Поэтому треугольник Δ измерим.

По теореме 6 а) получаем, что любой многоугольник, как объединение треугольников, является измеримым множеством.

б) Рассмотрим функцию $T_0(\Delta)=\frac{1}{2}\left|\det\begin{pmatrix}1&1&1\\a&b&c\end{pmatrix}\right|$, определенную на множестве U_0 всех треугольников. Значение этой функции T_0 на прямоугольном треугольнике Δ, очевидно, равно $\frac{a_0\cdot b_0}{2}$. Докажем, что и $\mu^*(\Delta)=M(\Delta)=\frac{a_0\cdot b_0}{2}$.

Рассмотрим покрытие данного треугольника Δ прямоугольниками из пункта а) этой теоремы. Тогда

$$\mu^*(\Delta)\leqslant\frac{b_0}{k}\cdot\frac{a_0}{k}+\frac{b_0}{k}\cdot\frac{2a_0}{k}+\ldots+\frac{b_0}{k}\cdot\frac{ka_0}{k}=$$

$$=\frac{a_0\cdot b_0}{k^2}(1+2+\ldots+k)=\frac{a_0\cdot b_0}{k^2}\cdot\frac{k+1}{2}\cdot k\to\frac{a_0b_0}{2}\text{ при }k\to\infty.$$

Рассмотрим объединение Y_k' ровно k прямоугольников, вписанных в Δ (они показаны горизонтальной штриховкой на рис. 80). Тогда

$$\mu^*(Y_k')=\frac{b_0}{k}\cdot\frac{a_0}{k}+\ldots+\frac{b_0\cdot(k-1)\cdot a_0}{k^2}=\frac{a_0\cdot b_0}{k^2}(1+2+\ldots+k-1)=$$

$$=\frac{a_0\cdot b_0}{k^2}\cdot\frac{k-1}{2}\to\frac{a_0\cdot b_0}{2}\text{ при }k\to\infty.$$

Поэтому $\mu^*(\Delta)\geqslant\frac{a_0\cdot b_0}{2}$. Итак, $M(\Delta)=\mu^*(\Delta)=\frac{a_0\cdot b_0}{2}=T_0(\Delta)$.

Функции T_0 и M, очевидно, инвариантны относительно параллельных переносов. Поэтому значения этих функций совпадают для любого прямоугольного треугольника со сторонами, параллельными координатным осям.

Пусть Δ произвольный треугольник. «Разложим» его на прямоугольные треугольники, как показано на рисунке 73б). Точнее, $\Delta=(\Delta_1+\Delta_2+(\Delta_3+\Delta_4))-\Delta_3$. Для каждого из треугольников Δ_1, Δ_2, $\Delta_3+\Delta_4$, Δ_4 значения T_0 и M равны. По теореме 6 б) функция M аддитивна, т. е. $M(\Delta)=M(\Delta_1)+M(\Delta_2)+M(\Delta_3+\Delta_4)-M(\Delta_4)$, а по предложению 4в) для данного конкретного разложения треугольника Δ функция T_0 аддитивна, т. е. $T_0(\Delta_1+\Delta_2)=T_0(\Delta_1)+T_0(\Delta_2)$, $T_0(\Delta_3+\Delta_4)=T_0(\Delta_3)+T_0(\Delta_4)$, $T_0(\Delta)=T_0(\Delta_1+\Delta_2)+T_0(\Delta_3)$.

Итак, $T_0(\Delta)=M(\Delta)$.

в) По предложению 4б) число $T(u)$ не зависит от выбора разложения многоугольника u на треугольники, если существует хотя бы одна конечно-аддитивная функция на множестве U всех многоугольников, которая совпадает с T_0 на множестве U_0 всех треугольников. Но такая функция нами построена! Ею является функция $M\restriction U$.

г) Функции T и M на множестве U_0 всех треугольников совпадают, так как на U_0 совпадают T и T_0. Обе функции T и M аддитивны (однако T конечно-аддитивна, а M счетно-аддитивна). Следовательно, они совпадают на множестве всех многоугольников. ■

Следствие 1. ***Верхнюю меру μ^* можно определить с помощью функции площади T и семейства U всех многоугольников. А именно:***

$$\mu^*(X)=\inf\left\{\sum_{n=1}^{\infty}T(u_n)\;\middle|\;X\subseteq\bigcup_{n=1}^{\infty}u_n,\ u_n\in U\right\}.$$

▷ Обозначим $\mu_1(X)$ правую часть этого равенства. В определении μ^* можно использовать только замкнутые прямоугольники. Поэтому $\mu_1(X)\leqslant\mu^*(X)$. Для любого покрытия $\{u_n\}$ многоугольниками по счетной полуаддитивности верхней меры μ^* получаем

$$\mu^*(X) \leqslant \sum_{n=1}^{\infty} \mu^*(u_n) = \sum_{n=1}^{\infty} T(u_n),$$

т. е. $\mu^*(X) \leqslant \mu_1(X)$. ■

Следствие 2. ***Верхняя мера μ^* является G_δ-инвариантной функцией.***

▷ Пусть $f_{A,\xi}: \boldsymbol{R}^2 \to \boldsymbol{R}^2$ эквиаффинное преобразование, т. е. $f_{A,\xi}(x) = A \cdot x + \xi$, где $A \in L_2$ и $\xi \in \boldsymbol{R}^2$. Если X покрывается семейством многоугольников $\{u_n\}$, то $f(X)$ покрывается семейством многоугольников $\{f(u_n)\}$ и $\sum_{n=1}^{\infty} T(u_n) = \sum_{n=1}^{\infty} T(f(u_n))$.

Отсюда непосредственно вытекает, что $\mu^*(X) = \mu^*(f(X))$. ■

Следствие 3. ***а) Семейство $\mathscr{L}$ измеримых множеств замкнуто относительно эквиаффинных преобразований.***

б) Мера M является G_δ-инвариантной функцией.

▷ а) Если $X \in \mathscr{L}$ и $f: \boldsymbol{R}^2 \xrightarrow{s} \boldsymbol{R}^2$ эквиаффинное преобразование, т. е. $f(x) = A \cdot x + \xi$, где $A \in L_2$ и $\xi \in \boldsymbol{R}^2$, то $f(X) \in \mathscr{L}$, так как соотношение $\mu^*(X \triangle Y) < \varepsilon$ (где Y — элементарная фигура) влечет соотношение $\mu^*(f(X) \triangle f(Y)) < \varepsilon$. Множество $f(Y)$ является многоугольником (но не обязательно элементарным множеством). Поэтому $f(Y) \in \mathscr{L}$, т. е. найдется элементарное множество Y_1, для которого $\mu^*(f(Y) \triangle Y_1) < \varepsilon$. Значит, $\mu^*(f(X) \triangle Y) < \varepsilon$. Здесь полезно разобрать

Упражнение 18. *Докажите, что пара $\langle \mathscr{L}/\sim, \rho \rangle$, где $\rho(X, Y) = \mu^*(X \triangle Y)$, метрическое пространство; $X \sim Y \rightleftarrows \mu^*(X \Delta Y) = 0$.*

б) Это утверждение — частный случай следствия 2. ■

Замечание. Итак, измерение $M: \mathscr{L} \to \boldsymbol{R}_{\geqslant 0} \cup \{\infty\}$, действительно, похоже на измерение $T: U \to \boldsymbol{R}_{>0}$; во-первых, измерение M обладает двумя главными свойствами измерения T: оно аддитивно и G_δ-инвариантно; во-вторых, на многоугольниках значения M совпадают со значениями T. В то же время определяются эти измерения весьма по-разному.

Следствие 4. ***а) Все открытые и замкнутые множества измеримы.***

б) Для любого непустого открытого множества $\mathscr{O}$ выполняется $M(\mathscr{O}) > 0$.

▷ а) Любое открытое в топологии $\boldsymbol{R}^2$ множество является объединением счетного числа прямоугольников со сторонами, параллельными координатным осям. По теореме 6а) получаем проверяемое утверждение.

б) Любое непустое открытое в $\boldsymbol{R}^2$ множество $\mathscr{O}$ содержит невырожденный прямоугольник u со сторонами, параллельными осям. По свойству монотонности меры получаем $M(\mathscr{O}) \geqslant M(u) = T(u) > 0$. ■

Следствие 5. ***Если λ произвольное вещественное число из $\boldsymbol{R}_{>0}$ и X — любое измеримое множество с непустой внутренностью $\mathring{X}$, то существует функция $\mu: \mathscr{L} \to \boldsymbol{R}_{\geqslant 0} \cup \{\infty\}$, обладающая свойствами:***

1) счетной аддитивности,

2) G_δ-инвариантности,

3) $\mu(\boldsymbol{X}) = \lambda$.

▷ Так как $M(X) \geqslant M(\mathring{X}) > 0$, то можно определить функцию

$$\mu(Y) \rightleftarrows \frac{\lambda}{M(X)} \cdot M(Y).$$

Ясно, что $\mu: \mathscr{L} \to \boldsymbol{R}_{\geqslant 0} \cup \{\infty\}$ и μ обладает свойствами 1), 2). Кроме того, $\mu(X) = \frac{\lambda}{M(X)} M(X) = \lambda$. ■

Следствие 6. ***Если μ — счетно-аддитивная функция вида $\mu: \mathscr{L} \to \boldsymbol{R}_{\geqslant 0} \cup \{\infty\}$, которая на множестве U_1 всех прямоугольников равна произведению длин сторон многоугольников, то $\mu = M$.*** (Это своеобразная форма для утверждения о единственности функции M.)

▷ Если $X \in \mathscr{L}$, то существует элементарное множество Y, для которого $\mu^*(X \triangle Y) < \varepsilon$. По лемме 5 выполняется $|M(X) - M(Y)| < \varepsilon$. Повторяя доказательство леммы 5, легко показать, что $|\mu(X) - \mu(Y)| \leqslant \mu(X \triangle Y)$, и по счетной полуаддитивности μ (вытекающей из ее σ-аддитивности), получаем $\mu(X \triangle Y) \leqslant \mu^*(X \triangle Y) < \varepsilon$. По аддитивности μ имеем $\mu(Y) = M(Y)$, т. е. $|M(X) - \mu(X)| \leqslant 2\varepsilon$. ■

Следствие 7. ***Мера M обладает свойством гомотетичности: $\forall \lambda \in \boldsymbol{R}_{>0} \forall X \in \mathscr{L} (M(X) = \lambda^2 \cdot M(H_\lambda(X)))$, где $H_\lambda(x) \rightleftarrows \lambda \cdot x$.***

▷ Доказательство сразу вытекает из гомотетичности площади T:

$$M(H_\lambda(X)) = \inf\left\{\sum_{n=1}^{\infty} T(H_\lambda(u_n))\right\} = \inf\left\{\lambda^2 \cdot \sum_{n=1}^{\infty} T(u_n)\right\} = \lambda^2 \cdot M(X). \quad ■$$

Обозначим $\mathfrak{B}$ наименьшее σ-кольцо, содержащее все открытые в топологии $\boldsymbol{R}^2$ множества. Такое $\mathfrak{B}$ существует и единственно. Его легко образовать обычным в таких случаях способом: пусть $\{R_\alpha\}$ семейство всех σ-колец, содержащих семейство $\mathscr{T}$ всех открытых множеств. Это семейство не пусто, так как содержит, например, $\mathscr{P}(\boldsymbol{R}^2)$. Поэтому $\bigcap_\alpha R_\alpha$ является семейством подмножеств в $\boldsymbol{R}^2$. Пересечение любых σ-колец, содержащих $\mathscr{T}$, является σ-кольцом, содержащим $\mathscr{T}$, которое содержится в любом σ-кольце, содержащем $\mathscr{T}$. Поэтому $\mathfrak{B} = \bigcap_\alpha R_\alpha$; в частности, семейство $\{R_\alpha\}$, содержит $\mathscr{L}$, поэтому $\mathfrak{B} \subseteq \mathscr{L}$.

σ-Кольцо $\mathfrak{B}$ содержит, кроме открытых множеств, все замкнутые множества (как $\boldsymbol{R}^2 \setminus \mathscr{O}$), все одноэлементные множества, все отрезки, все прямые и т. д. В частности, $U \subseteq \mathfrak{B}$.

Обозначим $\mu_{\mathscr{L}}$ сужение функции M на семейство $\mathfrak{B}$, т. е. $\mu_{\mathscr{L}} \rightleftarrows M \restriction \mathfrak{B}$. Измерение (функция) $\mu_{\mathscr{L}}$ называется борелевской мерой в $\boldsymbol{R}^2$.

Множества, принадлежащие σ-кольцу $\mathfrak{B}$, называются борелевскими множествами.

Обозначим $\boldsymbol{R}'_{\geqslant 0} \rightleftarrows \boldsymbol{R}_{\geqslant 0} \cup \{\infty\}$.

Следствие 8. ***Пусть λ произвольное число из $\boldsymbol{R}_{>0}$ и $\mathring{w}$ произвольное борелевское множество с непустой внутренностью $\mathring{w}$.***

Функция $\mu(X) \rightleftarrows \frac{\lambda}{\mu_{\mathscr{L}}(w)} \cdot \mu_{\mathscr{L}}(X) : \mathfrak{B} \to R'_{\geqslant 0}$ ***обладает свойствами:***

1) ***счетной аддитивности,***
2) ***G_s-инвариантности,***
3) $\mu(w) = \lambda$.

▷ Так как $\mu_{\mathscr{L}}(w) > 0$, то определение функции μ корректно.

Свойство 1) непосредственно вытекает из σ-аддитивности функции M, свойство 3) очевидно. Проверим свойство 2). Пусть $f : R^2 \to R^2$ эквиаффинное преобразование и $w \in \mathfrak{B}$. Если $f(w) \in \mathfrak{B}$, то из G_s-инвариантности функции M непосредственно вытекает

$$\mu_{\mathscr{L}}(w) = \mu_{\mathscr{L}}(f(w)).$$

Но почему $f(w) \in \mathfrak{B}$? Если R есть σ-кольцо множеств, содержащее $\mathscr{T}$, то $f(R)$ есть σ-кольцо множеств, содержащее $\mathscr{T}$; и любое σ-кольцо R_1, содержащее $\mathscr{T}$, представимо в виде $f(R)$, для некоторого σ-кольца R, содержащего $\mathscr{T}$ (так как f гомеоморфизм). Множество w принадлежит всем σ-кольцам, содержащим $\mathscr{T}$, поэтому $f(w)$ принадлежит всем σ-кольцам, содержащим $\mathscr{T}$, т. е. $f(w) \in \mathfrak{B}$. ■

Итак,

$$\begin{array}{ccccccc} U & \subset & \mathfrak{B} & \subset & \mathscr{L} & \subset & \mathscr{P}(R^2) \\ T \downarrow & & \mu_{\mathscr{L}} \downarrow & & M \downarrow & & ? \downarrow \\ R_{>0} & & R'_{\geqslant 0} & & R'_{\geqslant 0} & & R'_{\geqslant 0} \end{array}$$

Три построенные нами измерения T, $\mu_{\mathscr{L}}$ и M являются последовательными продолжениями друг друга на все более широкие классы множеств (фигур). Легко видеть, что $U \neq \mathfrak{B}$, и можно доказать, что $\mathfrak{B} \neq \mathscr{L}$ и $\mathscr{L} \neq \mathscr{P}(R^2)$.

Знак «?» указывает на естественно возникающий вопрос: нельзя ли определить такое широкое измерение f, которое пригодно сразу для всех фигур, т. е. $D(f) = \mathscr{P}(R^2)$, и f, обладающее основными свойствами измерений T, $\mu_{\mathscr{L}}$ и M; а именно, свойствами:

1) аддитивности (конечной или счетной),
2) G_s-инвариантности.

Если $f(\cdot) \equiv 0$ или $f(\cdot) \equiv +\infty$, то $D(f) = \mathscr{P}(R^2)$ и эти свойства выполняются. Но такие измерения не содержательны. Решение этого вопроса может зависеть от того, какой вид аддитивности имеется в виду в свойстве 1): конечная аддитивность (характерная для T) или счетная аддитивность (характерная для M и $\mu_{\mathscr{L}}$). Не входя в обсуждение этих двух случаев, условимся свойство 1) понимать как счетную аддитивность, которая отражает счетную аддитивность измерения M, давшего жизнь измерениям T и $\mu_{\mathscr{L}}$. Отрицательный ответ на вопрос о существовании такого широкого измерения содержится в следующем п. 2.

Замечание. 1) Задача, поставленная в начале этого пункта, выполнена: определено измерение μ^* (которое в сущности совпадает с M и $\mu_{\mathscr{L}}$), весьма похожее на измерение T, которое, однако, применимо для всех обычных криволинейных фигур. Возникает вопрос, чему численно равен результат этого измерения для круга, сектора, сегмента и т. п. Оказывается, что численно получаются те обычные ответы, которые хорошо известны из школы. Это будет доказано ниже в пункте 5 этого параграфа; к нему можно перейти прямо сейчас, пропустив пункты 2—4.

2) Определения и доказательства этого пункта и, прежде всего, основополагающее определение 7, используют последовательность квадратов $\square_n$. Убедитесь, что для любой другой последовательности прямоугольников u_n, если $u_n \subseteq u_{n+1}$, $\bigcup\limits_n u_n = R^2$, все результаты этого пункта сохраняются. В частности, получается то же самое семейство измеримых множеств $\mathscr{L}$ и те же функции μ^*, M, $\mu_{\mathscr{L}}$.

2. Неизмеримые множества. Итак, решим вопрос, поставленный в конце предыдущего пункта: раз уж мы хотим продолжить функцию площади T на более широкую область определения (как можно более широкую область определения), то нельзя ли ее продолжить на семейство всех фигур $\mathscr{P}(R^2)$ с сохранением ее основных свойств? По-видимому, был исторический момент, когда это казалось возможным. Потом было обнаружено, что не существует определенной на семействе всех фигур функции f (кроме тождественно равной нулю или $+\infty$), которая счетно-аддитивна и G_s-инвариантна (или хотя бы инвариантна относительно «маленьких» подгрупп группы G_s). Это соответствует интуитивному ощущению, что нельзя предложить содержательный процесс измерения, применимый для совершенно произвольной фигуры. Условия аддитивности и инвариантности измерения f представляются обязательными, так как они, по-видимому, соответствуют экспериментам, связанным с реальными измерениями, и вообще представляются основой всякого процесса измерения.

Утверждение о несуществовании такой функции f вытекает из следующего примера Витали. Обычно его рассматривают в курсе анализа—III для случая множеств, лежащих на прямой R. Поскольку мы занимаемся изучением множеств (фигур) в плоскости R^2, то рассмотрим двумерный вариант этого примера.

Пусть D замкнутый единичный круг в R^2 с центром в точке $\overline{0}$ единичного радиуса, т. е. $D \rightleftarrows \{(x, y) \mid x^2 + y^2 \leqslant 1\}$. Рассмотрим в множестве $D \setminus \{\overline{0}\}$ следующее отношение эквивалентности: точки a и b эквивалентны, если они получаются друг из друга поворотом на угол, главная мера которого (по основанию 2π) рациональна, т. е.

$$a \sim b \rightleftarrows \exists r \in ([0,\ 2\pi[\cap \boldsymbol{Q})(A_r(b) = a),$$

где $A_r = \begin{pmatrix} \cos r & -\sin r \\ \sin r & \cos r \end{pmatrix}$. Легко проверить, что это действительно отношение эквивалентности. Тем самым множество $D \setminus \{\overline{0}\}$ разбивается на непересекающиеся классы эквивалентности. Выберем в каждом из этих классов эквивалентности по одной, какой угодно точке, и множество всех этих и только этих точек обозначим X, $X \subseteq D \setminus \{\overline{0}\}$.

На рисунке 74 основанием «столбика» является точка $[a]$ из фактор-множества $D/\sim$, «столбик» над этой точкой совпадает с

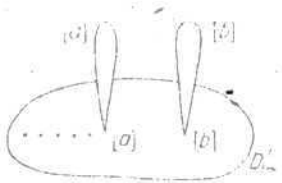


Рис. 74

множеством $[a]$. Если $\varphi: D \to D/\sim$ каноническое отображение, то столбик над $[a]$ совпадает с множеством $\varphi^{-1}([a])$. Столбики попарно не пересекаются. В каждом столбике мы выбрали по одной точке, и эти точки образуют множество X. Докажем, что

$$\bigcup_{r \in [0,\ 2\pi[\cap \boldsymbol{Q}} A_r(X) = D \setminus \{\bar{0}\}.$$

Действительно, любая точка a входит в какой-то класс эквивалентности. По определению множества X в этом классе эквивалентности найдется какой-то элемент $a' \in X$. Поэтому $a \sim a'$ и $a = A_r \cdot a'$ при некотором $r \in [0,\ 2\pi[\cap \boldsymbol{Q}$, т. е. $a \in A_r(X)$.

Докажем, что

$$A_{r_1}(X) \cap A_{r_2}(X) = \varnothing$$

при $r_1 \neq r_2$. Действительно, предположим, что $a \in (A_{r_1}(X)) \cap (A_{r_2}(X))$. Получаем $a = A_{r_1} \cdot a' = A_{r_2} \cdot a''$, где a' и $a'' \in X$. Пусть $r_1 \geqslant r_2$. Тогда $a'' = A_{r_2}^{-1} \circ A_{r_1} a' = A_{r_1 - r_2} a'$, т. е. $a'' \sim a'$, что противоречит определению множества X.

Воспользуемся инвариантностью функции f относительно поворотов, т. е. частным случаем G_s-инвариантности. Тогда значения функции f на всех множествах вида $A_r(X)$ одинаковы и равны $f(X)$. Поэтому проколотый круг $D \setminus \{\bar{0}\}$ представляется в виде объединения счетного числа (так как $\boldsymbol{Q}$ счетно) попарно непересекающихся множеств $A_r(X)$, для которых значения функции f равны между собой и равны $f(X)$. По свойству счетной аддитивности получаем:

$$f(D \setminus \{\bar{0}\}) = \sum_{r \in \boldsymbol{Q}} f(A_r(X)).$$

Следовательно, $f(D \setminus \{\bar{0}\})$ равно $+\infty$, если $f(X) > 0$, и равно 0, если $f(X) = 0$.

Кажется в равной мере странным, чтобы результат измерения «проколотого круга» $D \setminus \{0\}$ равнялся нулю или бесконечности. Допустим $f(D \setminus \{0\}) = 0$. Тогда на любом прямоугольнике, содержащемся в $D \setminus \{0\}$, значение функции f равно нулю. По инвариантности f относительно группы параллельных переносов $(\boldsymbol{R}^2)^+$, т. е. также по весьма частному случаю G_s-инвариантности, и по счетной аддитивности значения функции f на любом прямоугольнике равны 0. Отсюда значение функции f на любом открытом множестве и на любом многоугольнике равно нулю. (По свойству внешней регулярности, которое мы докажем в теореме 9, отсюда вытекает, что f тождественно равна 0.) Допустим, $f(D \setminus \{0\}) = \infty$. Тогда на любом прямоугольнике, содержащем D, значение функции f равно ∞. По инвариантности функции f относительно группы $(R^2)^+$ и свойству аддитивности значения функции f на всех прямоугольниках и, следовательно, на всех открытых множествах и всех многоугольниках равны ∞. (По тому же свойству внешней регулярности

отсюда вытекает, что f тождественно равна ∞). Итак, хотя мы строго не доказали, что $f \equiv 0$ или $f \equiv \infty$ (это сразу следует из теоремы 9а), получена следующая более слабая

Теорема 8. ***Не существует функции $f: \mathscr{P}(\boldsymbol{R}^2) \to R'_{\geqslant 0}$, обладающей свойствами счетной аддитивности и G_κ-инвариантности, которая хотя бы на одном многоугольнике (или на одном открытом множестве) имела значение, не равное 0 и $+\infty$.***

Эту теорему также можно доказать, используя вместо инвариантности функции f относительно поворотов, ее инвариантность относительно «параллельных переносов». Точнее, вместо $D \setminus \{\bar{0}\}$ рассмотрим квадрат $[0,\ 1[\times [0,\ 1]$ и вместо поворотов A_r рассмотрим по существу параллельные переносы:

$$B_r: \binom{x}{y} \mapsto \begin{cases} \binom{x+r}{y}, & \text{если } x + r < 1, \\ \binom{x+r-1}{y}, & \text{если } x + r \geqslant 1, \end{cases}$$

где $r \in [0,\ 1[\cap \boldsymbol{Q}$.

Если положить $(x \oplus y = z) \rightleftarrows (e^{2\pi i x} \cdot e^{2\pi i y} = e^{2\pi i z})$, где $x, y, z \in [0,\ 1[$, то легко проверить, что множество $[0,\ 1[$ является группой относительно бинарной операции $\oplus$ и $B_r\left(\binom{x}{y}\right) = \binom{x \oplus r}{y}$.

Упражнение 19. *Допустим, существует функция f, удовлетворяющая условиям теоремы 8. Докажите, что $f([0,\ 1[\times [0,\ 1[)$ равно 0 или $+\infty$. Отсюда получите еще одно доказательство теоремы 8.*

3. Аксиоматическое определение меры. Теперь мы хотим перейти от конструктивного определения измерения M (аналогичного конструктивному определению измерения T) к аксиоматическому определению измерения M (аналогичного аксиоматическому определению в § 3 измерения T как функции площади). Главная трудность состоит в том, что определение M существенно использует прямоугольники, определение которых связано с $\boldsymbol{R}^2$, точнее, с декартовой структурой в $\boldsymbol{R}^2$. Если рассмотреть произвольную плоскость, например векторную, то до тех пор, пока в ней не фиксирован базис, нельзя говорить о прямоугольниках и их площади. В пункте 1 мы выделили основные свойства функции M: это счетная аддитивность и G_s-инвариантность. Правда, мы отмечали, что многие свойства функции M вытекают из ее инвариантности относительно маленьких подгрупп группы G_s. «Самой маленькой» из таких подгрупп является группа параллельных переносов $(\boldsymbol{R}^2)^+$. Оказывается, что из свойств счетной аддитивности и $(\boldsymbol{R}^2)^+$-инвариантности уже вытекает свойство G_s-инвариантности! Поэтому в аксиоматическом определении функции M естественно сформулировать аксиомы в их более слабой форме, как поступают, например, при аксиоматизации элементарной геометрии.

Напомним, что семейство всех борелевских множеств в плоскости $\boldsymbol{R}^2$ обозначается $\mathfrak{B}$ — оно определяется с использованием только топологии в $\boldsymbol{R}^2$. Мы приходим к следующему определению.

Определение 9. **Борелевской мерой (короче — мерой) называется любая функция вида $\mu: \mathfrak{B} \to R'_{\geqslant 0}$ со свойствами:**

1) счетной аддитивности: если $\{w_n\} \subseteq \mathfrak{B}$, то

$$\mu\left(\bigcup_{n=1}^{\infty} w_n\right) = \sum_{n=1}^{\infty} \mu(w_n).$$

2) $(R^2)^+$-инвариантности: для любых $w \in \mathfrak{B}$ и $\xi \in R^2$ выполняется $\mu(w) = \mu(w+\xi)$.

Замечание. 1) Согласно определения 9 мерами являются функции, тождественно равные нулю или бесконечности. С точки зрения реальных измерений они малоинтересны. Условимся в дальнейшем называть борелевской мерой (мерой) только функции, удовлетворяющие определению 9 и отличные от двух упомянутых функций. В этом смысле меры также существуют: любая функция вида $c \cdot \mu_{\mathscr{I}}(\cdot)$, где c постоянная из $R_{>0}$, является мерой.

2) Почему в определении 9 в качестве области определения $D(\mu)$ выбирается σ-кольцо, содержащее топологию $\mathscr{T}$ в R^2? В случае измерения M было доказано, что $D(M)$ есть σ-кольцо, содержащее $\mathscr{T}$. Второй аргумент: так как конечные объединения и конечные пересечения многоугольников являются снова многоугольниками, то для получения даже простейших «криволинейных» фигур из «прямолинейных» фигур приходится рассматривать операции счетного объединения и счетного пересечения. Например, замкнутый круг D получается как пересечение счетного числа n-угольников, описанных около D при n, стремящемся к бесконечности, а открытый круг $\mathring{D}$ получается как объединение счетного числа n-угольников, вписанных в него. Допуская операции счетного объединения и счетного пересечения, естественно потребовать замкнутость $D(\mu)$ относительно них, т. е. потребовать, чтобы $D(\mu)$ было σ-кольцом. С другой стороны, все элементарные фигуры (после вычитания их границы, которая не влияет на измерение) являются открытыми множествами, а любое открытое множество является счетным дизъюнктным объединением прямоугольников; поэтому $\mathscr{T}$ включается в $D(\mu)$. Мы приходим к тому, что $D(\mu)$ есть σ-кольцо, содержащее $\mathscr{T}$, и нет причин считать, что $D(\mu)$ шире наименьшего σ-кольца, содержащего $\mathscr{T}$. Итак, $D(\mu) = \mathfrak{B}$.

Почему так выбрана область значений $R(\mu)$? Так как в $D(\mu)$ входят пустое множество, все прямоугольники и вся плоскость R^2 и мы ожидаем, что результаты их измерения будут соответственно равны 0, любому строго положительному числу и $+\infty$, то все элементы из $R'_{\geqslant 0}$ будут результатами некоторых измерений.

3) Определение 9 без всяких изменений переносится на любую топологическую группу вместо топологической группы $(R^2)^+$.

Цель оставшейся части этого пункта — вывести основные свойства произвольной борелевской меры из ее аксиоматического определения 9.

Упражнение 20. *Докажите, что: а) Всякое открытое в R^2 множество представимо в виде счетного объединения возрастающей последовательности компактных множеств.*

б) Всякое замкнутое в R^2 множество представимо в виде счетного пересечения убывающей последовательности открытых множеств. Если F компакт, то эту последовательность можно начать с ограниченного множества. Напомним, что возрастающая последовательность множеств определяется как такая, что $X_n \subseteq X_{n+1}$, $\forall n \in N$, а убывающая — как такая, что $X_n \supseteq X_{n+1}$, $\forall n \in N$.

в) Семейство борелевских множеств можно определить как наименьшее σ-кольцо множеств, содержащее семейство C всех компактов.

Предложение 10. ***Семейство борелевских множеств совпадает с наименьшим семейством множеств, содержащим все компакты и замкнутым относительно двух операций: счетного объединения и счетного пересечения.***

▷ Второе из этих семейств временно обозначим B'. Оно содержит все открытые множества. Так как $\mathfrak{B}$ содержит семейство C всех компактов и замкнуто относительно этих операций, то $B' \subseteq \mathfrak{B}$. Проверим, что B' замкнуто относительно операции дополнения $\urcorner$. После этого получим, что семейство B' замкнуто относительно операции разности $\setminus$ так, как $X \setminus Y = X \cap \urcorner Y$, и, значит, B' есть σ-кольцо, содержащее $\mathscr{T}$, т. е. $\mathfrak{B} \subseteq B'$, и, следовательно, $\mathfrak{B} = B'$. Итак, нужно только проверить, что семейство B' замкнуто относительно операции дополнения $\urcorner$.

Рассмотрим семейство $B'' \rightleftarrows \{w \in B' \mid \urcorner w \in B'\}$. Ясно, что $B'' \subseteq B'$. Любой компакт c содержится в B'', так как $\urcorner c$ — открытое множество.

Если $\{w_n\} \subseteq B''$, то множества $\bigcup_{n=1}^{\infty} w_n$ и $\bigcap_{n=1}^{\infty} w_n$ принадлежат B'', так как $\urcorner w_n \in B'$; например, в первом случае $\urcorner\left(\bigcup_{n=1}^{\infty} w_n\right) = \bigcap_{n=1}^{\infty}(\urcorner w_n) \in B'$.

Так как B' наименьшее семейство с данными свойствами, то $B' \subseteq B''$, что требовалось доказать. ■

Упражнение 21. *Проверьте, что следующие множества борелевские:* а) *множества точек плоскости, у которых соответственно обе координаты рациональны, обе иррациональны, одна (первая) — рациональна, а другая (вторая) — иррациональна;*

б) окружность, любая дуга окружности, эллипс, синусоида и их дуги и другие обычные кривые из курса анализа.

Нумерация свойств в следующем предложении продолжает их нумерацию из определения 9.

Предложение 11. ***Пусть w_n, w — произвольные борелевские множества. Любая мера $\mu: \mathfrak{B} \to R'_{\geqslant 0}$ обладает следующими свойствами:***

3) $\mu(\varnothing) = 0$,

если $w_1 \subseteq w$, то $\mu(w) = \mu(w_1) + \mu(w \setminus w_1)$; кроме того, выполняется $\mu(w_1 \cup w_2) + \mu(w_1 \cap w_2) = \mu(w_1) + \mu(w_2)$;

4) «монотонность»: если $w_1 \subseteq w_2$, то $\mu(w_1) \leqslant \mu(w_2)$;

5) «непрерывность снизу»: если $w_1 \subseteq w_2 \subseteq \ldots$ — возрастающая последовательность множеств, то

$$\mu\Big(\bigcup_{n=1}^{\infty} w_n\Big) = \lim_{n\to\infty} \mu(w_n);$$

6) «непрерывность сверху»: если $w_1 \supseteq w_2 \supseteq \ldots$ — убывающая последовательность множеств и $\mu(w_1) < +\infty$, то

$$\mu\Big(\bigcap_{n=1}^{\infty} w_n\Big) = \lim_{n\to\infty} \mu(w_n);$$

7) «счетная полуаддитивность»: если $w \subseteq \bigcup_{n=1}^{\infty} w_n$, то

$$\mu(w) \leqslant \sum_{n=1}^{\infty} \mu(w_n);$$

8) «внутренняя регулярность для открытых множеств»: если $\mathcal{O}$ — открытое множество, то

$$\mu(\mathcal{O}) = \sup\{\mu(c) \mid c \subseteq \mathcal{O},\ c \text{ — компакт}\};$$

9) $\mu(\mathcal{O}) > 0$ для любого непустого открытого множества $\mathcal{O}$;

10) если значение меры конечно хотя бы для какого-то одного непустого открытого множества, то ее значение конечно для любого ограниченного борелевского множества;

11) если $\mathcal{O}$ открытое множество, то

$$\mu(\mathcal{O}) = \inf\Big\{\sum_{n=1}^{\infty} \mu(u_n) \mid \mathcal{O} \subseteq \bigcup_{n=1}^{\infty} u_n,\ \{u_n\} \subseteq U_1\Big\},$$

где U_1 — семейство всех прямоугольников в R^2.

▷ 3) Выполняется $\mu(w) = \mu(w \sqcup \varnothing) = \mu(w) + \mu(\varnothing)$. Если $\mu(\varnothing) = +\infty$, то $\mu \equiv +\infty$, а такую функцию мы условились не называть мерой (см. замечание 2 к определению 9). Поэтому $0 \leqslant \mu(\varnothing) < +\infty$ и $\mu(\varnothing) = \mu(\varnothing \sqcup \varnothing) = \mu(\varnothing) + \mu(\varnothing)$ и $\mu(\varnothing) = 0$. Далее, $\mu(w) = \mu(w_1 \sqcup (w \setminus w_1)) = \mu(w_1) + \mu(w \setminus w_1)$ и $\mu(w_1 \cup w_2) = \mu(w_1 \sqcup (w_2 \setminus w_1)) = \mu(w_1) + \mu(w_2 \setminus w_1) = \mu(w_1) + \mu(w_2 \setminus (w_1 \cap w_2)) = \mu(w_1) + \mu(w_2) - \mu(w_1 \cap w_2)$. Отметим, что последнее равенство использует предположение о конечности $\mu(w_2)$, так как иначе $\mu(w_2 \setminus (w_1 \cap w_2)) = \infty - \infty = ?$ Однако проверяемое равенство остается верным и в том случае, если $\mu(w_1 \cap w_2) = \infty$, так как тогда его левая и правая части равны ∞. Последнее проверяется непосредственно по свойству монотонности.

4) Так как $\mu(\cdot) \geqslant 0$, то $\mu(w) = \mu(w_1) + \mu(w \setminus w_1) \geqslant \mu(w_1)$.

5) Если все значения меры $\mu(w_n)$ конечны, то $\mu\Big(\bigcup_{n=1}^{\infty} w_n\Big) = \mu(w_1 \sqcup (w_2 \setminus w_1) \sqcup (w_3 \setminus w_2) \sqcup \ldots) = \mu(w_1) + \sum_{k=1}^{\infty} \mu(w_{k+1} \setminus w_k) = \lim_{n\to\infty}\Big[\mu(w_1) + \sum_{k=1}^{n-1} \mu(w_{k+1} \setminus w_k)\Big] = \lim_{n\to\infty} [\mu(w_1) + \mu(w_2) - \mu(w_1) + \ldots + \mu(w_n) - \mu(w_{n-1})] = \lim_{n\to\infty} \mu(w_n)$. Если хотя бы одно значение $\mu(w_n)$ бесконечно, то проверяемое равенство непосредственно очевидно по свойству монотонности.

6) По свойству 4) имеем: $\forall n(\mu(w_n) < +\infty)$ и $\mu\Big(\bigcap_{n=1}^{\infty} w_n\Big) < +\infty$. Последовательность множеств $\{w_1 \setminus w_n\}$ возрастающая и $\bigcap_{n=1}^{\infty} (w_1 \setminus w_n) = w_1 \setminus \bigcup_{n=1}^{\infty} w_n$. По свойствам 3) и 5) получаем $\mu\Big(\bigcap_{n=1}^{\infty} w_n\Big) = \mu(w_1) - \mu\Big(\bigcup_{n=1}^{\infty} (w_1 \setminus w_n)\Big) = \lim_{n\to\infty} [\mu(w_1) - \mu(w_1 \setminus w_n)] = \lim_{n\to\infty} \mu(w_n)$.

(Заметим, что условие $\mu(w_1) < \infty$ существенно: если w_n — полуплоскость справа от прямой $x = n$, то $\bigcap_{n=1}^{\infty} w_n = \varnothing$ и левая часть равна 0, а правая равна ∞. Это условие доставит нам много хлопот в дальнейшем.)

7) Докажем сначала, что $\mu\Big(\bigcup_{n=1}^{m} w_n\Big) \leqslant \sum_{n=1}^{m} \mu(w_n)$. При $m = 2$ это неравенство проверялось при доказательстве свойства 3. Допустим, такое неравенство выполняется для какого-то числа m. Тогда

$$\mu\Big(\bigcup_{n=1}^{m+1} w_n\Big) = \mu\Big(\Big(\bigcup_{n=1}^{m} w_n\Big) \cup w_{m+1}\Big) \leqslant \mu\Big(\bigcup_{n=1}^{m} w_n\Big) + \mu(w_{m+1}) \leqslant \sum_{n=1}^{m+1} \mu(w_n).$$

По индукции замечаем, что такое неравенство выполняется для всех m. По свойству 5) получаем:

$$\mu\Big(\bigcup_{n=1}^{\infty} w_n\Big) = \lim_{m\to\infty} \mu\Big(\bigcup_{n=1}^{\infty} w_n\Big) \leqslant \lim_{m\to\infty} \sum_{n=1}^{m} \mu(w_n) = \sum_{n=1}^{\infty} \mu(w_n).$$

Теперь из свойства монотонности следует, что

$$\mu(w) \leqslant \mu\Big(\bigcup_{n=1}^{\infty} w_n\Big) \leqslant \sum_{n=1}^{\infty} \mu(w_n).$$

8) По свойству монотонности число $\mu(\mathcal{O})$ — верхняя грань множества чисел $\{\mu(c) \mid c \subseteq \mathcal{O},\ c$ — компакт$\}$. С другой стороны, по упражнению 20 а) $\mathcal{O} = \bigcup_{n=1}^{\infty} c_n$, где $c_n \subseteq c_{n+1}$ и c_n — компакты. По свойству 5) $\mu(\mathcal{O}) = \lim_{n\to\infty} \mu(c_n)$. Следовательно, $\mu(\mathcal{O})$ точная верхняя грань множества чисел $\{\mu(c) \mid c \subseteq \mathcal{O},\ c$ — компакт$\}$.

9) Предположим, что для некоторого непустого открытого множества $\mathcal{O}$ выполняется $\mu(\mathcal{O}) = 0$. По свойству монотонности $\mu(\mathring{D}) = 0$, где $\mathring{D}$ — некоторый открытый круг, содержащийся в $\mathcal{O}$. Параллельно перенося $\mathring{D}$, получим круги с центрами во всех точках с рациональными координатами, которые по свойству $(R^2)^+$-инвариантности имеют одну и ту же меру, равную мере $\mu(\mathring{D})$. Получим покрытие всей плоскости R^2 счетным числом множеств, каждое из которых имеет меру нуль. По свойствам полуаддитивности и мо-

нотонности получаем, что функция μ тождественно равна нулю. Однако мы условились не считать такую функцию мерой.

10) Если для некоторого непустого открытого множества $\mathscr{O}$ выполняется $\mu(\mathscr{O})<+\infty$, то $\mu(\mathring{D})<+\infty$, где $\mathring{D}$ — открытый круг, содержащийся в $\mathscr{O}$. Любое ограниченное множество w можно покрыть конечным числом открытых кругов, являющихся параллельными переносами круга $\mathring{D}$. Действительно, в силу ограниченности, w содержится в каком-то замкнутом круге D_1. Такой круг является компактом. Покроем D_1 параллельными переносами открытого круга $\mathring{D}$. Из этого покрытия выберем конечное подпокрытие. В силу полуаддитивности получим, что $\mu(w)<+\infty$.

11) Правую часть проверяемого равенства обозначим λ. По свойству полуаддитивности $\mu(\mathscr{O})\leqslant\lambda$. Как легко проверить, любое открытое множество $\mathscr{O}$ является дизъюнктным объединением счетного числа прямоугольников, т. е. $\mathscr{O}=\bigsqcup_{k=1}^{\infty} w_k$, и по свойству σ-аддитивности

$$\mu(\mathscr{O})=\sum_{k=1}^{\infty}\mu(u_k). \quad \blacksquare$$

Замечание. Условимся в дальнейшем называть борелевской мерой (мерой) только те меры в смысле определения 9 и замечания после него, у которых значение хотя бы на одном открытом множестве конечно. Иными словами, мы будем в дальнейшем рассматривать только меры, значения которых конечны на любых ограниченных борелевских множествах (см. предложение 11.10)). Отсюда сразу вытекает, что меры любого одноточечного множества, любого отрезка и любой прямой равны нулю (проверьте). Это означает, в частности, что в предложении 11.11) можно считать все прямоугольники u_k, покрывающие множество $\mathscr{O}$, замкнутыми прямоугольниками.

Теорема 9. ***Для всякого борелевского множества $w\in\mathfrak{B}$ выполняются свойства:***

а) $\mu(w)=\inf\{\mu(\mathscr{O}) \mid w\subseteq\mathscr{O},\ \mathscr{O}$ — открыто$\}$;

б) $\mu(w)=\sup\{\mu(c) \mid c\subseteq w,\ c$ — компактно$\}$;

в) $\mu(w)=\inf\left\{\sum_{k=1}^{\infty}\mu(u_k) \mid w\subseteq\bigcup_{k=1}^{\infty}u_k,\ u_k\in U_1\right\}$.

Борелевское множество w называется внешне регулярным, если для него выполняется свойство из пункта а). Борелевское множество w называется внутренне регулярным, если для него выполняется свойство из пункта б). Таким образом, пункты а) и б) утверждают, что любое борелевское множество внешне и внутренне регулярно. Мы докажем только внешнюю регулярность произвольного борелевского множества, оставляя утверждение о его внутренней регулярности в качестве упражнения 23 г). Для доказательства теоремы 9 нам потребуются следующие леммы.

Лемма 7. ***Если каждое множество w_n внешне регулярно, то и множество $\bigcup_{n=1}^{\infty} w_n$ внешне регулярно.***

▷ Для любого $\varepsilon>0$ выберем такие открытые множества $\mathscr{O}_n$, что $w_n\subseteq\mathscr{O}_n$ и $\mu(\mathscr{O}_n)\leqslant\mu(w_n)+\frac{\varepsilon}{2^n}$. Используя включение $\left(\left(\bigcup_{n=1}^{\infty}\mathscr{O}_n\right)\setminus\left(\bigcup_{n=1}^{\infty}w_n\right)\right)\subseteq\bigcup_{n=1}^{\infty}(\mathscr{O}_n\setminus w_n)$, получим $\mu\left(\bigcup_{n=1}^{\infty}\mathscr{O}_n\right)-\mu\left(\bigcup_{n=1}^{\infty}w_n\right)=\mu\left(\left(\bigcup_{n=1}^{\infty}\mathscr{O}_n\right)\setminus\left(\bigcup_{n=1}^{\infty}w_n\right)\right)\leqslant\mu\left(\bigcup_{n=1}^{\infty}\left(\mathscr{O}_n\setminus w_n\right)\right)\leqslant\sum_{n=1}^{\infty}\mu(\mathscr{O}_n\setminus w_n)<\varepsilon\sum_{n=1}^{\infty}\frac{1}{2^n}=\varepsilon$.

Следовательно, множество $\bigcup_{n=1}^{\infty} w_n$ внешне регулярно. $\blacksquare$

Лемма 8. ***Если $\{w_n\}$ — убывающая последовательность внешне регулярных борелевских множеств и значение меры $\mu(w_1)$ конечно, то множество $w=\bigcap_{n=1}^{\infty} w_n$ внешне регулярно.***

▷ Согласно предложению 11.6 выполняется $\mu(w)=\lim_{n\to\infty}\mu(w_n)$. По любому числу $\lambda>\mu(w)$ найдется число n, для которого $\lambda>\mu(w_n)$. В силу внешней регулярности множества w_n найдется открытое множество $\mathscr{O}$, для которого $w_n\subset\mathscr{O}$ и $\lambda>\mu(\mathscr{O})$. Поэтому $w\subset\mathscr{O}$ и $\mu(\mathscr{O})-\mu(w)\leqslant\lambda-\mu(w)$. Последнее означает внешнюю регулярность множества w. (Условие $\mu(w_1)<+\infty$ существенно, так как оно используется в предложении 11.) $\blacksquare$

Лемма 9. ***Всякое открытое множество и всякий компакт внешне регулярны.***

▷ Первое утверждение очевидно. Всякий компакт c по упражнению и замечанию после предложения 11 представим в виде $c=\bigcup_{n=1}^{\infty}\mathscr{O}_n$, где $\mathscr{O}_n$ — открытые множества и $\mathscr{O}_{n+1}\subseteq\mathscr{O}_n$ и значение меры $\mu(\mathscr{O}_1)$ конечно. По лемме 8 компакт c внешне регулярен. $\blacksquare$

Обозначим $\mathfrak{B}_0$ семейство всех множеств, представимых в виде конечного дизъюнктного объединения собственных разностей компактов. Напомним, что разность множеств $X\setminus Y$ называется собственной, если $Y\subseteq X$. Конечно, $C\subset\mathfrak{B}_0\subset\mathfrak{B}$.

Упражнение 22. *Проверьте, что $\mathfrak{B}_0$ — кольцо множеств.*

Лемма 10. ***Всякое множество из $\mathfrak{B}_0$ внешне регулярно.***

▷ По определению запись $w\in\mathfrak{B}_0$ означает, что $w=w_1\sqcup\ldots\sqcup w_n$, где каждое w_i вида $w_i=c_i\setminus c_i'$, $c_i'\subset c_i$ и c_i, c_i' — компакты. Если все множества w_i внешне регулярны, то по лемме 7 и множество w внешне регулярно. Поэтому остается проверить только внешнюю регулярность множеств w_i. Например, пусть $i=1$. По лемме 9 компакты c_1 и c_1' внешне регулярны. Поэтому для любого $\varepsilon>0$ найдется открытое множество $\mathscr{O}$, для которого $c_1\subset\mathscr{O}$ и $\mu(\mathscr{O})-\mu(c_1)\leqslant\varepsilon$. Так как $c_1\setminus c_1'\subset\mathscr{O}\setminus c_1'$, то $\mu(\mathscr{O}\setminus c_1')-\mu(c_1\setminus c_1')=\mu((\mathscr{O}\setminus c_1')\setminus(c_1\setminus c_1'))=\mu(\mathscr{O}\setminus c_1)=\mu(\mathscr{O})-\mu(c_1)\leqslant\varepsilon$. Так как множество $\mathscr{O}\setminus c_1'$ открыто, то полученное неравенство означает, что множество $c_1\setminus c_1'$ внешне регулярно. $\blacksquare$

Условимся обозначать $\searrow_{n=1}^{\infty} w_n$ пересечение убывающей последовательности множеств. Иными словами, $\left(\searrow_{n=1}^{\infty} w_n\right) \rightleftarrows \bigcap_{n=1}^{\infty} w_n$, когда $\forall n(w_{n+1} \subseteq w_n)$. Аналогично условимся обозначать $\left(\nearrow_{n+1}^{\infty} w_n\right) \rightleftarrows \bigcup_{n=1}^{\infty} w_n$, когда $\forall n(w_n \subseteq w_{n+1})$. Конечно, операции $\nearrow_{n=1}^{\infty}$ и $\searrow_{n=1}^{\infty}$ являются частными случаями операций $\bigcup_{n=1}^{\infty}$ и $\bigcap_{n=1}^{\infty}$.

Следующая лемма по-новому характеризует борелевские множества и поэтому представляет самостоятельный интерес.

Лемма 11. ***Семейство $\mathfrak{B}$ борелевских множеств совпадает с наименьшим семейством фигур, содержащим семейство $\mathfrak{B}_0$ и замкнутым относительно операций счетного объединения возрастающей последовательности множеств и счетного пересечения убывающей последовательности множеств.***

Смысл такой характеристики семейства борелевских множеств в том, что за счет некоторого усложнения исходного семейства множеств — вместо C берется $\mathfrak{B}_0$ — можно упростить операции, порождающие $\mathfrak{B}$, — вместо $\bigcup_{n=1}^{\infty}$ и $\bigcap_{n=1}^{\infty}$ берутся $\nearrow_{n=1}^{\infty}$ и $\searrow_{n=1}^{\infty}$.

▷ Это второе семейство множеств обозначим A. Как всегда, легко проверяется, что оно существует. Ясно, что $A \subseteq \mathfrak{B}$, так как $\mathfrak{B}_0 \subset \mathfrak{B}$ и $\mathfrak{B}$ замкнуто относительно более общих операций $\bigcup_{n=1}^{\infty}$ и $\bigcap_{n=1}^{\infty}$, а A — наименьшее семейство. Докажем, что A является σ-кольцом. Тогда, так как $\mathfrak{B}$ наименьшее σ-кольцо, содержащее C (см. упражнение 20в)), и A σ-кольцо, содержащее C, получим, что $\mathfrak{B} \subseteq A$ и, следовательно, $A = \mathfrak{B}$.

Сначала докажем, что A — кольцо. Обозначим

$$K(X) = \{Y \mid Y \setminus X,\ X \setminus Y,\ X \cup Y \in A\},$$

где X какое-то фиксированное множество. Ясно, что $(Y \in K(X)) \Leftrightarrow (X \in K(Y))$. Если $\{Y_n\}$ — какая-то возрастающая или убывающая последовательность множеств из $K(X)$, то (для определенности рассмотрим второй случай)

$$(\searrow_{n=1}^{\infty} Y_n) \setminus X = \searrow_{n=1}^{\infty} (Y_n \setminus X) \in A, \quad X \setminus (\searrow_{n=1}^{\infty} Y_n) = \searrow_{n=1}^{\infty} (X \setminus Y_n) \in A,$$

$$X \cup (\searrow_{n=1}^{\infty} Y_n) = \searrow_{n=1}^{\infty} (X \cup Y_n) \in A,$$

т. е. множество $(\searrow_{n=1}^{\infty} Y_n)$ принадлежит семейству множеств $K(X)$. Следовательно, семейство множеств $K(X)$ замкнуто относительно операций $\nearrow_{n=1}^{\infty}$ и $\searrow_{n=1}^{\infty}$. По упражнению 22 семейство $\mathfrak{B}_0$ является кольцом. Поэтому, если w_1, $w_2 \in \mathfrak{B}_0$, то $w_1 \setminus w_2$, $w_2 \setminus w_1$, $w_1 \cup w_2 \in \mathfrak{B}_0$. Следовательно, $(w_1,\ w_2 \in \mathfrak{B}_0) \Rightarrow w_1 \in K(w_2)$ и $\mathfrak{B}_0 \subseteq K(w_2)$ для любого фиксированного множества w_2 из $\mathfrak{B}_0$. Так как A — наименьшее семейство множеств, то $A \subseteq K(w_2)$. Поэтому $\forall w_1 \in A\,(w_1 \in K(w_2),\ w_2 \in K(w_1))$. Так как w_2 — любое множество

из $\mathfrak{B}_0$, то $\forall w_1 \in A(\mathfrak{B}_0 \subseteq K(w_1),\ A \subseteq K(w_1))$, т. е. $\forall w_1,\ w_2 \in A(w_2 \in K(w_1))$. Последнее означает, что семейство множеств A — кольцо.

Теперь проверим, что A является σ-кольцом. Пусть $\{X_n\}$ — произвольная последовательность множеств из A. Множества $X_n' \rightleftarrows X_1 \cup \ldots \cup X_n$ принадлежат A, так как оно является кольцом. Тогда $\bigcup_{n=1}^{\infty} X_n = (\nearrow_{n=1}^{\infty} X_n') \in A$ по определению семейства A. ■

Лемма 12. ***Пусть c — произвольный компакт в R^2. Обозначим $\mathfrak{B}(c)$ наименьшее семейство подмножеств множества c, содержащее все компакты c_0, $c_0 \subseteq c$ и замкнутое относительно операций счетного объединения и разности. Обозначим $\mathfrak{B}_0(c)$ семейство всех множеств вида $(c_1 \setminus c_1') \sqcup \ldots \sqcup (c_n \setminus c_n')$, где $c_i' \subset c_i$ и все c_i, c_i' — компакты, содержащиеся в c. Обозначим $A(c)$ наименьшее семейство множеств, содержащее $\mathfrak{B}_0(c)$ и замкнутое относительно операций $\nearrow_{n=1}^{\infty}$ и $\searrow_{n=1}^{\infty}$. Тогда***

$$\mathfrak{B}(c) = A(c) \text{ и } \mathfrak{B}(c) = \{w \cap c \mid w \in \mathfrak{B}\}.$$

Отметим, что компакты в R^2, содержащиеся в c, можно характеризовать как компакты в топологическом пространстве c с индуцированной топологией. Эта лемма (кроме последнего равенства) говорит то же самое, что лемма 11, с той разницей, что в ней R^2 заменено на c. Доказательство сохраняется без всяких изменений (проверьте!).

▷ Доказательство теоремы 9. а) Произвольное борелевское множество w представимо в виде $w = \bigcup_{n=1}^{\infty} (w \cap c_n)$, где c_n — возрастающая последовательность компактов, например концентрических кругов, для которых $\bigcup_{n=1}^{\infty} c_n = R^2$. Если любое ограниченное борелевское множество внешне регулярно, то по лемме 7 произвольное борелевское множество внешне регулярно. Таким образом, доказательство теоремы 9а) свелось к доказательству ее частного случая, когда рассматривается ограниченное борелевское множество. Ограниченное борелевское множество w удобно тем, что для него значение меры $\mu(w)$ конечно (по замечанию после предложения 11), и потому условие конечности значения меры в лемме 8 обеспечено. Итак, идея доказательства теоремы 9а) в следующем: все борелевские множества и, в частности, все ограниченные борелевские множества по лемме 11 получаются из множеств, входящих в $\mathfrak{B}_0$, операциями $\bigcup_{n=1}^{\infty}$ и $\searrow_{n=1}^{\infty}$; все множества, входящие в $\mathfrak{B}_0$, внешне регулярны и операции $\bigcup_{n=1}^{\infty}$ и $\searrow_{n=1}^{\infty}$ почти сохраняют внешнюю регулярность. Поэтому все ограниченные борелевские множества внешне регулярны. Повторим это рассуждение строго.

Пусть w произвольное ограниченное борелевское множество. Какой-то замкнутый круг c содержит w в своей внутренности,

т. е. $w \subseteq \mathring{c}$. Поскольку $\mu(c) < \infty$, то любое борелевское множество, содержащееся в c, имеет конечную меру. В частности, любое борелевское множество вида $w_1 \cap c$ имеет конечную меру, т. е. все значения меры μ на семействе $\mathfrak{B}(c)$ конечны. Рассмотрим семейство всех внешне регулярных подмножеств круга c. Оно содержит семейство $\mathfrak{B}_0(c)$ и по леммам 7 и 8 замкнуто относительно операций $\bigcup_{n=1}^{\infty}$ и $\bigcap_{n=1}^{\infty}$. Поэтому оно содержит семейство $A(c)$ и по лемме 12 содержит $\mathfrak{B}(c)$. С другой стороны, оно содержит w. Итак, w внешне регулярно.

Пункт б) теоремы 9 мы не доказываем.

в) Эта принципиально важная формула непосредственно вытекает из пункта а) теоремы 9 и внешней регулярности меры на открытых множествах. ■

Упражнение 23. *Докажите, что*

а) *пересечение любой последовательности внутренне регулярных борелевских множеств конечной меры внутренне регулярно;*

б) *объединение возрастающей последовательности внутренне регулярных борелевских множеств внутренне регулярно;*

в) *объединение двух не пересекающихся внутренне регулярных множеств внутренне регулярно;*

г) *любое борелевское множество внутренне регулярно.*

Из этого упражнения вытекает доказательство теоремы 9б).

Теорема 9 имеет принципиальное значение для всей концепции меры. В ее пункте а) устанавливается тесная связь между топологией в $\boldsymbol{R}^2$ и мерой в $\boldsymbol{R}^2$. А именно, любое значение меры однозначно определяется по ее значениям на открытых множествах — маленькой части семейства всех борелевских множеств. Подчеркнем, что понятия меры и топологии, судя по их определениям, не связаны между собой, не предполагают друг друга и наличие тесной связи между ними является неожиданностью.

В пункте в) теоремы 9 устанавливается тесная связь между декартовым строением множества $\boldsymbol{R}^2$ и мерой в $\boldsymbol{R}^2$. А именно, любое значение меры однозначно определяется по ее значениям на простейших фигурах — замкнутых прямоугольниках, которые определяются за счет того, что $\boldsymbol{R}^2$ — декартово произведение $\boldsymbol{R} \times \boldsymbol{R}$. Связь между декартовым строением множества и мерой в нем опять-таки является неожиданностью.

Итак, если указана инструкция для измерения всего лишь замкнутых прямоугольников, то любое значение любой меры однозначно определяется по этой инструкции. В частности, однозначно определяется мера круга, сектора, сегмента и т. д.

4. Сравнение конструктивного и аксиоматического определений меры. Пусть μ произвольная аксиоматически определенная мера. По теореме 9в), любое ее значение $\mu(w)$ однозначно вычисляется по значениям вида $\mu(u)$, где u произвольный замкнутый прямоугольник. Если смотреть на меру μ, как на измерение, то ее значения на замкнутых прямоугольниках — простейших многоугольниках — естественно считать равными значениям какой-то функции площади s. Функция площади s аксиоматически определялась в § 4, п. 1 и там же доказывалось, что она имеет вид $s(\cdot) = c \cdot T(\cdot)$, $\forall u \in U$. При этом U_1 — небольшая часть U. Итак, если $\forall u \in U_1 (\mu(u) = s(u))$, то $\mu(w) = \inf \{\sum_n c \cdot T(u_n)\} = c \cdot M(w)$, т. е. μ по существу совпадает с M. Это мотивирует следующую теорему.

Теорема 10. *а) Фиксируем любое борелевское множество w_0 конечной меры с непустой внутренностью и любое положительное число λ_0. Существует мера $\mu: \mathfrak{B} \to R'_{\geqslant 0}$, для которой $\mu(w_0) = \lambda_0$. А именно, $\mu \rightleftarrows c \cdot \mu_{\mathscr{L}}$, где $c \rightleftarrows \frac{\lambda_0}{\mu(w_0)}$.*

б) Для любых двух мер μ_1 и μ_2 выполняется: если их значения совпадают на каком-то борелевском множестве конечной меры с непустой внутренностью, то они всюду совпадают, т. е. $\mu_1 = \mu_2$.

▷ а) По определению меры $\mu_{\mathscr{L}}$, приведенному в конце пункта 1, утверждение этого пункта очевидно.

б) Начнем со следующей леммы.

Лемма 13. ***Пусть μ — любая мера и $\mu([0,1] \times [0,1]) \rightleftarrows c$. Тогда для любого прямоугольника $u_{x,y}$, длины сторон которого равны числам x и y, выполняется $\mu(u_{x,y}) = c \cdot x \cdot y$.***

▷ Разобьем квадрат $[0,1] \times [0,1]$ на n^2 маленьких квадратов со сторонами $\frac{1}{n}$, не имеющих общих внутренних точек. По $(\boldsymbol{R}^2)^+$ инвариантности меры получаем, что значения меры на этих квадратиках равны между собой. Тогда по конечной аддитивности меры получаем:

$$\mu(u_{\frac{1}{n},\frac{1}{n}}) = c \cdot \frac{1}{n^2}.$$

Если x и y — рациональные числа вида $x = \frac{k}{n}$ и $y = \frac{s}{m}$, то аналогичным образом прямоугольник $u_{x,y}$ разбиваем на $(km) \cdot (sn)$ квадратиков со стороной $\frac{1}{mn}$, не имеющих общих внутренних точек. Тогда

$$\mu(u_{x,y}) = (km)(sn) \cdot \mu(u_{\frac{1}{mn},\frac{1}{mn}}) = c \cdot (km) \cdot (sn) \cdot \left(\frac{1}{mn}\right)^2 = c \cdot \frac{kmsn}{m^2 n^2} = c \cdot x \cdot y.$$

Если x и y — произвольные вещественные числа, то $u_{x,y} = \bigcup_{n=1}^{\infty} u_{x_n,y_n}$, где $u_{x_1,y_1} \subseteq u_{x_2,y_2} \subseteq \ldots$ возрастающая последовательность и x_n, y_n — рациональные числа такие, что $\lim\limits_{n\to\infty} x_n = x$ и $\lim\limits_{n\to\infty} y_n = y$.

Тогда по предложению 11. 5) получаем

$$\mu(u_{x,y}) = \lim \mu(u_{x_n,y_n}) = c \cdot \lim x_n \cdot y_n = c \cdot x \cdot y. \quad ■$$

Перейдем к доказательству пункта б) теоремы. Рассмотрим две произвольные меры: μ_1 и μ_2. Положим

$c_1 \rightleftarrows \mu_1([0,1]\times[0,1])$, $c_2 \rightleftarrows \mu_2([0,1]\times[0,1])$ и $c \rightleftarrows \frac{c_1}{c_2}$.

Тогда для любого прямоугольника $u_{x,y}$ выполняется $\mu_1(u_{x,y}) = c_1 \times x \cdot y = c \cdot c_2 \cdot x \cdot y = c \cdot \mu_2(u_{x,y})$, т. е. меры μ_1 и μ_2 почти совпадают на всех прямоугольниках. Если борелевское множество w представимо в виде $\bigcup_{k=1}^{\infty} u_k$, где $\{u_k\} \subseteq U_1$, то по счетной аддитивности мер выполняется $\mu_1(w) = c \cdot \mu(w)$. Так как в этом виде представляется любое открытое множество в $\boldsymbol{R}^2$, то меры μ_1 и μ_2 почти совпадают на всех открытых множествах.

По внешней регулярности мер получаем, что для любого борелевского множества w выполняется $\mu_1(w) = \inf\{\mu_1(\mathcal{O})/w \subseteq \mathcal{O}, \mathcal{O} \in \mathcal{T}_{R^2}\} = c \cdot \inf\{\mu_2(\mathcal{O}) | w \in \mathcal{O}, \mathcal{O} \in \mathcal{T}_{R^2}\} = c \cdot \mu_2(w)$, т. е. μ_1 и μ_2 почти совпадают всюду.

По условию для некоторого w_0 из $\mathfrak{B}$ с непустой внутренностью выполняется $\mu_1(w_0) = \mu_2(w_0)$. По предложению 11 $\mu_i(w_0) > 0$ и получаем, что $c = 1$. Следовательно, меры μ_1 и μ_2 в точности совпадают. ■

Следствие 1. ***а) Любая аксиоматически определенная мера μ равна $c \cdot \mu_{\mathcal{L}}$.***

б) Любые две аксиоматически определенные меры μ_1 и μ_2 отличаются самое большее на постоянную, т. е. $\mu_1 \equiv c \cdot \mu_2$, где $c > 0$.

▷ а) Выберем любое ограниченное борелевское множество w_0 с непустой внутренностью, например квадрат или круг. Пусть $\mu(w_0) = \lambda_0$. По теореме 10а) существует мера вида $c \cdot \mu_{\mathcal{L}}$, имеющая одинаковое с мерой μ значение на w_0. По теореме 10б) $\mu = c \cdot \mu_{\mathcal{L}}$.

б) Это утверждение непосредственно вытекает из пункта а). ■

Следствие 2. ***а) Любая мера G_s-инвариантна, т. е. свойства счетной аддитивности и R^2-инвариантности функции вида $f : \mathcal{L} \to R'_{\geqslant 0}$ эквивалентны свойствам счетной аддитивности и G_s-инвариантности этой функции.***

б) Сужение любой аксиоматически определенной меры на множество всех многоугольников является функцией площади.

Итак, установлена связь аксиоматических определений функции площади и меры.

▷ а) По следствию 1а) любая мера μ имеет вид $c \cdot \mu_{\mathcal{L}}$. По следствию 7 к теореме 7 функция $\mu_{\mathcal{L}}$ является G_s-инвариантной функцией. Следовательно, такова же функция μ.

б) Это утверждение непосредственно вытекает из пункта а). ■

Упражнение 24. *Если f — гомеоморфизм вида $f : \boldsymbol{R}^2 \leftrightarrow \boldsymbol{R}^2$, то для любого борелевского множества w множество $f(w)$ борелевское.*

Следствие 3. ***Любая мера μ обладает свойством гомотетичности: если H_k — отображение гомотетии с коэффициентом k из $\boldsymbol{R}_{>0}$, то $\forall w \in \mathfrak{B}(\mu(H_k(w))) = k^2 \cdot \mu(w)$.***

▷ Доказательство вытекает из следствия 1а) и следствия 6 к теореме 7. ■

5. Вычисление меры простейших криволинейных фигур.

Предложение 12. ***Мера образа любой гладкой кривой равна нулю.***

▷ Пусть $\alpha : R \to R^2$ гладкая кривая, т. е. $\alpha(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$, где функции $x(\cdot)$ и $y(\cdot)$ имеют непрерывные производные при всех значениях аргумента t. Сначала разберем случай, когда t меняется на отрезке $[0, 1]$. Поскольку функции x' и y' непрерывны на отрезке $[0, 1]$, то они ограничены на нем. Обозначим

$$M \rightleftarrows \max\{(x'(t))^2 + (y'(t))^2 | t \in [0, 1]\} + 1.$$

Пусть $t_1, t_2 \in [0, 1]$ и $t_1 < t_2$. По теореме Лагранжа, применяемой к отрезку $[t_1, t_2]$ и функциям $x(t)$ и $y(t)$, получаем:

$$\|(x(t_2), y(t_2)) - (x(t_1), y(t_1))\|^2 = [x(t_2) - x(t_1)]^2 + [y(t_2) - y(t_1)]^2 = [(t_2 - t_1) \cdot x'(t_3)]^2 + [(t_2 - t_1) y'(t_4)]^2 < M(t_2 - t_1)^2,$$

где t_3 и t_4 — некоторые точки отрезка $[t_1, t_2]$. Разобьем отрезок $[0, 1]$ на n равных частей. По доказанному неравенству образ отрезка $\left[\frac{k}{n}, \frac{k+1}{n}\right]$ целиком содержится в открытом круге с центром в точке $\left(x\left(\frac{k}{n}\right), y\left(\frac{k}{n}\right)\right)$ и радиусом $\frac{\sqrt{M}}{n}$ (рис. 75). Тем более этот образ содержится внутри квадрата со стороной $\frac{2 \cdot \sqrt{M}}{n}$. Так как значение меры μ на этом квадрате равно значению функции $c \cdot T$, где c — некоторая постоянная, то образ всей кривой покрыт n квадратами, мера каждого из которых равна $c \cdot \frac{4M}{n^2}$. Поэтому $\mu(R(\alpha)) \leqslant \frac{4c \cdot M}{n}$. Устремляя n к бесконечности, получаем, что $\mu(R(\alpha)) = 0$.

В общем случае кривая $\alpha(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ определена на всей прямой $\boldsymbol{R}$, интервале или полуинтервале. Рассмотрим, например, первый из этих случаев. Разобьем R на отрезки ..., $[-1, 0]$, $[0, 1]$, ... и затем воспользуемся счетной полуаддитивностью любой меры. ■

Следствие. ***Мера любой окружности равна нулю.***

Отметим, что существует непрерывное, но не гладкое отобра-

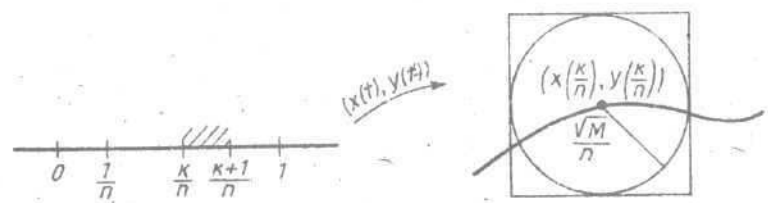


Рис. 75

жение вида $[0, 1]\to\boldsymbol{R}^2$, образ которого совпадает со всем квадратом $[0, 1]\times[0, 1]$. Оно называется кривой Пеано. Поэтому в предложении 12 условие гладкости кривой существенно.

До конца этого параграфа будем обозначать μ меру, значение которой на квадрате $[0, 1]\times[0, 1]$ равно числу 1. Ясно, что эта мера совпадает с $\mu_{\mathcal{L}}$, а на всех многоугольниках совпадает с функцией площади T.

Предложение 13. ***а) Мера единичного круга D равна*** π.
б) Мера сектора в круге D, опирающегося на центральный угол, главная мера которого по какому-то основанию $a>0$ ***равна*** x, ***равна*** $\frac{x}{a}\pi$.

▷ а) Так как мера обладает свойством $(\boldsymbol{R}^2)^+$-инвариантности, то утверждение этого пункта не зависит от того, в какой точке находится центр круга.

Экспонента e_a с основанием $a>0$ является непрерывным гомоморфизмом вида $e_a:\boldsymbol{R}^+\to S$, который биективно отображает полуинтервал $[0, a[$ на единичную окружность S, являющуюся границей D. Пусть

$$z\rightleftarrows e_a\left(\frac{a}{n}\right)=e^{\frac{2\pi}{a}i\left(\frac{a}{n}\right)}=e^{\frac{2\pi i}{n}}=\left\langle\cos\frac{2\pi}{n},\ \sin\frac{2\pi}{n}\right\rangle,$$

где $n\geqslant 3$. Если Δ_n — треугольник с вершинами в точках $\{0, 1, z\}$, то

$$T(\Delta_n)=\frac{1}{2}\left|\det\begin{pmatrix}\cos\frac{2\pi}{n} & 1\\ \sin\frac{2\pi}{n} & 0\end{pmatrix}\right|=\frac{1}{2}\sin\frac{2\pi}{n}=\mu(\Delta_n).$$

Пусть

$$W\rightleftarrows\begin{pmatrix}\cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n}\\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n}\end{pmatrix}\in SO_2.$$

Тогда $W(\Delta_n)$ — треугольник с вершинами в точках $\{0, z, z^2\}$ и в силу G_s-инвариантности меры $\mu(W(\Delta_n))=\mu(\Delta_n)=\frac{1}{2}\sin\frac{2\pi}{n}$.

Положим $u_n\rightleftarrows\Delta_n+W(\Delta_n)+W^2(\Delta_n)+\ldots+W^{n-1}(\Delta_n)$ и получим правильный n-угольник u_n, вписанный в D. При этом $\mu(u_n)=n\cdot\frac{1}{2}\sin\frac{2\pi}{n}=\pi\cdot\frac{\sin\frac{2\pi}{n}}{\frac{2\pi}{n}}\to\pi$ при $n\to\infty$ (отсюда $\forall n(\mu(u_n)\leqslant\leqslant\mu(D))$ и $\pi\leqslant\mu(D)$).

Если $n=2^i$, где $i=2, 3, 4, \ldots$, то многоугольники u_n образуют возрастающую последовательность. Очевидно, $\mathring{D}=\bigcup_n\mathring{u}_n$ и по непрерывности снизу $\mu(\mathring{D})=\lim\mu(u_n)=\pi$. Мера окружности равна нулю. Поэтому $\mu(D)=\pi$.

Упражнение 25. *Докажите, что для любой точки* $z\in\mathring{D}$ *существует номер* n, *для которого* $z\in\mathring{u}_n$.

б) Сначала разберем случай, когда главная мера по основанию $a>0$ угла, на который опирается сектор, равна числу $\frac{a}{n}$. Другими словами, найдем меру сектора w с вершинами $\{0, 1, z\}$, где $z=e_a\left(\frac{a}{n}\right)=e^{\frac{2\pi}{a}i\left(\frac{a}{n}\right)}=\left\langle\cos\frac{2\pi}{n},\ \cos\frac{2\pi}{n}\right\rangle$ (см. доказательство, пункт а)).

Тогда $D=w\cup W(w)\cup\ldots\cup W^{n-1}(w)$, где $W\rightleftarrows\begin{pmatrix}\cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n}\\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n}\end{pmatrix}\in SO_2$.

Пусть $\tilde{D}$ — круг без отрезков $[0, 1], [0, z], [0, z^2], \ldots, [0, z^{n-1}]$ и $\tilde{w}$ — сектор без отрезков $[0, 1]$ и $[0, z]$. В силу G_s-инвариантности меры получаем $\mu(\tilde{w})=\mu(W(w))=\ldots=\mu(W^{n-1}(\tilde{w}))$. Так как значение меры на любом отрезке равно нулю, то по равенству

$$\tilde{D}=\tilde{w}\cup W(\tilde{w})\cup\ldots\cup W^{n-1}(\tilde{w})$$

получаем, что $\pi=\mu(D)=\mu(\tilde{D})=n\cdot\mu(\tilde{w})$.

Следовательно, $\mu(w)=\mu(\tilde{w})=\frac{1}{n}\pi=\frac{x}{a}\pi$, где $x=\frac{a}{n}$.

Если главная мера по основанию $a>0$ угла, на который опирается сектор w', равна числу $\frac{m}{n}a$, то $m<n$, так как $\frac{m}{n}a\in[0, a[$. Для такого сектора w' выполняется $w'=w\cup W(w)\cup\ldots\cup W^m(w)$, где w и W описаны выше. При этом слагаемые пересекаются только по отрезкам $[0, z], \ldots, [0, z^{m-1}]$. Тогда исходный сектор w' (без конечного числа отрезков) представим в виде $\tilde{w}\cup W(\tilde{w})\cup\ldots\cup W^{m-1}(\tilde{w})$. Следовательно, $\mu(w')=m\cdot\mu(\tilde{w})=m\cdot\mu(w)=\frac{m}{n}\pi=\frac{x}{a}\pi$, где $x=\frac{m}{n}\cdot a$.

Наконец, если главная мера по основанию $a>0$ угла, образующего сектор w, равна иррациональному числу x, то сектор w без одного отрезка можно представить в виде объединения возрастающей последовательности секторов w_n, опирающихся на углы, главные меры которых по основанию a являются рациональными числами r_n, и числа r_n сколь угодно приближаются x по недостатку. Тогда

$$\mu(w)=\mu(w\setminus[0, z])=\mu(\bigcup_n w_n)=\lim_{n\to\infty}\mu(w_n)=\lim_{n\to\infty}\frac{r_n}{a}\pi=\frac{x}{a}\pi.\ \blacksquare$$

Замечания: 1) Значения меры на круге произвольного радиуса легко получаются по свойству гомотетичности меры.

2) Значения меры на кольце, образованном концентрическими окружностями радиусов R и r, легко получить, используя свойство аддитивности меры: $\mu(w)=\pi(R^2-r^2)$.

3) Меру сектора w чаще всего находят при измерении углов по основанию $a=2\pi$ или $a=360°$. Если $a=360°$, то $\mu(w)=\frac{\pi}{360}x°$.

4) Мера сегмента находится как разность мер сектора и треугольника:

$$\mu(w)=\frac{1}{2}x\,(\text{рад})-\frac{1}{2}|\det(ab)|.$$

6. Сравнение борелевской меры с мерами Жордана и Лебега. Обозначим G семейство всех множеств, которые одновременно являются борелевскими множествами и измеримы по Жордану (см. пункт 3г в § 4). По предложению 9 и теореме 7а выполняется $U\subseteq G\subseteq\mathscr{B}$. В определении 6 функция T_4 зависела от числа s_0 и базиса $\{\bar{e}_1, \bar{e}_2\}$. В плоскости $\boldsymbol{R}^2$ выберем канонический базис $\bar{e}_1=\begin{pmatrix}1\\0\end{pmatrix}$, $\bar{e}_2=\begin{pmatrix}0\\1\end{pmatrix}$ и число $s_0=1$. Иными словами, выберем ту единственную функцию T_4, для которой $T_4([0,\,1]^2)=1$. В примере а) после определения 6 подсчитано, что $T_4([a,\,b]\times[c,\,d])=(c-a)\times(d-b)$. Сужение этой функции на G обозначим μ_G и назовем мерой Жордана.

Предложение 14. ***Мера $\mu_{\mathscr{L}}: \mathscr{B}\to\boldsymbol{R}'_{\geqslant 0}$ является продолжением меры Жордана $\mu_G: G\to\boldsymbol{R}_{\geqslant 0}$.***

▷ Пусть $w\in G$. По определению 6 $\mu_G(w)$ является пределом числовой последовательности $\frac{N_i}{100^i}$. При любом i объединение всех ячеек i-го ранга, пересекающихся с w, покрывает w; поэтому

$$\mu_{\mathscr{L}}(w)\leqslant\frac{N_i}{100^i},\ \mu_{\mathscr{L}}(w)\leqslant s^*(w).$$

По тому же определению 6 $\mu_G(w)$ является пределом числовой последовательности $\frac{n_i}{100^i}$. При любом i объединение всех ячеек i-го ранга, содержащихся в w, содержится в w, поэтому $\mu_{\mathscr{L}}(w)\geqslant\frac{n_i}{100^i}$, $\mu_{\mathscr{L}}(w)\geqslant s_*(w)$. Поскольку $\mu_G(w)=s^*(w)=s_*(w)$, то $\mu_G(w)=\mu_{\mathscr{L}}(w)$, $\forall w\in G$. ■

Следствие 1. ***Функции T_4 и T тождественно совпадают.***

▷ Обе эти функции определены на множестве U. И обе они совпадают с $\mu_{\mathscr{L}}\restriction U$. ■

Следствие 2. ***Функция T_4, рассматриваемая на множестве G, G_s-инвариантна.***

▷ Если $w\in\mathscr{B}$ и $A\in G_s$, то $A(w)\in\mathscr{B}$, так как A гомеоморфизм $\boldsymbol{R}^2$ на себя. Если w измеримо по Жордану, то по предложению 9б) для любых $\varepsilon>0$ найдутся такие многоугольники u_1 и u_2, что $u_1\subseteq w\subseteq u_2$ и $T_4(u_2)-T_4(u_1)<\varepsilon$. Тогда $A(u_1)\leqslant A(w)\leqslant A(u_2)$. Так как $A(u_1)$ и $A(u_2)$ — многоугольники и для многоугольников функция T_4 совпадает с функцией T, которая G_s-инвариантна, то $T_4(A(u_2))-T_4(A(u_1))<\varepsilon$. Следовательно, $A(w)$ измеримо по Жордану. В силу G_s-инвариантности меры $\mu_{\mathscr{L}}$ (следствие 3б) теоремы 7) получаем:

$$T_4(A(w))=\mu_G(A(w))=\mu_{\mathscr{L}}(A(w))=\mu_{\mathscr{L}}(w)=\mu_G(w)=T_4(w).$$ ■

Мера Жордана μ_G, как и мера $\mu_{\mathscr{L}}$, является G_s-инвариантной функцией, совпадающей на семействе всех многоугольников U с функцией T. Однако — и это ее принципиальное отличие от меры $\mu_{\mathscr{L}}$ — мера Жордана лишь конечно-аддитивна, в то время как $\mu_{\mathscr{L}}$ σ-аддитивна.

Множество всех точек квадрата $[0,\,1]\times[0,\,1]$, у которых обе координаты рациональны и строго положительны, счетно. Поэтому оно борелевское и его мера $\mu_{\mathscr{L}}$ равна нулю. В § 3 показано, что это множество неизмеримо по Жордану. Следовательно, G строго содержится в $\mathscr{B}$. С другой стороны, G строго содержит U, так как, например, круг — борелевское множество, измеримое по Жордану, которое не является многоугольником.

Соотношение между площадью, мерой Жордана, мерой $\mu_{\mathscr{L}}$ и мерой M представлено на следующей коммутативной диаграмме:

$$\begin{array}{lccccccc}
 & U & \subset & G & \subset & B & \subset & \mathscr{L}\\
\text{Функция площади } T & \downarrow & & \downarrow & & \downarrow & & \downarrow\\
 & \text{мера Жордана } \mu_G & & & & \text{мера } \mu_{\mathscr{L}} & & \text{мера } M\\
 & \boldsymbol{R}_{>0} & \subset & \boldsymbol{R}'_{\geqslant 0} & = & \boldsymbol{R}'_{\geqslant 0} & = & \boldsymbol{R}'_{\geqslant 0}.
\end{array}$$

При этом мера Жордана является продолжением функции площади T с U на G, а мера $\mu_{\mathscr{L}}$ — продолжение меры Жордана μ_G с G на $\mathscr{B}$, а мера Лебега M — продолжение меры $\mu_{\mathscr{L}}$ с $\mathscr{B}$ на $\mathscr{L}$.

Глава IV

АЛГЕБРАИЧЕСКИЕ УРАВНЕНИЯ СТЕПЕНЕЙ, МЕНЬШИХ ИЛИ РАВНЫХ 5, И ГЕОМЕТРИЧЕСКИЕ ПОСТРОЕНИЯ

§ 1. СВЯЗЬ МЕЖДУ РАЗРЕШИМОСТЬЮ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ В РАДИКАЛАХ И ВЫПОЛНИМОСТЬЮ ТРАДИЦИОННЫХ ГЕОМЕТРИЧЕСКИХ ПОСТРОЕНИЙ

1. Кубические уравнения и квадратичные расширения. Начнем этот параграф с рассмотрения алгебраической стороны вопроса о возможности того или иного традиционного геометрического построения. Под традиционными геометрическими построениями понимается построение, выполняемое последовательным применением конечное число раз циркуля и линейки. Что это точно значит, мы обсудим ниже.

Все построения в этом параграфе производятся в арифметической плоскости $\boldsymbol{R}^2$.

Числовым полем называется любое подполе поля $\boldsymbol{C}$. Например, числовым полем является само поле $\boldsymbol{C}$, поле $\boldsymbol{R}$, поле $\boldsymbol{Q}$, поля $\boldsymbol{Q}(\sqrt{2})$, $\boldsymbol{Q}(\sqrt{3})$, $\boldsymbol{Q}(i)$ и т. п.

Пусть P — числовое поле и z_0 — фиксированное комплексное число. Напомним, что расширением поля P, порожденным числом z_0, называется такое наименьшее (относительно отношения порядка $\subseteq$) числовое поле P_1, которое содержит поле P и число z_0, т. е. $P\subseteq P_1$ и $z_0\in P$. Легко доказать, что такое поле P_1 существует, единственно, так как оно совпадает с пересечением всех подполей P' поля $\boldsymbol{C}$, для которых $P\subseteq P'$ и $z_0\in P'$. Оно обозначается $P(z_0)$. Это обозначение было использовано выше.

Пусть P — числовое поле и $z_0, z_1, \ldots, z_n$ — любые фиксированные комплексные числа. Расширением поля P, порожденным числами $z_0, z_1, \ldots, z_n$, называется наименьшее числовое поле P', содержащее P (т. е. $P\subseteq P'$) и $z_0, \ldots, z_n$ — элементы P'. Такое расширение обозначается $P(z_0, \ldots, z_n)$. Ясно, что оно совпадает с пересечением всех числовых полей, содержащих поле P и числа $z_0, \ldots, z_n$. Следовательно, поле $P(z_0, \ldots, z_n)$ существует и единственно.

Упражнение 1. *Докажите, что:*

а) пересечение любого числа подполей поля C является подполем поля C;

б) найдется хотя бы одно подполе P' поля C, для которого $P\subseteq P'$ и $z_0\in P$, где P — фиксированное числовое поле и z_0 — фиксированное комплексное число;

в) расширение $P(z_0)$ поля P, порожденное числом z_0, содержит все числа вида $k_1+k_2\cdot z_0$, где $k_1, k_2\in P$;

г) если $z_0^2\in P$, то $P(z_0)=\{k_1+k_2\cdot z_0 \mid k_1, k_2\in P\}$;

д) выполняются соотношения $P(z_0, z_1)=(P(z_0))(z_1)$, $P(z_0, z_1, z_2)=(P(z_0, z_1))(z_2)$ и т. д.

Все поля, которые рассматриваются в этой главе, являются числовыми полями. Поэтому числовое поле будем называть просто полем. Собственным называется любое расширение P_1 поля P, для которого $P\neq P_1$.

Напомним два определения из курса алгебры:

а) Простым квадратичным расширением поля P называется любое поле вида $P(z_0)$, где $z_0^2\in P$.

б) Поле P' называется квадратичным расширением поля P, если найдется такая цепочка полей $P\neq P_0\subset P_1\subset\ldots\subset P_s\neq P'$, вложенных друг в друга, что любое поле P_i является простым квадратичным расширением поля P_{i-1}, $i=1, 2, \ldots, s$.

Предложение 1. ***Если уравнение $z^3+az^2+bz+c=0$ с коэффициентами из поля P не имеет решений в поле P, то это уравнение не имеет решений и в любом его квадратичном расширении P'.***

▷ Рассмотрим сначала случай, когда поле P' является собственным простым квадратичным расширением вида $P(z_0)$ поля P. Допустим, что уравнение $z^3+az^2+bz+c=0$ $(*)$ имеет корень z_1 в поле $P'\neq P(z_0)$, где $z_0\notin P$, $z_0^2\in P$. Согласно упражнению 1,г существуют числа k_1, k_2 из P, для которых

$$z_1\neq k_1+k_2\cdot z_0.$$

Проверим, что z_1 является корнем квадратного уравнения

$$z^2+Az+B=0, \qquad (**)$$

где $A, B\in P$.

Действительно, $(z-(k_1+k_2z_0))\cdot(z-(k_1-k_2z_0))=z^2-2k_1z+(k_1^2-k_2^2z_0^2)=z^2+Az+B$.

Здесь $A\neq -2k_1\in P$ и $B\neq k_1^2-k_2^2z_0^2\in P$, так как $z_0^2\in P$.

Если многочлен z^3+az^2+bz+c разделить на многочлен z^2+Az+B, то получим $z^3+az^2+bz+c=(z^2+Az+B)(z+a-A)+Cz+D$, где $C, D,\in P$.

Так как число z_1 — корень одновременно уравнений $(*)$ и $(**)$, то z_1 — корень уравнения $Cz+D=0$. При $C\neq 0$ получаем равенство $z_1=-\frac{D}{C}\in P$. Но тогда и $z_0=\frac{z_1-k_1}{k_2}\in P$, что противоречит условию $z_0\notin P$.

Если допустить, что $C=0$, то из равенства $Cz_1+D=0$ следует $D=0$. Тогда из равенства

$$z^3+az^2+bz+c=(z^2+Az+B)(z+a-A)$$

вытекает, что уравнение $(*)$ имеет корень $z=A-a$, принадлежащий P, что также противоречит условиям теоремы.

Пусть теперь поле P' — квадратичное расширение поля P и

$P \neq P_0 \subset P_1 \subset \ldots \subset P_s \neq P'$ — цепочка таких вложенных друг в друга полей, что для всех $i=1, 2, \ldots, s$ поле P_i является простым квадратичным расширением поля P_{i-1}.

В соответствии с доказанным выше уравнение $z^3+az+bz+c=0$ не имеет корней в поле P_1. Его коэффициенты a, b, c содержатся в P и, следовательно, содержатся в P_1. Поэтому уравнение $(*)$ не имеет корней и в поле P_2, и т. д. рассуждаем для всех чисел $3, \ldots, s$. ■

Следствие. а) Пусть a и b — целые числа, сумма которых отлична от -2 и разность которых отлична от нуля. Тогда уравнение $z^3+az^2+bz+1=0$ не имеет решений ни в каком квадратичном расширении поля $\boldsymbol{Q}$.

б) Пусть n — целое число, не являющееся кубом никакого другого целого числа. Тогда уравнение $z^3-n=0$ не имеет решений ни в каком квадратичном расширении поля $\boldsymbol{Q}$.

▷ а) По предложению 1 достаточно показать, что уравнение $z^3+az^2+bz+1=0$ не имеет рациональных корней. Но всякий рациональный корень многочлена с целыми коэффициентами и старшим коэффициентом, равным единице, является делителем свободного члена. Поэтому рациональными корнями уравнения $z^3+az^2+bz+1=0$ могут быть только числа 1 и -1. При $z=1$ получаем $z^3+az^2+bz+1=a+b+2=0$ и при $z=-1$ получаем $z^3+az^2+bz+1=a-b=0$, что противоречит условию.

б) Так же как и в случае а, достаточно доказать отсутствие рациональных корней у уравнения $z^3-n=0$. Если z_0 — рациональный корень этого уравнения, то z_0 — делитель числа n, т. е. $n=z_0 \cdot k$, где $k \in \boldsymbol{Z}$. Тогда $z_0^2-k=0$. Аналогичным образом $k=z_0 \times m$, где $m \in \boldsymbol{Z}$. Тогда $z_0=m$ и $n=m^3$, что противоречит условию следствия. ■

2. Построение циркулем и линейкой. Рассмотрим задачу трисекции угла, т. е. задачу о построении с помощью циркуля и линейки такого угла β, для которого выполняется равенство $3 \cdot \beta \neq \beta+\beta+\beta=\alpha$, где α — некоторый наперед заданный угол. Конечно, выражение «построить с помощью» заранее неясно и требует уточнения и определения. Необходимость такого уточнения особенно видна из способа трисекции произвольного угла, который предложил еще Архимед. В его построении используются только циркуль и линейка. Но линейка используется особым образом. Напомним этот способ Архимеда.

Пусть фиксирован угол $0<\alpha<180°$. Можно считать, что он задается парой лучей, один из которых совпадает с положительным направлением оси Ox (см. рис. 76). Продолжим с помощью линейки горизонтальную сторону угла влево и проведем с помощью циркуля полуокружность произвольного радиуса R с центром в вершине угла O.

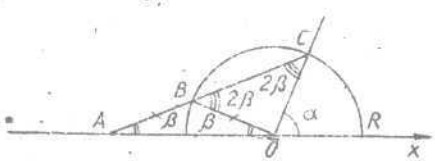


Рис. 76

На линейке отметим такие две точки A и B, что расстояние $|A, B|$ равняется R. Затем подберем такое положение линейки, при котором точка A окажется на продолженной стороне угла и точка B окажется на полуокружности и в то же время линейка пройдет через точку пересечения полуокружности со второй стороной угла. Таким образом задается угол β (см. рис. 76). Угол β как раз равен третьей части угла α, так как $\alpha=180° -(180° -4\beta)-\beta=3\beta$.

Можно ли считать, что таким образом задача о трисекции угла α с помощью циркуля и линейки решена? Ответ зависит от того, что же назвать «построением с помощью циркуля и линейки».

Определим это важное понятие. Для этого сначала определим три отображения φ_1, φ_2, φ_3, которые будем называть простыми CL-построениями. Здесь буква C от английского слова *circle* — окружность, а буква l от *line* — прямая.

Пусть M_1, M_2, M_3, M_4 — четыре произвольные точки в плоскости $\boldsymbol{R}^2$. Тогда:

1) значение $\varphi_1(M_1, M_2, M_3, M_4)$ по определению равно точке пересечения прямых (M_1, M_2) и (M_3, M_4), проходящих соответственно через точки M_1 и M_2, M_3 и M_4;

2) значение $\varphi_2(M_1, M_2, M_3, M_4)$ по определению равно неупорядоченной паре точек, в которых пересекаются прямая (M_1, M_2) и окружность с центром в M_3 радиуса $\|M_3, M_4\|$; если прямая $(M_1 M_2)$ и окружность $(M_3, \|M_3, M_4\|)$ касаются, то $\varphi_2(M_1, M_2, M_3, M_4)$ равно их точке касания;

3) значение $\varphi_3(M_1, M_2, M_3, M_4)$ по определению равно неупорядоченной паре точек, в которых пересекаются окружности с центрами в точках M_1 и M_3 радиусов $\|M_1, M_2\|$ и $\|M_3, M_4\|$; если эти окружности касаются, то $\varphi_3(M_1, M_2, M_3, M_4)$ равно их точке касания.

Прямые и окружности, определяемые точками M_1, M_2, M_3, M_4, могут совпадать, или не пересекаться, или вообще не быть определенными, например, если $M_1=M_2$ или $M_3=M_4$. В этих случаях считаем, что отображения φ_1, φ_2, φ_3 не определены. Поэтому отображения φ_1, φ_2, φ_3 являются частично определенными. Итак, $\varphi_{1,2,3}: (\boldsymbol{R}_2)^4 \to \boldsymbol{R}^2 \cup \mathscr{P}_2(\boldsymbol{R}^2)$ — частично определенное отображение (частично определенная функция), а $P_2(X)$ обозначает семейство всех двуэлементных подмножеств множества X.

По аналогии с определением расширения поля (см. п. 1) сформулируем определение CL-расширения (расширения с помощью циркуля и линейки) какого-либо множества в арифметической плоскости $\boldsymbol{R}^2$.

Определение 1. **а) Простым CL-расширением множества $X \subseteq R^2$ называется любое множество вида**

$$X \cup \varphi_i(M_1, M_2, M_3, M_4), \qquad (*)$$

где φ_i — одно из простых CL-построений $i=1, 2, 3$ и M_1, M_2, M_3, M_4 — такие фиксированные точки из X, для которых определено данное отображение $\varphi_i(\cdot)$.

б) Множество $X' \subseteq R^2$ называется CL-расширением множества $X \subseteq R^2$, если найдется такая цепочка вложенных друг в друга множеств вида $X \rightleftarrows X_0 \subset X_1 \subset \ldots \subset X_s \rightleftarrows X'$, что при всех $i=1, \ldots, s$ множество X_i является простым CL-расширением множества X_{i-1}.

Запись $(*)$ нуждается в пояснении: если $\varphi_i(M_1, \ldots, M_4)$ — неупорядоченная пара, т. е. множество, состоящее из двух точек, то запись $(*)$ обозначает объединение двух множеств X и $\varphi_i(M_1, \ldots, M_4)$; если $\varphi_i(M_1, \ldots, M_4)$ — отдельная точка из R^2, то она обозначает объединение X и $\{\varphi_i(M_1, \ldots, M_4)\}$.

Определение 2. а) Будем говорить, что точку M в плоскости R^2 можно CL-построить (построить циркулем и линейкой), если найдется такое CL-расширение X множества, состоящего ровно из двух точек $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ и $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, что $M \in X$.

б) Будем говорить, что вещественное число x можно (CL)-построить, если можно CL-построить точку $\begin{pmatrix} x \\ 0 \end{pmatrix}$.

в) Будем говорить, что комплексное число $x+iy$ можно CL-построить, если можно CL-построить точку $\begin{pmatrix} x \\ y \end{pmatrix}$.

Предложение 2. ***а) Точку $\begin{pmatrix} x \\ y \end{pmatrix}$ можно CL-построить тогда и только тогда, когда можно CL-построить обе ее координаты x и y.***

б) Если числа a и b можно CL-построить, то можно CL-построить и числа $a+b$; $a-b$; $a \cdot b$; $\frac{a}{b}$, где $b \neq 0$. Если при этом число $a \geqslant 0$, то можно CL-построить и число $\sqrt{a}$.

▷ а) Допустим, точку $M(x, y)$ можно CL-построить. Обозначим $M_1 \rightleftarrows \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $M_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ и рассмотрим пару точек $\varphi_3(M_1, M, M_2, M) \rightleftarrows \rightleftarrows \{M, M'\}$ (см. рис. 77). Точка $\varphi_1(M_1, M_2, M, M')$ равна точке $\begin{pmatrix} x \\ 0 \end{pmatrix}$.

Итак, доказано, что число x можно CL-построить. В дальнейшем мы не будем столь формально записывать ход CL-построения. И в необходимых случаях будем просто приводить чертеж.

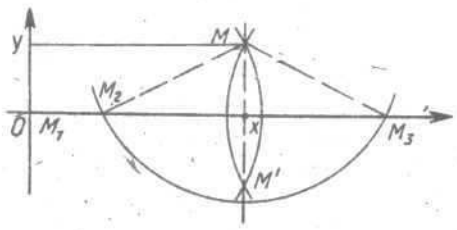


Рис. 77

Продолжим доказательство. Построим ось ординат, т. е. перпендикуляр к прямой (M_1, M_2) в точке M_1, и, опустив на нее из точки M перпендикуляр, получим точку $\begin{pmatrix} 0 \\ y \end{pmatrix}$. Образуем точки пересечения прямой (M_1, M_2) и окружности $\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ y \end{pmatrix}\right)$ и получим точки $\begin{pmatrix} y \\ 0 \end{pmatrix}$ и $\begin{pmatrix} -y \\ 0 \end{pmatrix}$. Поэтому число y можно CL-построить.

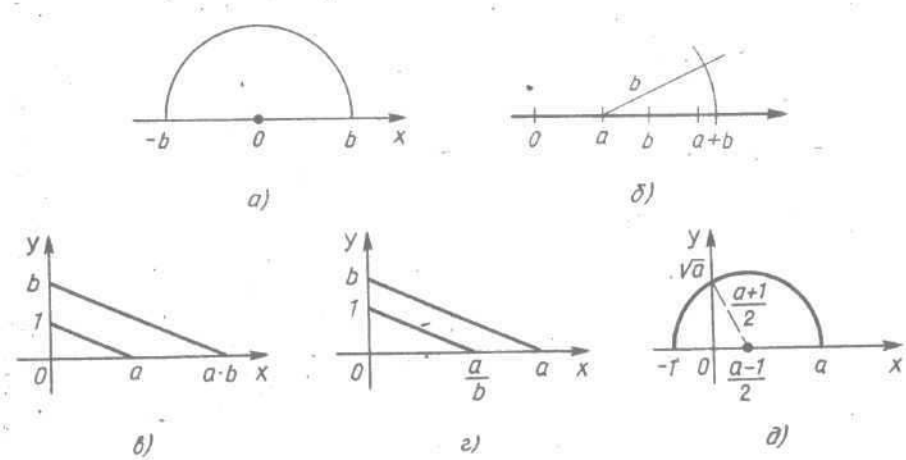


Рис. 78

Наоборот, пусть числа x и y можно CL-построить. Проведем перпендикуляры к координатным осям соответственно в точках $\begin{pmatrix} x \\ 0 \end{pmatrix}$ и $\begin{pmatrix} 0 \\ y \end{pmatrix}$ и получим точку $M = \begin{pmatrix} x \\ y \end{pmatrix}$.

б) На рисунке 78а) приведено построение числа $-b$, на рисунке 78б) — построение числа $a+b$; далее построение числа $a \cdot b$ сводится к проведению прямой, параллельной данной прямой и проходящей через данную точку (рис. 78в).

в) Построение числа $\frac{a}{b}$ также сводится к проведению параллельной прямой, проходящей через данную точку (рис. 78,г).

г) Построение числа $\sqrt{a}$ сводится к делению данного отрезка пополам (рис. 78д).

Следствие 1. Любое рациональное число можно CL-построить. Любой элемент из квадратичного расширения поля $\boldsymbol{Q}$ можно CL-построить.

Следствие 2. Если вещественные или комплексные числа a, b, c можно CL-построить, то можно CL-построить все корни квадратного уравнения $ax^2+bx+c=0$.

Упражнение 2. *Запишите с помощью цепочки простых CL-построений следующие конкретные построения:*

а) построение перпендикуляра к прямой (M_1, M_2) в точке M_1, где M_1 и M_2 — уже построенные точки и $M_1 \neq M_2$;

б) построение прямой, проходящей через точку M_3, параллельно прямой (M_1, M_2), где M_1, M_2, M_3 — построенные точки и $M_1 \neq M_2$;

в) деление отрезка пополам;

г) проведение биссектрисы угла $M_2M_1M_3$, где M_1, M_2, M_3 — уже построенные точки;

д) построение чисел $\sqrt{2}$, $\sqrt{3}$, $\sqrt{17}$, $\sqrt{23}$.

Следующая лемма и теорема 1 являются основными утверждениями этого параграфа. Они устанавливают тесную связь между квадратичными расширениями поля $\boldsymbol{Q}$ и CL-расширениями множества $\left\{\binom{0}{0}; \binom{1}{0}\right\}$.

Лемма 1. ***Пусть X — множество в арифметической плоскости R^2, координаты всех точек которого лежат в некотором поле P, и X' — любое простое CL-расширение этого множества. Тогда найдется такое простое квадратичное расширение P' поля P, что координаты всех точек из X' лежат в P'.***

▷ По определению множество X' получается из множества X добавлением значения одного из трех отображений φ_1, φ_2, φ_3, которые и были названы простыми CL-построениями. Поэтому достаточно построить поле P' в каждом из следующих трех случаев:

1) Пусть $X'=X\cup(\varphi_1(M_1, M_2, M_3, M_4))$. По определению простого CL-расширения отображение φ_1 определено в точках M_1, M_2, M_3, M_4. Обозначим x, y координаты точки $\varphi_1(M_1, M_2, M_3, M_4)$ пересечения прямых (M_1, M_2) и (M_3, M_4) (см. рис. 79). В этом случае к множеству X добавляется точка (x, y) и поле P' должно содержать подполе $P(x, y)$.

Вычислим координаты x и y. Из условия принадлежности трех точек одной прямой имеем:

$$\begin{vmatrix} x_1-x & y_1-y \\ x_2-x & y_2-y \end{vmatrix}=0 \text{ и } \begin{vmatrix} x_3-x & y_3-y \\ x_4-x & y_4-y \end{vmatrix}=0.$$

Поэтому для неизвестных x и y получается система из двух линейных уравнений

$$\begin{cases}(y_1-y_2)x+(x_2-x_1)y=x_2y_1-x_1y_2,\\(y_3-y_4)x+(x_4-x_3)y=x_4y_3-x_3y_4.\end{cases}$$

Все коэффициенты и свободные члены этой системы лежат в поле P. Переобозначив их, получим систему

$$\begin{cases}a_1x+b_1y=c_1,\\a_2x+b_2y=c_2,\end{cases} \text{ где } a_i,\ b_i,\ c_i\in P \text{ и } \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}\neq 0.$$

Отсюда получаем:

$$x=\frac{c_1b_2-c_2b_1}{a_1b_2-a_2b_1} \text{ и } y=\frac{a_1c_2-a_2c_1}{a_1b_2-a_2b_1}.$$

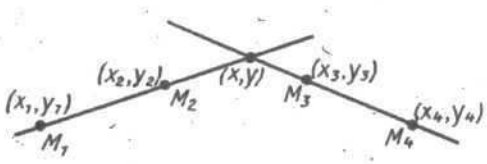


Рис. 79

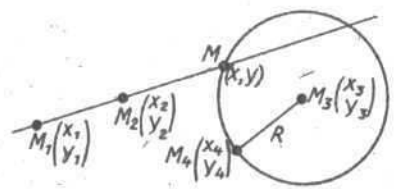


Рис. 80

Следовательно, $x\in P$ и $y\in P$. Поэтому в данном случае $P(x, y)=P$ и в качестве поля P' можно взять само поле P, так как поле P является простым квадратичным расширением самого себя.

2) Пусть $X'=X\cup(\varphi_2(M_1, M_2, M_3, M_4))$. Отображение φ_2 определено для данных аргументов, и к множеству X добавляется точка (x, y), показанная на рисунке 80, и, может быть, еще другая точка (x', y'). Поле P' должно содержать подполе $P(x, y, x', y')$.

Для неизвестных x и y получается система из двух уравнений

$$\begin{cases}(y_1-y_2)x+(x_2-x_1)y=x_2y_1-x_1y_2,\\(x-x_3)^2+(y-y_3)^2=(x_4-x_3)^2+(y_4-y_3)^2.\end{cases}$$

Переобозначив ее коэффициенты, получим систему

$$\begin{cases}ax+by=c,\\(x-A)^2+(y-B)^2=C^2,\end{cases} \text{ где } a,\ b,\ c,\ A,\ B,\ C^2\in P.$$

Так как точки M_1 и M_2 различны, то число a или число b не равно нулю.

Пусть, например, $b\neq 0$. Тогда из первого уравнения $y=\frac{c-ax}{b}$, и, подставив это выражение во второе уравнение, получим квадратное уравнение с коэффициентами α, β, γ из поля P:

$$\alpha\cdot x^2+\beta\cdot x+\gamma=0.$$

Это уравнение должно иметь решение (иначе отображение φ_2 не определено). Поэтому его дискриминант $\delta=\beta^2-4\alpha\gamma\geqslant 0$. Из формулы для корней квадратного уравнения следует, что x является элементом поля $P(\sqrt{\delta})$, которое есть простое квадратичное расширение поля P.

Поскольку $y=\frac{c-a\cdot x}{b}$, то и $y\in P(\sqrt{\delta})$. Аналогично получаем, что x', $y'\in P(\sqrt{\delta})$. Итак, $P'=P(\sqrt{\delta})$.

3) Пусть $X'=X\cup(\varphi_3(M_1, M_2, M_3, M_4))$. Отображение φ_3 определено для данных аргументов, и к множеству X добавляется точка (x, y), показанная на рисунке 81, и может быть еще другая точка (x', y'). Поле P' должно содержать подполе $P(x, y, x', y')$.

Для неизвестных x и y получается система из двух уравнений

$$\begin{cases}(x-x_1)^2+(y-y_1)^2=(x_2-x_1)^2+(y_2-y_1)^2=c^2,\\(x-x_3)^2+(y-y_3)^2=(x_4-x_3)^2+(y_4-y_3)^2=r^2.\end{cases}$$

Заменяя первое уравнение на разность первого и второго уравнений, получаем

$$\begin{cases}2(x_3-x_1)x+2(y_3-y_1)y=\\=c^2-r^2+x_3^2-x_1^2+y_3^2-y_1^2,\\(x-x_3)^2+(y-y_3)^2=r^2.\end{cases}$$

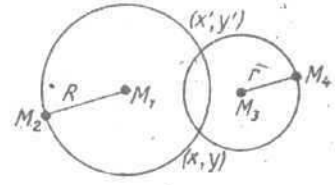


Рис. 81

Переобозначая коэффициенты, получаем:

$$ax+by=C,$$
$$(x-A)^2+(y-B)^2=r^2, \text{ где } a,\ b,\ C,\ A,\ B,\ r^2\in P.$$

Как и во втором случае, число a или число b не равно нулю, так как точки M_1 и M_3 различны. Поэтому, рассуждая аналогично второму случаю, получаем, что x и y лежат в некотором простом квадратичном расширении поля P. ■

Теорема 1. ***Точку $M(x, y)$ арифметической плоскости R^2 можно CL-построить в том и только в том случае, когда найдется такое квадратичное расширение P поля Q, что $x\in P$ и $y\in P$.***

▷ Достаточность. Пусть $\boldsymbol{Q}\rightleftarrows P_0\subset P_1\subset\ldots\subset P_s\rightleftarrows P$ — такая цепочка вложенных друг в друга полей, что при всех $i=1, 2, \ldots, s$ поле P_i является простым квадратичным расширением поля P_{i-1} и $x, y\in P$. По следствию 1 к предложению 2 все рациональные числа можно CL-построить.

Так как любой элемент из простого квадратичного расширения P_1 поля Q является корнем квадратного уравнения с рациональными коэффициентами (см. упражнение 1г), то по следствию 2 к предложению 2 все числа из поля P_1 можно CL-построить. Аналогичным образом все числа из полей $P_2, P_3, \ldots, P_s=P$ можно CL-построить. Если теперь $x\in P$ и $y\in P$, то по предложению 2,а можно CL-построить и точку $M=(x, y)$.

Необходимость. Пусть точку M можно CL-построить, т. е. найдется такая цепочка множеств

$$\left\{\binom{0}{0}, \binom{1}{0}\right\}\rightleftarrows X_0\subset X_1\subset\ldots\subset X_s\rightleftarrows X,$$

вложенных друг в друга, что при всех $i=1, 2, \ldots, s$ множество X_i является простым (CL)-расширением множества X_{i-1} и $M\in X$. Координаты всех точек множества X_1 лежат в некотором простом квадратичном расширении P_1 поля $\boldsymbol{Q}$, так как координаты всех точек множества X_0 лежат в $\boldsymbol{Q}$ и применима лемма 1.

Аналогично координаты всех точек множества X_2 лежат в некотором простом квадратичном расширении P_2 поля P_1 и т. д. Поэтому, пройдя всю цепочку множеств $X_0\subset X_1\subset\ldots\subset X_s=X$, получим цепочку $\boldsymbol{Q}\rightleftarrows P_0\subset P_1\subset\ldots\subset P_s\rightleftarrows P$ простых квадратичных расширений поля $\boldsymbol{Q}$ и, следовательно, получим квадратичное расширение P поля $\boldsymbol{Q}$, в котором лежат координаты всех точек множества X. ■

3. Проблемы удвоения куба, трисекции угла и построения правильного семиугольника с помощью циркуля и линейки. В этом пункте мы уточним постановки упомянутых классических проблем и докажем невозможность их положительного решения. План доказательства невозможности соответствующих построений будет таким:

— сначала покажем, что если какая-то из этих проблем имеет

положительное решение, то можно CL-построить хотя бы один вещественный корень определенного кубического уравнения с целыми коэффициентами;

— по следствию к предложению 1 получим, что эти кубические уравнения не имеют ни одного решения ни в каком квадратичном расширении поля $\boldsymbol{Q}$;

— отсюда по основной теореме 1 получим невозможность CL-построения корней этих кубических уравнений и заключим, что положительное решение каждой из этих проблем невозможно.

Предложение 3. ***Если проблемы удвоения куба, трисекции угла и построения правильного семиугольника имеют положительное решение, то соответственно следующие кубические уравнения имеют вещественные корни, которые можно CL-построить:***

а) $x^3-2=0$;
б) $x^3-3x-1=0$;
в) $x^3+x^2-2x-1=0$.

▷ а) Проблема удвоения куба заключается в следующем.

Дан куб единичного объема, точнее, дан отрезок, являющийся ребром такого куба, т. е. отрезок единичной длины. Требуется найти куб, объем которого вдвое больше объема исходного куба. Точнее, найти ребро такого куба.

Длина искомого ребра является вещественным корнем кубического уравнения $x^3-2=0$. Слова «требуется найти» означают «найти циркулем и линейкой». Поэтому нужно найти конечную последовательность простых CL-построений, которые начинаются с единичного отрезка и заканчиваются отрезком длиной $\sqrt[3]{2}$. Поэтому положительное решение проблемы удвоения куба эквивалентно CL-построению вещественного числа $\sqrt[3]{2}$.

б) Если возможна трисекция любого угла, то она возможна, в частности, для угла в 60°, одна из сторон которого совпадает с положительным направлением оси Ox. При этом слово «возможна», как и в п. а, означает существование некоторой последовательности простых CL-построений, которые начинаются с точек (0, 0) и (1, 0) и закончатся любой точкой M на луче, образующем угол в 20° с положительным направлением оси Ox (см. рис. 82).

Допустим, что такую точку можно CL-построить. Тогда число $\cos 20°$ можно CL-построить, так как можно CL-построить точку $(\cos 20°, \sin 20°)$ в виде значения отображения $\varphi_2\left(\binom{0}{0}, M, \binom{0}{0}\binom{1}{0}\right)$.

По известной тригонометрической формуле получаем:

$$\frac{1}{2}=\cos 60°=4\cos^3 20°-3\cos 20°$$

или

$$1=(2\cos 20°)^3-3(2\cos 20°).$$

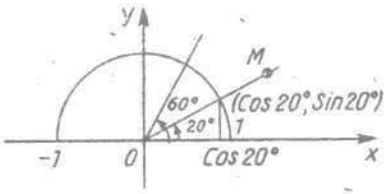


Рис. 82

Поэтому число $2\cos 20°$ является вещественным корнем уравнения

$$x^3-3x-1=0.$$

Итак, из возможности положительного решения проблемы трисекции угла следует, что можно CL-построить хотя бы один вещественный корень уравнения

$$x^3-3x-1=0.$$

в) Проблема построения циркулем и линейкой правильного семиугольника состоит в CL-построении правильного семиугольника, вписанного в окружность единичного радиуса. Без ограничения общности можно считать, что центр окружности является началом координат и одна из вершин правильного семиугольника совпадает с точкой $\binom{1}{0}$.

Что означают слова «CL-построить правильный семиугольник»? Естественно считать, что они означают возможность CL-построить все его вершины или, что эквивалентно (докажите), означают возможность CL-построить его вершину в верхней полуплоскости, ближайшую к точке 1.

Этой вершиной является точка $e^{i\frac{2\pi}{7}}$, так как $(e^{i\frac{2\pi}{7}})^7=e^{2\pi i}=1$. Если можно CL-построить точку $e^{i\frac{2\pi}{7}}$, то можно построить и точку $e^{i\frac{\pi}{7}}$, так как она является пересечением биссектрисы угла $\angle 1, \bar{0}, e^{i\frac{2\pi}{7}}$ и единичной окружности с центром в точке $\bar{0}$ (см. упражнение 2, г).

Итак, из положительного решения этой проблемы вытекает возможность CL-построения точки $e^{i\frac{\pi}{7}}=\left(\cos\frac{\pi}{7}, \sin\frac{\pi}{7}\right)$ и CL-построения вещественного числа $2\cdot\cos\frac{\pi}{7}$.

Рассмотрим комплексное число $z=\cos\frac{\pi}{7}+i\sin\frac{\pi}{7}$. Тогда $z^7=1$ или $(z^7-1)=(z-1)(z^6+z^5+z^4+z^3+z^2+z+1)=0$. Так как $z\neq 1$, то нулю равен второй сомножитель. Поделим его на z^3 и получим

$$\left(z^3+\frac{1}{z^3}\right)+\left(z^2+\frac{1}{z^2}\right)+\left(z+\frac{1}{z}\right)+1=0.$$

Слагаемые, стоящие в скобках, являются вещественными числами, так как числа z^n и z^{-n} комплексно сопряжены.

Рассмотрим вещественное число $x_1=z+\frac{1}{z}=2\cdot\cos\frac{\pi}{7}$. Тогда

$$x_1^2=\left(z^2+\frac{1}{z^2}\right)+2 \text{ и } x_1^3=\left(z^3+\frac{1}{z^3}\right)+\frac{3z^2}{z}+\frac{3z}{z^2}=\left(z^3+\frac{1}{z^3}\right)+3x_1.$$

Поэтому $x_1^3-3x_1+x_1^2-2+x_1+1=x_1^3+x_1^2-2x_1-1=0$.

Итак, из возможности положительного решения проблемы построения правильного семиугольника следует возможность CL-построения одного из вещественных корней уравнения $x^3+x^2-2x-1=0$. ■

Теорема 2. ***Проблемы удвоения куба, трисекции угла и построения правильного семиугольника с помощью циркуля и линейки неразрешимы.***

▷ По предложению 3 для доказательства теоремы 2 нужно установить, что нельзя CL-построить ни один из вещественных корней следующих кубических уравнений:

а) $x^3-2=0$;

б) $x^3-3x-1=0$;

в) $x^3+x^2-2\cdot x-1=0$.

По следствию к предложению 1 ни один из этих вещественных корней не лежит ни в каком квадратичном расширении поля Q. По теореме 1 это как раз и означает невозможность их (CL)-построения. ■

Замечание. Иногда в понятие «построение циркулем и линейкой» включают еще одну операцию: выбор произвольной точки плоскости. Эта операция обсуждается в следующем пункте.

4. Геометрические построения, включающие операцию выбора произвольной точки в заданной фигуре. В задачах на построение, кроме первоначальных точек $\binom{0}{0}$ и $\binom{1}{0}$, часто с самого начала задаются еще некоторые простейшие плоские фигуры, например точки, прямые, окружности, многоугольники и т. п.

Например, в такой задаче на построение, как «вписать в данный треугольник квадрат, все вершины которого лежат на сторонах исходного треугольника», явно подразумевается наличие наперед заданного произвольного треугольника. Фактически такой треугольник произвольно выбирается перед началом построения. Для многих других построений также характерно, что на некоторых их шагах используются операции типа операций *выбора произвольной точки, проведения произвольной прямой, выбора произвольной точки на данной прямой* и т. п. При этом понятие CL-построения на первый взгляд заменяется на некоторое более широкое понятие построения.

Например, решение упомянутой выше задачи выглядит так. Выбирается произвольная точка D на одной из сторон треугольника ABC. Затем CL-строится квадрат с вершиной в точке D, с двумя другими вершинами на сторонах треугольника ABC и еще с одной вершиной D' внутри треугольника ABC (рис. 83). После

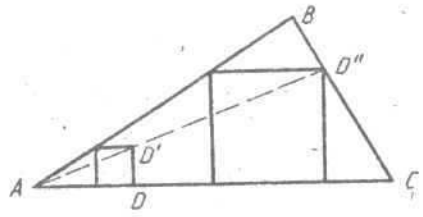


Рис. 83

этого CL-строится точка $D'' \rightleftarrows (BC) \cap (AD')$ и затем строится искомый квадрат.

Отметим, что все упомянутые выше операции произвольного выбора сводятся к операции выбора произвольной точки. Например, выбор произвольной прямой означает выбор двух различных точек этой прямой, выбор произвольной окружности означает выбор ее центра и любой точки, лежащей на ней и т. д. Поэтому ниже будем говорить лишь об операции выбора произвольной точки.

Пусть различные точки A и B уже CL-построены и для решения некоторой задачи на построение нужно выбрать произвольную точку отрезка $[A, B]$. Как понимать термин «произвольную»? Если его понимать как «любую точку, лежащую на этом отрезке», то мы попадаем в трудное положение: таких точек континуальное множество.

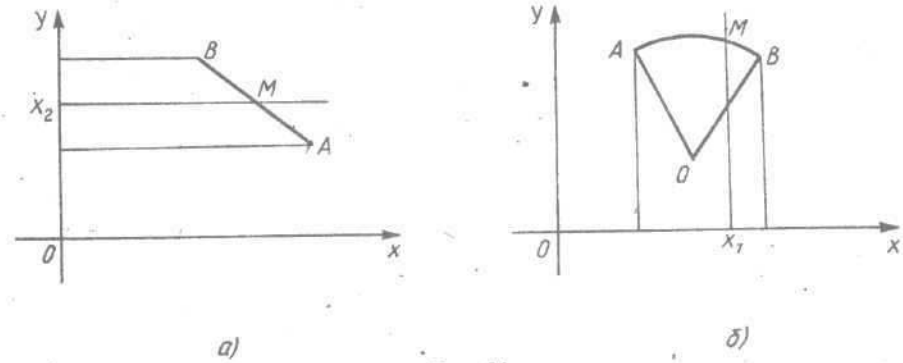
Упражнение 3. *Докажите, что множество всех точек в $\boldsymbol{R}^2$, которые можно CL-построить, счетно.*

Поэтому нельзя CL-построить в этом смысле произвольную точку, лежащую даже на одном каком-то отрезке $[A, B]$, где $A \neq B$. Впрочем, интуитивно ясно, что нельзя явно описать, построить, сконструировать каждую по отдельности точку из несчетного множества точек. Поэтому слова «построить произвольную точку» отрезка $[A, B]$ понимают как возможность построить хотя бы одну точку данного отрезка, отличную от его концов A и B. Как мы сейчас увидим, на любом отрезке $[A, B]$ имеется хотя бы одна точка, которую можно CL-построить. Поэтому в качестве произвольной точки данного отрезка можно выбирать именно ее. Итак, последовательное добавление на каждом шаге CL-построения в таком смысле произвольной точки не приводит к получению точек, которые нельзя было бы построить без операции выбора произвольной точки. Таким образом, операция выбора произвольной точки не расширяет множества всех точек, которые можно CL-построить!

Чтобы доказать упомянутые выше утверждения, спроектируем отрезок $[A, B]$ на каждую из осей координат. Одна из этих проекций является невырожденным отрезком, в котором (по свойству плотности множества $\boldsymbol{Q}$ в множестве $\boldsymbol{R}$) содержится рациональное число x_2 (рис. 84, *а*). Рациональное число, конечно, можно CL-построить. Восстановив перпендикуляр из этой точки до пересечения с отрезком $[A, B]$, найдем в этом отрезке CL-построенную точку M.

Аналогичным образом, если требуется выбрать произвольную точку на уже CL-построенной окружности или дуге окружности, ее можно выбрать из множества точек плоскости, которые сами по себе могут быть CL-построены (рис. 84, *б*).

Свойство плотности множества $\boldsymbol{Q}$ в множестве $\boldsymbol{R}$ помогает и в тех случаях, когда требуется выбрать произвольную точку внутри или вне заданной плоской фигуры ω. При этом всегда подразумевается, что внутренности фигуры ω и ее дополнения $C\omega$



а) *б)*

Рис. 84

не пусты. Тогда плотность множества $\boldsymbol{Q}^2$ в $\boldsymbol{R}^2$ означает, что в ω и в $C\omega$ найдутся точки с рациональными координатами и, следовательно, найдутся точки, которые сами по себе могут быть CL-построены.

Итак, выбор в таком смысле (!) произвольных точек на отрезках, прямых, дугах и т. д. приводит к образованию точек, которые могут быть сами CL-построены.

В заключение коснемся еще одного расширения понятия CL-построения. Если изменить определения 1, б и 2, а в том и только в том отношении, что на 0-м шаге CL-расширения, кроме точек $\binom{0}{0}$ и $\binom{1}{0}$, по определению считается построенной еще одна фиксированная, наперед заданная точка M_1 с координатами x_1 и y_1, то получится новое, более широкое понятие построения, которое назовем параметрическим CL-построением с числовыми параметрами x_1 и y_1. Если на 0-м шаге по определению считаются построенными несколько таких точек $M_1 \rightleftarrows \binom{x_1}{y_1}$, ..., $M_n \rightleftarrows \binom{x_n}{y_n}$, то получается понятие параметрического CL-построения с числовыми параметрами $x_1, y_1, \ldots, x_n, y_n$. Аналогично доказательству теоремы 2 легко проверить, что точку M в арифметической плоскости $\boldsymbol{R}^2$ можно параметрически CL-построить с параметрами $x_1, y_1, \ldots, x_n, y_n$ в том и только в том случае, когда обе ее координаты лежат в некотором квадратичном расширении поля $\boldsymbol{Q}(x_1, y_1, \ldots, x_n, y_n)$. Далее получаются результаты, аналогичные тем, которые изложены выше. В частности, если понятие параметрического CL-построения расширить, допуская выбор в том смысле, как и раньше, произвольных точек, то множество всех точек, которые можно параметрически CL-построить, не расширится.

Поле $\boldsymbol{Q}(x_1, y_1, \ldots, x_n, y_n)$ состоит из счетного числа элементов. Более того, счетно любое квадратичное расширение этого поля и счетно объединение всех его квадратичных расширений. Поэтому множество всех точек, которые можно параметрически CL-построить с любыми наперед фиксированными параметрами $x_1, y_1, \ldots,$

x_n, y_n, счетно. Несчетность множества всех точек арифметической плоскости $\boldsymbol{R}^2$ означает, что для любых фиксированных параметров x_1, y_1, ..., x_n, y_n найдутся точки, которые нельзя параметрически CL-построить с этими параметрами.

5. Геометрические построения с помощью одного циркуля. В этом пункте покажем, что любую точку, которую можно CL-построить (т. е. построить с помощью циркуля и линейки), можно и C-построить, т. е. построить с помощью только одного циркуля. Этот факт был доказан Мором в 1672 году. Изложим доказательство Г. Моро о совпадении множества точек, которые можно C-построить, и множества точек, которые можно CL-построить.

Начнем с определения C-построения. Оно формулируется в точности так же, как определение 1, с единственным отличием: из функций φ_1, φ_2, φ_3 разрешается использовать только функцию φ_3. Остальные определения п. 2 остаются без всяких изменений.

Идея доказательства совпадения множеств точек плоскости, которые можно CL-построить и C-построить, состоит в том, что если коэффициенты квадратного уравнения можно C-построить, то оба его корня можно C-построить. Кроме того, все целые числа можно C-построить. Отсюда следует, что все точки любого квадратичного расширения поля $\boldsymbol{Q}$ можно C-построить (объясните почему); по теореме 1 получаем совпадение двух интересующих нас множеств точек плоскости.

Лемма 2. ***Если A и B — любые точки на плоскости, то можно C-построить две точки, расстояние между которыми равно $2\|\overline{A, B}\|$, и две точки, расстояние между которыми равно $\sqrt{3}\times$ $\times\|\overline{A, B}\|$.***

▷ Обозначим $p \rightleftarrows \|\overline{A, B}\|$ (см. рис. 85).

Проведем окружность радиуса p с центром в точке B, и пусть:

а) C — точка пересечения исходной окружности с окружностью радиуса p с центром в точке A;

б) D — точка пересечения исходной окружности с окружностью радиуса p с центром в точке C, отличная от точки A;

в) E — точка пересечения исходной окружности с окружностью радиуса p с центром в точке D, отличная от точки C.

Тогда $\|\overline{A, E}\| = 2\cdot p$ и $\|\overline{A, D}\| = \sqrt{3}\cdot p$. ■

Лемма 3. ***Если на плоскости заданы две пары точек, расстояние между первыми двумя точками равно p и между вторыми двумя точками равно g, то можно C-построить пары точек, расстояния между которыми равны соответственно числам:***

а) $\sqrt{p^2-g^2}$ при $p>g$,

б) $\sqrt{2}\cdot p$,

в) $\sqrt{p^2+g^2}$,

г) $p+g$ и $p-g$ при $p>g$,

д) $2\sqrt{pg}$.

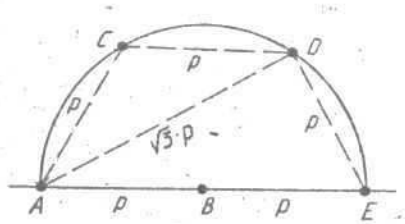


Рис. 85

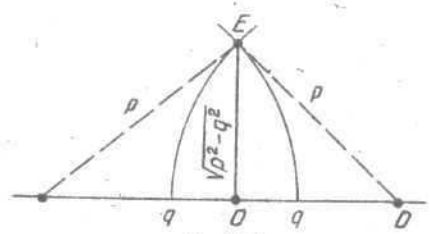


Рис. 86

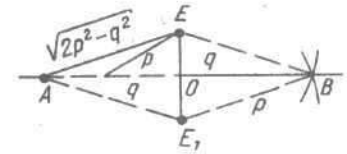


Рис. 87

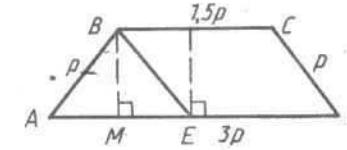


Рис. 88

▷ Ограничимся рисунком 86 с изображениями искомых построений. Отметим, что фактически не нужно соединять отрезками никакие точки, показанные на этом рисунке.

а) По лемме 2 построим отрезок длины $2\cdot g$ (его середину обозначим O) и из его концов проведем окружности радиса p. Возьмем любую из точек пересечения этих окружностей (такая точка существует, так как $p>g$, обозначим ее E). Пара точек $\langle O, E\rangle$ искомая. Только для наглядности искомые пары точек соединены сплошными отрезками, а вспомогательные точки соединены пунктирными отрезками.

б) По лемме 2 построим отрезок $\sqrt{3}\cdot p$ и затем применим п. а) к расстояниям $\sqrt{3}\cdot p$ и p, получим $\sqrt{2}\cdot p = \sqrt{(\sqrt{3}p)^2 - p^2}$.

в) Применим а) к расстояниям $\sqrt{2}\cdot p$ (п. б) и $\sqrt{p^2-g^2}$ (п. а), получим $\sqrt{p^2+g^2} = \sqrt{(\sqrt{2}p)^2 - (\sqrt{p^2-g^2})^2}$.

г) По п. а) (см. рис. 87) по отрезкам p и g образуем отрезок OE и симметричный ему отрезок OE_1 (точка E_1 определяется как вторая точка пересечения окружностей, о которых говорилось в п. а). Из точек E и E_1 проведем окружности радиуса $\sqrt{2\cdot p^2 - g^2}$ (такой радиус C-построим по пунктам б), а)). Обозначим точки пересечения этих окружностей A и B. Расстояние между A и началом отрезка g равно $p-g$, а расстояние между B и началом отрезка g равно $p+g$.

д) Все получается из равенства $2\sqrt{pg} = \sqrt{(p+g)^2 - (p-g)^2}$. ■

Лемма 4. ***Если в трапеции (рис. 88) расстояние $|A, B| = |C, D| = p$, и $|B, C| = 1{,}5p$, и $|A, D| = 3p$, и E принадлежит $[AD]$, а $|B, E| = |C, E| = p$, то E — середина $[AD]$.***

▷ Если точка M — основание перпендикуляра, опущенного из точки B на отрезок $[A, D]$, то, очевидно, $|M, E| = \frac{3}{4}p$, и так как $|A, E| =$

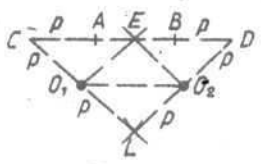


Рис. 89

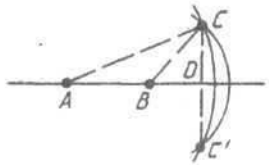


Рис. 90

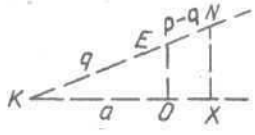


Рис. 91

$=\frac{3}{2}p$, то треугольник $\triangle ABE$ — равнобедренный. Поэтому $|B, E| = |A, B| = p$. ■

Лемма 5. ***Если на плоскости заданы две точки A и B на расстоянии p друг от друга, то можно C-построить точки, находящиеся на расстоянии $\frac{p}{2}$ друг от друга.***

▷ По лемме 2 построим такие две точки C и D, которые расположены, как показано на рисунке 89.

Пусть L — любая из двух точек пересечения окружностей радиусов $2p$ с центрами в точках C и D. Проведем окружности радиуса p с центрами в точках C и L. Получим точку их касания — точку O_1. Она является серединой отрезка $[C, L]$. Аналогично найдем точку O_2 — середину отрезка $[D, L]$. Наконец, пусть E — точка пересечения окружностей с центрами в точках O_1 и O_2 и радиусов p, отличная от L. По лемме 4 точка E — середина отрезка $[A, B]$. ■

Лемма 6. ***а) Для любых трех точек A, B, C, не лежащих на одной прямой, можно C-построить основание D перпендикуляра, опущенного из C на прямую, проходящую через точки A и B (рис. 90).***

б) Если на плоскости даны точки A, B, C, D и $|A, B| = p$, $|C, D| = g$, то можно C-построить точки на расстоянии $\sqrt{p \cdot g}$.

в) Если на плоскости даны точки A, B, C, D, E, F и $|A, B| = p$, $|C, D| = g$, $|E, F| = a$, то можно C-построить точки на таком расстоянии x, что $\frac{x}{a} = \frac{p}{g}$ (рис. 91).

▷ а) Пусть C' — точка пересечения окружностей с центрами A и B и радиусов $|A, C|$ и $|B, C|$ соответственно, отличная от C. По лемме 5 можно C-построить середину D отрезка $[C, C']$. Она является искомым основанием перпендикуляра.

б) Следует из леммы 3д и леммы 5.

в) Можно считать $g > a$, так как в противном случае заменим p и g (по лемме 3г) на соответственно $n \cdot p$ и $n \cdot g$. Тогда $|K, X| = \frac{a \cdot p}{q}$. ■

Теорема 3. ***Произвольную точку в арифметической плоскости R^2 можно CL-построить в том и только в том случае, когда ее можно C-построить.***

▷ По определению возможность C-построения некоторой точки

влечет возможность ее CL-построения. Докажем обратную импликацию.

По лемме 3г можно C-построить все целые числа, т. е. точки вида $\binom{n}{0}$ при всех $n \in \mathbf{Z}$. По лемме 6в можно C-построить все рациональные числа, т. е. точки вида $\binom{r}{0}$ при всех $r \in \mathbf{Q}$. Если числа a, b, c можно C-построить, то по леммам 3, 5, 6 из формулы $z = \frac{-b \pm \sqrt{b^2 - (2\sqrt{ac})^2}}{2a}$ для корней квадратного уравнения $az^2 + bz + c = 0$ следует возможность C-построения его корней.

Следовательно, доказана возможность C-построения любого числа из любого квадратичного расширения поля $\mathbf{Q}$. По теореме 1, если точку M можно CL-построить, то обе ее координаты лежат в некотором квадратичном расширении поля $\mathbf{Q}$. Поэтому ее координаты можно C-построить. По лемме 3в можно C-построить и саму точку. ■

§ 2. ЗАДАЧА О РАЗРЕШИМОСТИ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ В РАДИКАЛАХ. КРИТЕРИЙ РАЗРЕШИМОСТИ. ПРИМЕР НЕРАЗРЕШИМОГО В РАДИКАЛАХ АЛГЕБРАИЧЕСКОГО УРАВНЕНИЯ 5-Й СТЕПЕНИ

1. Постановка задачи о разрешимости алгебраического уравнения в радикалах. Пусть $f(x) = x^n + a_1 \cdot x^{n-1} + \ldots + a_n$ — произвольный фиксированный многочлен с комплексными (или, в частности, с вещественными) коэффициентами $a_1, \ldots, a_n$ и старшим коэффициентом, равным 1. Иными словами $f \in \mathbf{C}[x]$ или, в частности, $f \in \mathbf{R}[x]$. Нас интересует очень старый, классический вопрос: можно ли комплексные корни уравнения $f(x) = 0$ (которые по основной теореме алгебры, разумеется, существуют) выразить через коэффициенты данного уравнения с помощью таких рациональных действий над числами, как сложение $+$, вычитание $-$, умножение $\cdot$, деление:, а также извлечение корня любой степени? Точнее, существует ли такая конечная последовательность множеств, у которой 0-е множество состоит из всех коэффициентов исходного уравнения, а $(i+1)$-е множество получается из i-го множества добавлением ровно одного числа; это число является результатом выполнения одного из рациональных действий или действия извлечения корня, которые применяются к числам, содержащимся в i-м множестве; при этом последнее множество в этой конечной последовательности множеств содержит хотя бы один корень исходного уравнения? Другой вопрос: возможна ли аналогичная последовательность множеств, у которой последнее множество содержит все корни исходного уравнения? Подчеркнем: речь здесь идет о конечных последовательностях конечных множеств.

Например, если $f=x^2+2x+\frac{1}{2}$, то эта последовательность конечных множеств выглядит следующим образом:

0) $\{1, 2, 1/2\}$,
1) $\{1, 2, 1/2, 0\}$, где $0=1-1$,
2) $\{1, 2, 1/2, 0, -2\}$, где $-2=0-2$,
3) $\{1, 2, 1/2, 0, -2, \sqrt{2}\}$,
4) $\{1, 2, 1/2, 0, -2, \sqrt{2}, -2+\sqrt{2}\}$,
5) $\{1, 2, 1/2, 0, -2, \sqrt{2}, -2+\sqrt{2}, \frac{-2+\sqrt{2}}{2}\}$

и последнее число в 5-ом множестве является корнем исходного уравнения. Второй корень исходного уравнения получается аналогичной цепочкой множеств, в которой 4-е множество вместо числа $(-2+\sqrt{2})$ содержит число $(-2-\sqrt{2})$. Конечно, эти последовательности множеств не являются ни единственно возможными, ни наиболее короткими, однако мы не интересуемся этой стороной дела.

В рассматриваемой задаче дано уравнение $f(x)=0$. Конкретно это значит, что даны его коэффициенты и, в частности, старший коэффициент 1. Поэтому не вызывает сомнения то положение, что 0-е множество состоит из чисел, 1, a_1, ..., a_n. Рациональные действия отражают структуру полей $\boldsymbol{R}$ и $\boldsymbol{C}$, которые являются основными полями, по крайней мере в элементарной математике (обоснование этого тезиса содержится в главе V). Иными словами, рациональные действия над числами — простейшие действия (операции, функции), по крайней мере в элементарной математике. Поэтому не вызывает сомнения и другое положение: если образовано множество P_{i-1}, содержащее числа a_1 и a_2, то законно его расширение одним из чисел a_1+a_2, a_1-a_2, $a_1\cdot a_2$, $a_1:a_2$.

Расширенное множество обозначается P_i. Переход от множества P_{i-1} к множеству P_i можно аргументировать и таким образом: если числа a_1 и a_2 CL-построены, то можно CL-построить и любое из чисел a_1+a_2, a_1-a_2, $a_1\cdot a_2$, $a_1:a_2$ (см. § 1.2). Почему при переходе от множества P_{i-1} к множеству P_i, кроме рациональных операций, допускается еще операция извлечения корня? Отчасти это объясняется возможностью CL-построить число $\sqrt{a}$ (см. § 1.2), отчасти историческими причинами, а отчасти тем, что число $\sqrt{a}$ определяется простейшим уравнением $x^2=a$, часто возникающим на практике.

Современные теоретико-множественные традиции создали привычку рассматривать любой интересующий нас объект как элемент совокупности всех подобных ему объектов. (Это часто приводит к возникновению бесконечных множеств там, где по существу нужны только конечные множества.) Переформулируем в этом духе обсуждаемую нами задачу. Нулевое множество — это поле $P_0 \rightleftarrows \boldsymbol{Q}(a_1, \ldots, a_n)$. Оно называется полем коэффициентов исходного уравнения $f(x)=0$. (Если $a_1, \ldots, a_n \in \boldsymbol{R}$, то $P_0 \subseteq \boldsymbol{R}$.)

Иногда из каких-то внешних соображений разумно считать, что 0-е множество содержит, кроме коэффициентов, еще какие-то числа. Иными словами, 0-е множество совпадает с каким-то числовым полем P, включающим поле коэффициентов P_0. В дальнейшем, как правило, будем рассматривать такой более общий случай.

Если уже образовано $(i-1)$-е множество — поле P_{i-1}, то переходят к i-му множеству — полю $P_i \rightleftarrows P_{i-1}(\sqrt[m]{a_{i-1}})$, где a_{i-1} — какой-то фиксированный элемент из P_{i-1} и, как всегда, $\sqrt[m]{\cdot}$ по определению обозначает один из корней уравнения $x^m=a_{i-1}$. Удобно определить поле P_i несколько шире: как $P_i \rightleftarrows P_{i-1}(\zeta, \sqrt[m]{a_{i-1}})$, где $\zeta \rightleftarrows e^{i\cdot\frac{2\pi}{m}}$.

Как мы знаем из главы I, выполняется $\xi^m=\left(e^{i\frac{2\pi}{m}}\right)^m=e^{2\pi i}=\cos 2\pi+\sin 2\pi=1$, т. е. ξ — один из корней $\sqrt[m]{1}$. Кажется вполне естественным получать корни исходного уравнения, используя не только корень m-й степени из какого-то заранее образованного числа a_{i-1}, но и корни той же степени из числа 1. Кроме того, число ξ не просто один из корней m-й степени из числа 1: любой корень m-й степени из 1 получается из числа ξ в виде ξ^j, где $1\leqslant j\leqslant m$. Действительно, $\xi^j=e^{i\frac{2\pi j}{m}}$ и $(\xi^j)^m=e^{2\pi ji}=\cos 2\pi j+i\sin 2\pi j=1$. Поскольку все числа ξ^j при $1\leqslant j\leqslant m$ различны и всего их m, то они исчерпывают все m чисел вида $\sqrt[m]{1}$, т. е. все m корней уравнения $x^m=1$. Число ξ называется первообразным корнем m-й степени.

Такое определение поля P_i имеет то важное техническое преимущество, что все числа вида $\sqrt[m]{a_{i-1}}$, т. е. все m корней уравнения $x^m=a_{i-1}$, содержатся в P_i. Действительно, пусть θ — какой-то один корень m-й степени из числа a_{i-1}. Тогда все m чисел $\xi\cdot\theta$, ..., $\xi^j\cdot\theta$, ..., $\xi^m\cdot\theta$ различны и для них выполняется $(\xi^j\theta)^m=\xi^{jm}\cdot\theta^m=(\xi^m)^j\cdot a_{i-1}=1\cdot a_{i-1}=a_{i-1}$. Поэтому они исчерпывают все корни уравнения $x^m=a_{i-1}$.

Можно поступить еще следующим эквивалентным образом: добавим к нулевому множеству P все числа вида $e^{i\frac{2\pi}{m}}$, где $m=2, 3, 4 \ldots$. Иными словами, с самого начала заменить 0-е поле P на его расширение, содержащее множество $\left\{e^{i\frac{2\pi}{m}} \mid m\in\boldsymbol{N}_{\geqslant 2}\right\}$. Тогда для любого i-го множества выполняется

$$P_{i-1}(\xi, \sqrt[m]{a_{i-1}})=P_{i-1}(\sqrt[m]{a_{i-1}}),$$

так как все такие числа ξ содержатся в P и, следовательно, в P_{i-1}. То обстоятельство, что поле P_i содержит соответствующие первообразные корни, будет существенно использоваться в дальнейшем.

(Мы использовали понятие поля, порожденного числовым полем P и некоторым множеством X, где $X\subseteq\boldsymbol{C}$. Оно определяется как наименьшее числовое поле, содержащее P и X. Ясно, что такое поле совпадает с пересечением всех числовых полей, содержащих множества P и X.)

Поле $P_i \neq P_{i-1}(\zeta \sqrt[m]{a_{i-1}})$ называется простым радикальным расширением поля P_{i-1} (ср. § 1.1). За s шагов образуется цепочка (последовательность) полей

$$P \subset P_1 \subset \ldots \subset P_s, \qquad (*)$$

которая называется радикальной цепочкой или радикальным рядом длины s. Последнее из этих полей, т. е. поле P_s, называется радикальным расширением исходного поля P.

Определение 3. **а) Уравнение $f(x)=0$ называется разрешимым в радикалах, если существует хотя бы одна радикальная цепочка вида $(*)$, для которой поле P_s содержит все корни этого уравнения.**

б) Говорят, что уравнение $f(x)=0$ имеет корень, выражающийся в радикалах, если существует хотя бы одна радикальная цепочка вида $(*)$, в которой P_s содержит какой-то корень этого уравнения.

Если уравнение разрешимо в радикалах, то оно, конечно, имеет *корень, выражающийся в радикалах*. А именно таковым является любой из его корней. Оказывается, что при некотором условии на исходное уравнение выполняется и обратное утверждение (см. следствие к теореме 6 в п. 7).

Для уравнения $x^2+2x+\frac{1}{2}=0$, с которого мы начинали этот пункт, радикальная цепочка полей имеет вид:

$$\boldsymbol{Q} \subset \boldsymbol{Q}(\sqrt{2}),\ s=1.$$

Для любого квадратного уравнения $x^2+bx+c=0$ с конкретными фиксированными коэффициентами b и c радикальная цепочка полей имеет вид:

$$\boldsymbol{Q}(b,\ c) \subset \boldsymbol{Q}(b,\ c,\ \sqrt{b^2-4c}),\ s=1.$$

Мы использовали утверждение следующего упражнения.

Упражнение 4. *Докажите, что для любого числового поля P и любых чисел $a_1, \ldots, a_{m+1}$ выполняется*

$$P(a_1, \ldots, a_m)(a_{m+1}) = P(a_1, \ldots, a_{m+1}).$$

Для любого кубического уравнения $x^3+ax^2+bx+c=0$ с фиксированными коэффициентами a, b, c также выпишем радикальную цепочку полей:

$$\boldsymbol{Q}(a,\ b,\ c) \subset \boldsymbol{Q}\left(a,\ b,\ c,\ \sqrt{\left(\frac{q}{2}\right)^2+\left(\frac{p}{3}\right)^3}\right) \subset$$

$$\subset \boldsymbol{Q}\left(a,\ b,\ c,\ d,\ \sqrt[3]{-\frac{q}{2}+d}\right) \subset P_3 \neq \boldsymbol{Q}\left(a,\ b,\ c,\ d,\ \sqrt[3]{-\frac{q}{2} \pm d}\right),$$

$$p=-\frac{a^2}{3}+b,\ q=\frac{2a^3}{27}-\frac{ab}{3}+c \text{ и } d=\sqrt{\left(\frac{q}{2}\right)^2+\left(\frac{p}{3}\right)^3}.$$

Ясно, что эта радикальная цепочка выписана на основе формулы Кардано: $x_1=\sqrt[3]{-\frac{q}{2}+d}+\sqrt[3]{-\frac{q}{2}-d}-\frac{a}{3}$, которая дает один из корней x_1 исходного уравнения.

Над полем P_3 исходное уравнение нацело делится на многочлен $(x-x_1)$ и в результате получается квадратное уравнение, коэффициенты которого содержатся в поле P_3. Делая еще один, 4-й шаг, образуем расширение P_4, содержащее квадратный корень из дискриминанта этого квадратного уравнения. Таким образом получим радикальную цепочку полей, означающую, что исходное кубическое уравнение разрешимо в радикалах.

Аналогично для любого уравнения четвертой степени $x^4+ax^3+bx^2+cx+d=0$, используя формулу Феррари, легко выписать радикальное расширение поля коэффициентов, содержащее два корня x_1 и x_2 исходного уравнения. Затем, поделив исходное уравнение на $(x-x_1)\cdot(x-x_2)$ над этим радикальным расширением, получим квадратное уравнение и продолжим радикальную цепочку до такого радикального расширения, которое содержит все корни исходного уравнения.

Действительно, образуем радикальное расширение P_s, содержащее по крайней мере один корень y_1 следующего кубического уравнения («резольвенты Феррари»):

$$y^3-b\cdot y^2+(ac-4d)\cdot y+(4bd-a^2d-c^2)=0.$$

Соответствующий ему радикальный ряд обозначим:

$$P_0 \subset P_1 \subset \ldots \subset P_s.$$

Продолжим этот ряд

$$P_0 \subset P_1 \subset \ldots \subset P_s \subset P_s(\sqrt{D}),$$

где D — дискриминант следующего квадратного (!) уравнения:

$$x^2+\frac{a}{2}\cdot x+\frac{y_1}{2}=$$

$$=\sqrt{\left(\frac{a^2}{4}-b+y_1\right)\cdot x^2+\left(\frac{a}{2}\cdot y_1-c\right)\cdot x+\frac{y_1^2}{4}-d}.$$

(Подкоренное выражение в правой части является полным квадратом.) Мы получили искомое радикальное расширение, так как $y_1 \in P_s$ и два корня x_1 и x_2 исходного уравнения 4-й степени являются корнями этого квадратного уравнения и, следовательно, принадлежат $P_s(\sqrt{D})$.

Итак, любое (с конкретными числовыми коэффициентами) уравнение степени, меньшей или равной 4, разрешимо в радикалах. Однако если мы напишем уравнение 5-й степени, например уравнение

$$x^5+kx+k=0,$$

где k — фиксированное натуральное число, то совершенно не вид-

но, разрешимо ли это уравнение в радикалах и от чего зависит наличие или отсутствие такой разрешимости.

Более того, для уравнений степени $\leqslant 4$ и даже для большинства уравнений 3-й степени совсем не ясно, чем объясняется их разрешимость в радикалах; да и формулы Кардано и Феррари, позволившие выше построить соответствующие радикальные расширения, выводятся довольно искусственными преобразованиями.

Ответы на эти вопросы даются ниже, в теореме Э. Галуа (п. 5). В § 3 на ее основе будут получены формулы для решения в радикалах уравнений степени, меньшей или равной 4. Основным понятием в теореме Э. Галуа является понятие разрешимой группы, к которому мы вынуждены сейчас перейти.

2. Понятие разрешимой группы.

Определение 4. **а) Разрешимым рядом конечной группы G называется последовательность (цепочка) групп вида**

$$G \rightleftarrows G_0 \supset G_1 \supset G_2 \supset \ldots \supset G_s \rightleftarrows \{e\},$$

в которой любая группа G_i является нормальным делителем предыдущей группы G_{i-1} (где $i=1, \ldots, s$) и все фактор-группы G_{i-1}/G_i являются циклическими группами.

б) Конечная группа G, обладающая хотя бы одним разрешимым рядом, называется разрешимой группой. Конечная группа, не обладающая ни одним таким рядом, называется неразрешимой группой.

Для бесконечных групп также определяется понятие разрешимой группы; мы не будем его рассматривать.

Напомним, что нормальным делителем группы G называется любая ее подгруппа H со свойством $\forall g \in G (gHg^{-1} \subseteq H)$. Например, любая подгруппа абелевой группы является ее нормальным делителем; сама группа G и ее травиальная подгруппа $\{e\}$ всегда являются нормальными делителями исходной группы G.

Упражнение 5. *Приведите примеры нормальных делителей группы* GL_2 *и примеры ее подгрупп, не являющихся ее нормальными делителями.*

Напомним, что циклической называется группа, все элементы которой являются степенями какого-то одного ее элемента. Этот элемент называется образующим группы.

Например $\boldsymbol{Z}^+$ есть циклическая группа с образующей 1 и F_n^+ — множество вычетов по модулю n с операцией сложения вычетов — циклическая группа с образующей 1. В частности, $F_5 = \{0, 1, 2, 3, 4\}$ и $2 = 1+1$, $3 = 1+1+1$ и т. д.

Упражнение 6. *Докажите, что:*

а) всякая циклическая группа абелева;

б) любая подгруппа и любая фактор-группа циклической группы сами являются циклическими группами.

Важным примером циклической группы является множество всех корней m-й степени из числа 1 с операцией умножения их как комплексных чисел. Как мы видели в п. 1, все корни уравнения $x^m = 1$,

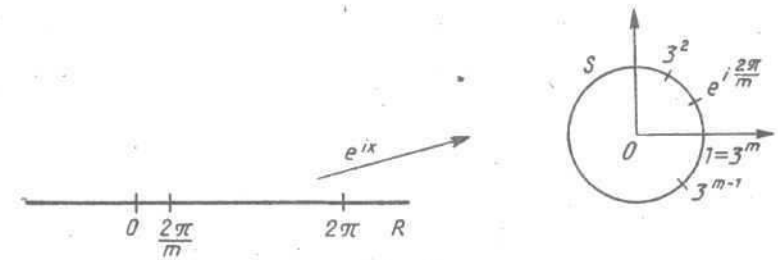


Рис. 92

т. е. числа вида $\sqrt[m]{1}$, образуют множество $\{\zeta, \zeta^2, \ldots, \zeta^{m-1}, 1\}$, где $\zeta = e^{i\frac{2\pi}{m}}$ (рис. 92). Таким образом, число ζ — образующий элемент этой группы. Легко проверить, что эта группа изоморфна F_m^+.

Конечной называется группа, содержащая конечное число элементов. Число элементов в ней иногда называется порядком группы.

Примером разрешимой группы является любая конечная циклическая группа G, так как в этом случае разрешимый ряд имеет вид $G \supset \{e\}$.

Простой называется группа G, не имеющая нормальных делителей кроме $\{e\}$ и G. Например, легко доказать, что F_p^+ — простая группа для любого простого числа p.

Важным примером неразрешимой группы является любая конечная, простая и не циклическая группа. Действительно, для такой группы G любой ее разрешимый ряд может иметь только вид $G \supset \{e\}$.

Но $G/\{e\} \cong G$; поэтому G — циклическая группа, что противоречит исходным условиям на группу, т. е. она не может иметь никакого разрешимого ряда. В следующем пункте приводится, к сожалению, не простой пример группы с таким набором свойств.

3. Определение симметрической и знакопеременной групп. Пусть фиксировано натуральное число $n \geqslant 2$ и множество X состоит из n элементов. Например, можно считать, что $X \rightleftarrows \{1, 2, \ldots, n\}$. Обозначим S_n семейство всех биекций множества X на себя. Раньше мы неоднократно встречались с семейством всех биекций произвольного (не обязательно конечного) множества Y на себя и обозначали это семейство $\mathrm{Is}(Y)$. Поэтому $S_n = \mathrm{Is}(X)$. В семействе биекций $\mathrm{Is}(Y)$ и, в частности, в S_n рассматривается операция композиции двух биекций. Относительно этой операции семейство (множество) $\mathrm{Is}(Y)$ и, в частности, S_n является группой. В ней нейтральным элементом является тождественная биекция id и симметричным элементом к биекции f является обратная биекция f^{-1}.

Если множества X_1 и X_2 состоят из одинакового числа элементов, то группы $\mathrm{Is}(X_1)$ и $\mathrm{Is}(X_2)$ изоморфны. Поскольку изоморфные группы обычно не различаются, то природа элементов множества X не играет никакой роли. Любую биекцию σ из S_n записывают в

виде $\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$ или $\begin{pmatrix} 1 & 2 & \dots & n \\ k_1 & k_2 & \dots & k_n \end{pmatrix}$, где $k_n = \sigma(n)$. Ее называют перестановкой (подстановкой). Ясно, что группа S_n имеет порядок $n!$

Группа S_n, как и матричная группа и произвольная группа вида $\mathrm{Is}(X)$, действует на множестве X, которое для определенности можно представить себе как множество, состоящее из первых n натуральных чисел, т. е. $X = \{1, \dots, n\}$. А именно, определим отображение $F(\cdot) : \sigma \mapsto f_\sigma$, где функция $f_\sigma(i) \rightleftarrows \sigma(i)$. Иными словами, биекция σ как элемент группы S_n переводит точку i в точку $\sigma(i)$. Ясно, что F — эффективное и транзитивное действие группы S_n на множестве $X = \{1, \dots, n\}$. Напомним, что действие F называется эффективным, если функция f инъективна; действие F называется транзитивным, если выполняется $\forall i,\ j \in X \exists \sigma \in S_n (f_\sigma(i) = j)$. Если G произвольная подгруппа S_n, то ограничение действия F на G является действием G на том же множестве $X = \{1, \dots, n\}$, т. е. элементу σ из G по-прежнему соответствует функция f_σ, где $f_\sigma(i) \rightleftarrows \sigma(i)$. Обычно объекты σ и f_σ отождествляются и обозначаются одинаковой буквой σ. Произвольная подгруппа G группы S_n действует эффективно, но не обязательно транзитивно. Например, подгруппа $G \rightleftarrows \{\mathrm{id}\}$ не может перевести число 1 в число 2. Подгруппа G называется транзитивной, если ее действие транзитивно. Иными словами, подгруппа G группы S_n называется транзитивной, если

$$\forall i,\ j \in X \exists \sigma \in G (\sigma(i) = j).$$

В множестве $\{+1, -1\}$ в качестве групповой операции определяется обычное умножение чисел $+1$ и -1. Полученная группа изоморфна группе F_2.

Перестановка σ, оставляющая на месте все элементы множества $\{1, \dots, n\}$, кроме каких-то двух его элементов, называется транспозицией. Иными словами, перестановка σ — транспозиция, если для фиксированных i и j выполняется $\sigma(i) = j$, $\sigma(j) = i$ и $\forall k \neq i,\ j (\sigma(k) = k)$, где $i \neq j$. Например, перестановка $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ транспозиция, а перестановки $\mathrm{id} \rightleftarrows \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ и $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ не транспозиции. Транспозицию, переставляющую элементы i и j, обозначим σ_{ij}. Порядок индексов i и j, очевидно, не важен.

Если i, j, k, l различные числа, то

$$\sigma_{ij} \cdot \sigma_{kl} = \sigma_{kl} \cdot \sigma_{ij},$$

но в общем случае это равенство не выполняется. Например, $\sigma_{ij} \circ \sigma_{jk}(i) = k$ и $\sigma_{jk} \circ \sigma_{ij}(i) = j$ и j может не равняться k. Напомним, что запись $f \cdot g$ по определению означает то же самое, что и $g \circ f \rightleftarrows g(f(\cdot))$.

Обозначим $\nu(\sigma)$ число пар i, j в верхней строке перестановки σ, для которых $i < j$ и $\sigma(i) > \sigma(j)$.

Теорема 4. ***а) Множество всех транспозиций порождает группу S_n, т. е. любой элемент σ из S_n представим в виде***

$$\sigma = \sigma_{i_1 i_2} \cdot \dots \cdot \sigma_{i_l i_{l+1}}. \qquad (*)$$

б) Функция $\varepsilon(\sigma) \rightleftarrows (-1)^{\nu(\sigma)}$ является гомоморфизмом групп $\varepsilon(\cdot) : S_n \to \{-1;\ +1\}$.

в) Для любой транспозиции σ_{ij} выполняется $\varepsilon(\sigma_{ij}) = -1$.

▷ а) Докажем это утверждение индукцией по n. Если $n = 2$, то любое σ из S_2 вида $\begin{pmatrix} 1 & 2 \\ k_1 & k_2 \end{pmatrix}$ и, следовательно, доказываемое утверждение верно.

Если для всех σ из S_n это утверждение уже доказано, то рассмотрим $\sigma \in S_{n+1}$. Разберем два случая: σ содержит неподвижную точку и σ произвольная перестановка. В первом случае пусть для конкретности $\sigma(1) = 1$. Рассмотрим перестановку σ', состоящую из всех столбцов σ, кроме первого. Такое $\sigma' \in S_n$. Точнее, в верхней строке σ' находятся числа $2, \dots, n+1$, и мы заменяем их соответственно числами $1, \dots, n$. Для такого σ' применяем индуктивное предположение. Получаем $\sigma' = \pi_{i_1 i_2} \cdot \dots \cdot \pi_{i_l i_{l+1}}$ и затем заменяем все элементы всех этих перестановок на соответствующие прежние числа $2, \dots, n+1$. Таким образом, $\forall i (2 \leqslant i \leqslant n+1 \Rightarrow \sigma'(i) = (\pi_{i_1 i_2} \cdot \dots \cdot \pi_{i_l i_{l+1}})(i)$. Добавим ко всем транспозициям $\pi_{i_1 i_2}, \dots, \pi_{i_l i_{l+1}}$ столбец $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Получаются снова транспозиции, которые мы обозначим $\pi'_{i_1 i_2}, \dots, \pi'_{i_l i_{l+1}}$. Тогда $\sigma = \pi'_{i_1 i_2} \cdot \dots \cdot \pi'_{i_l i_{l+1}}$.

Пусть теперь σ произвольная перестановка и в ней число 1 переходит в какое-то число k_1. Умножим σ на транспозицию σ_{1k_1}, тогда транспозиция $\sigma \cdot \sigma_{1, k_1}$ содержит неподвижную точку 1. Применим к ней уже рассмотренный случай, т. е. представим ее в виде $\sigma \cdot \sigma_{1, k_1} = \sigma_1 \cdot \dots \cdot \sigma_l$, где $\sigma_1, \dots, \sigma_l$ — транспозиции. Тогда $\sigma = \sigma_1 \cdot \dots \cdot \sigma_l \cdot \sigma_{1, k_1}^{-1}$ и $\sigma_{1, k_1}^{-1} = \sigma_{1, k_1}$ — транспозиция.

б) Пары чисел i, j из верхней строки перестановки σ, для которых $i < j$ и $\sigma(i) > \sigma(j)$, называются инверсиями перестановки σ. Поэтому число $\nu(\sigma)$ равно числу инверсий в перестановке σ.

Рассмотрим произведение $P \rightleftarrows \prod\limits_{1 \leqslant i < j \leqslant n} (j - i)$. Положим $\sigma(P) \rightleftarrows \rightleftarrows \prod\limits_{i<j} (\sigma(j) - \sigma(i))$. Легко проверить соотношение

$$\sigma(P) = \varepsilon(\sigma) \cdot P.$$

Если τ — любая перестановка, то аналогично проверяется

$$(\tau \cdot \sigma) \cdot P = \prod_{i<j} \sigma(\tau(j)) - \sigma(\tau(i)) = \varepsilon(\sigma) \cdot \prod_{i<j} (\tau(j) - \tau(i)).$$

Поэтому

$$\varepsilon(\tau \cdot \sigma) \cdot P = (\tau \cdot \sigma)(P) = \varepsilon(\sigma) \cdot \prod_{i<j} (\tau(j) - \tau(i)) =$$

$$= \varepsilon(\sigma) \cdot \tau(P) = \varepsilon(\sigma) \cdot \varepsilon(\tau) \cdot P.$$

Сокращая на число P и переставляя сомножители в правой части этого равенства, получим $\varepsilon(\tau\cdot\sigma)=\varepsilon(\tau)\cdot\varepsilon(\sigma)$.

Далее, $\varepsilon(\mathrm{id})\cdot P=\mathrm{id}\,(P)=P=1\cdot P$ и $\varepsilon(\mathrm{id})=1$. Отсюда, как и для любого гомоморфизма, вытекает:

$$\varepsilon(\sigma^{-1})=(\varepsilon(\sigma))^{-1}.$$

Итак, функция $\varepsilon(\cdot)$ — гомоморфизм.

в) Рассмотрим транспозицию $\sigma_{i_1 i_2}$, где можно всегда считать, что $i_1<i_2$. Тогда $\varepsilon(\sigma_{i_1 i_2})\cdot P=\sigma_{i_1 i_2}(P)=-P$, так как произведение $\sigma_{i_1 i_2}(P)$ ровно в одном сомножителе отличается от произведения P и эти сомножители отличаются знаком $-$. Поэтому $\varepsilon(\sigma_{i_1 i_2})=-1$. ■

Следствие. а) Перестановка id не представима в виде произведения нечетного числа транспозиций.

б) Четность или нечетность числа транспозиций в представлении $(*)$ не зависят от выбора такого представления.

▷ а) Если $\mathrm{id}=\sigma_1\circ\ldots\circ\sigma_{2l+1}$, то $\varepsilon(\mathrm{id})=\varepsilon(\sigma_1)\ldots\varepsilon(\sigma_{2l+1})$, что приводит к противоречию.

б) Если $\sigma=\sigma_1\circ\ldots\circ\sigma_l$ и $\sigma=\sigma'_1\circ\ldots\circ\sigma'_{l'}$, то $\varepsilon(\sigma)=\varepsilon(\sigma_1)\ldots\varepsilon(\sigma_l)$ и l четно, если $\varepsilon(\sigma)=+1$. С другой стороны, $\varepsilon(\sigma)=\varepsilon(\sigma'_1)\ldots\varepsilon(\sigma'_{l'})$ и l' четно, если $\varepsilon(\sigma)=+1$. Поэтому l и l' оба четны и оба нечетны. ■

Множество $\{\sigma\in S_n\mid\varepsilon(\sigma)=+1\}$ с операциями, индуцированными в него из группы S_n, называется знакопеременной группой степени n и обозначается A_n.

Так как $A_n=\mathrm{Ker}\,\varepsilon$, то группа A_n есть подгруппа S_n. Иными словами, A_n состоит из всех перестановок с четным числом инверсий. Такие перестановки называются четными.

Упражнение 7. *Докажите, что ядро любого гомоморфизма является нормальным делителем области определения этого гомоморфизма. В частности, группа A_n есть нормальный делитель группы S_n. Если $f: G_1\to G_2$ гомоморфизм групп, то группы $G_1/\mathrm{Ker}\,f$ и $R(f)$ изоморфны.*

4. Разрешимость симметрической и знакопеременной групп. Рассмотрим вопрос о том, при каких значениях $n\in\boldsymbol{N}_{\geqslant 2}$ группы S_n и A_n являются простой и разрешимой группой.

Для группы S_n вопрос о разрешимости легко сводится к такому же вопросу для группы A_n. А именно, если группа A_n неразрешима, то по следующему упражнению группа S_n также неразрешима.

Упражнение 8. *Докажите, что любая подгруппа разрешимой группы разрешима. Если упражнение покажется трудным, то можно обратиться к лемме 12 из приложения 3.*

Если группа A_n разрешима, то и группа S_n разрешима, т. к. разрешимый ряд для A_n можно продолжить до разрешимого ряда для S_n. Действительно, если разрешимый ряд для A_n имеет вид

$$A_n\supset G_1\supset\ldots\supset\{e\},$$

то разрешимый ряд для S_n имеет вид

$$S_n\supset A_n\supset\ldots\supset G_n\supset\ldots\supset\{e\},$$

т. к. A_n нормальный делитель S_n (упражнение 8) и S_n/A_n состоит из двух элементов. (По упражнению 7 получаем $S_n/A_n=S_n/\mathrm{Ker}\,\varepsilon(\cdot)\cong R(\varepsilon)=\{-1;\ +1\}$, т. е. группа S_n/A_n изоморфна циклической группе $\{-1;\ +1\}$, и, следовательно, сама является циклической группой.)

Итак, решим поставленный выше вопрос для группы A_n. При $n=2$ группа A_2 состоит из одного элемента id, так как группа $S_2=\left\{\begin{pmatrix}1&2\\1&2\end{pmatrix},\begin{pmatrix}1&2\\2&1\end{pmatrix}\right\}$ и вторая перестановка нечетная, т. е. A_2 простая и разрешимая группа.

При $n=3$ группа A_3 состоит из трех элементов, которые легко найти перебором всех элементов группы S_3. Впрочем, легко доказать, что A_n состоит из $\frac{1}{2}n!$ элементов. Поэтому в A_3 три элемента. Итак, выпишем все элементы групп S_3 и A_3:

$$S_3=\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\right\},$$

$$A_3=\left\{\mathrm{id},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\right\}.$$

Очевидно, что A_3 циклическая группа с образующим элементом, например, $\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$. Поэтому она, как и любая циклическая группа, имеет тривиальный разрешимый ряд

$$A_3\supset\{e\},$$

т. е. A_3 простая и разрешимая группа.

При $n=4$ группа A_4 состоит из 12 элементов, которые нетрудно выписать:

$$\begin{array}{llll} e, t_1=(1\,2)(3\,4), & t_2=(1\,3)(2\,4), & t_3=(1\,4)(2\,3), & \\ s_1=(1\,2\,3), & s_2=(1\,2\,4), & s_3=(1\,3\,2), & s_4=(1\,3\,4), \\ s_5=(1\,4\,5), & s_6=(1\,4\,3), & s_7=(2\,3\,4), & s_8=(2\,4\,3), \end{array}$$

где используется удобный способ записи перестановок: запись $(ij)(kl)$ обозначает, что элементы i и j переходят друг в друга, а элементы k и l переходят друг в друга, т. е. обозначает перестановку $\begin{pmatrix}i&j&k&l\\j&i&l&k\end{pmatrix}$. Запись $(i\,j\,k)$ обозначает, что i переходит в j, j в k, k в i, а все оставшиеся элементы переходят в себя, т. е. обозначает перестановку $\begin{pmatrix}i&j&k&l\\j&k&i&l\end{pmatrix}$. Нейтральный элемент любой группы обычно обозначают e. В нашем случае $e=\mathrm{id}$.

Множество $\{\mathrm{id}, t_1, t_2, t_3\}$ образует, как легко проверить, непосредственными подсчетами, подгруппу в A_4, которая является абелевой и имеет порядок 4. Она называется клейновской группой и обозначается B_4.

Прямым подсчетом проверяется, что B_4 нормальный делитель в A_4.

Упражнение 9. *Проверьте сформулированные выше утверждения относительно группы A_4.*

Поскольку фактор-группа A_4/B_4 состоит из трех элементов, то из пункта а) следующей леммы вытекает ее цикличность.

Лемма 7. ***а) Любая конечная группа простого порядка циклическая.***

б) Любая конечная, циклическая и простая группа имеет простой порядок.

▷ а) В любой конечной группе G, порядок которой является простым числом p, всякий ее элемент g, отличный от нейтрального элемента l, имеет порядок $n(g)$, равный порядку p всей группы. Действительно, всевозможные степени элемента g образуют подгруппу в G. Эта подгруппа состоит из $n(g)$ элементов.

По теореме Лагранжа число $n(g)$ делит число p (в кольце $\boldsymbol{Z}$), т. е. $n(g)=1$ или $n(g)=p$. Если $n(g)=1$, то $g=e$, что исключается выбором элемента g. Поэтому $n(g)=p$.

Итак, выберем в G любой не нейтральный элемент g. Тогда $n(g)=p$ и множество $\{g^1, g^2, \ldots, g^{n(g)}\}$, содержащееся в G, состоит из p попарно различных элементов. Но в G всего p элементов. Поэтому это множество совпадает с G и G циклическая группа с образующей g.

б) Пусть g — образующая группы G и порядок группы G равен n. Допустим, что число m делит число n (в кольце $\boldsymbol{Z}$). Тогда множество всех степеней элемента $g^{\frac{n}{m}}$ образует подгруппу в G, состоящую ровно из m элементов (почему?). Так как G простая группа, то $m=1$ или $m=n$. ■

Поскольку B_4 абелева группа, то все ее подгруппы являются нормальными делителями группы B_4 (но не обязательно всей группы A_4). Например, образуем ее подгруппу $C_4 \rightleftarrows \{e, t_1\}$. Она является циклической группой, так как ее порядок — простое число 2.

Порядок фактор-группы B_4/C_4 также равен 2, и поэтому она циклическая группа. Итак, найден разрешимый ряд для группы A_4:

$$A_4 \supset B_4 \supset C_4 \supset \{e\}.$$

Следовательно, A_4 разрешимая, но не простая группа.

Теорема 5. ***а) Все группы A_n и S_n при $2 \leqslant n \leqslant 4$ разрешимы.***

б) Группы A_5 и S_5 неразрешимы.

▷ Пункт а) уже доказан выше. Осталось проверить первое утверждение пункта б), так как из него, как мы видели, вытекает второе утверждение.

Для этого докажем, что произвольный нормальный делитель G_0 в группе A_5 тривиален, т. е. $G_0=\{e\}$ или $G_0=A_5$. Это означает, что группа A_5 проста. Если для группы A_5 существует разрешимый ряд, то он имеет вид: $A_5 \supset \{e\}$. Тогда группа $A_5=A_5|\{e\}$ циклична и одновременно проста, ее порядок равен $\frac{1}{2}5!=60$. Такое сочетание свойств группы невозможно по лемме 7б.

Итак, вся оставшаяся часть доказательства теоремы 5 посвящена проверке утверждения: если G_0 нормальный делитель группы A_5 и $G_0 \neq \{e\}$, то $G_0=A_5$. При первом чтении эту проверку можно опустить.

Лемма 8. ***Нормальный делитель G_0 группы A_5, $G_0 \neq \{e\}$, содержит перестановку***

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ k_1 & k_2 & k_3 & k_4 & k_5 \end{pmatrix},$$

для которой $\sigma(i)=i$ при некотором i, $1 \leqslant i \leqslant 5$.

▷ Пусть $\sigma \in G_0$ и $\sigma \neq e$; σ — четная перестановка вида

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ k_1 & k_2 & k_3 & k_4 & k_5 \end{pmatrix}.$$

Допустим, что в σ нет неподвижной точки, т. е. $\forall i(\sigma(i) \neq i)$. Тогда в σ элемент 1 переходит в какой-то элемент $k_i \neq 1$; если k_i переходит в 1, то оставшиеся три элемента i_1, i_2, i_3 должны переходить по закону $(i_1\ i_2\ i_3)$, так как иначе в σ будет неподвижная точка. В этом случае (назовем его первым) σ вида $\sigma = \begin{pmatrix} 1 & k_i & i_1 & i_2 & i_3 \\ k_i & 1 & i_2 & i_3 & i_1 \end{pmatrix}$.

Пусть k_i переходит в $k_j \neq 1, k_i$; если k_j переходит в k_i или k_j переходит в 1, то мы приходим к первому случаю (проверьте). Значит, k_j переходит в один из двух оставшихся элементов i_1 и i_2, например в i_1. Если i_1 переходит в 1, то i_2 неподвижная точка. Поэтому получаем второй случай:

$$\sigma = \begin{pmatrix} 1 & k_i & k_j & i_1 & i_2 \\ k_i & k_j & i_1 & i_2 & 1 \end{pmatrix}.$$

Перестановка $\sigma'\sigma(\sigma')^{-1}$ принадлежит G_0, т. к. G_0 нормальный делитель, поэтому $\sigma_1 \rightleftarrows \sigma'\sigma(\sigma')^{-1}\sigma^{-1} \in G_0$. В первом случае положим $\sigma' \rightleftarrows (i_2 i_3 k_i)$ и, подсчитав, получим, что перестановка σ_1 относится ко второму случаю. Во втором случае положим $\sigma' \rightleftarrows (k_i k_j i_1)$ и, подсчитав, получим перестановку $\sigma_1 \rightleftarrows (1\ k_j\ i_1)$, которая имеет неподвижные точки. ■

Если в перестановке σ имеется неподвижная точка, то, переобозначая элементы множества X, на котором действуют перестановки, можно считать, что неподвижно число 5. Столбец $\begin{pmatrix} 5 \\ 5 \end{pmatrix}$ в такой перестановке будем опускать.

Пусть G_0 нормальный делитель A_5 и $G_0 \neq \{e\}$. Выберем в G_0 какую-то четную перестановку $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ k_1 & k_2 & k_3 & k_4 \end{pmatrix}$, $\sigma \neq e$, четвертого порядка и рассмотрим вопрос о том, какой вид она может иметь.

Если относительно перестановки σ никакое число i не переходит ни в какое другое число j, т. е. $\forall i(\sigma(i)=i)$, то $\sigma=e$. Это невозможно. Поэтому обозначим i число, которое σ переводит в какое-то другое число j, где $i \neq j$. Остальные два числа в верхней строке этой перестановки σ обозначим k и l.

1) Чему равняется значение перестановки σ на числе j? Иными словами, в какое число переводит перестановка σ число j? Назовем первым тот случай, когда $\sigma(j)=i$. Тогда σ — перестановка вида $\begin{pmatrix} i & j & k & l \\ j & i & . & . \end{pmatrix}$. Число k, не равное числам i и j, не может перейти ни в i, ни в j. Следовательно, число k переходит в k (и тогда число l переходит в число l) или число k переходит в число l (тогда число l переходит в k). При первой возможности перестановка σ — транспозиция, что противоречит ее четности. Остается только вторая возможность.

Итак, в первом случае перестановка σ имеет вид:

$$\sigma=\begin{pmatrix} i & j & k & l \\ j & i & l & k \end{pmatrix}.$$

2) Если первый случай не имеет места, то число j переходит в какое-то число k, где $k\neq i$, j. При этом само k переходит или в число i, или в число l. Назовем вторым случаем тот, когда k переходит в i. Тогда l может перейти только в себя. Следовательно, во втором случае перестановка σ имеет вид:

$$\sigma=\begin{pmatrix} i & j & k & l \\ j & k & i & l \end{pmatrix}.$$

3) Третьим случаем назовем тот, когда число k переходит в число l и, следовательно, l может перейти только в i. Итак, в третьем случае перестановка σ имеет вид:

$$\sigma=\begin{pmatrix} i & j & k & l \\ j & k & l & i \end{pmatrix}.$$

Поскольку G_0 — нормальный делитель группы G, то для любой перестановки σ' перестановка вида $\sigma'\sigma(\sigma')^{-1}$ принадлежит G_0. Следовательно, и перестановка вида $\sigma_1\rightleftarrows(\sigma'\sigma(\sigma')^{-1})\cdot\sigma^{-1}$ принадлежит G_0.

В первом случае положим $\sigma'=(j\,k\,l)$, во втором и третьем случаях положим $\sigma'=(i)\,(jkl)$.

В каждом из этих случаев подсчитаем σ_1:

1) $\sigma_1=(j\,k\,l)\,((i\,j)\,(k\,l))\,(l\,k\,j)\,((i\,j)\,(k\,l))=\begin{pmatrix} i & j & k & l \\ l & k & j & i \end{pmatrix}=(i\,l)\,(j\,k)$;

2) $\sigma_1=(i)\,(j\,k\,l)\,((i\,j\,k)\,(l))\,((i)\,(l\,k\,j))\,((k\,j\,i)\,(l))=\begin{pmatrix} i & j & k & l \\ l & k & j & i \end{pmatrix}=(i\,l)\,(j\,k)$;

3) $\sigma_1=(i)\,(j\,k\,l)\,((i\,j\,k\,l))\,((i)\,(l\,k\,j))\,((l\,k\,j\,i))=\begin{pmatrix} i & j & k & l \\ j & k & i & l \end{pmatrix}=(i\,j\,k)\,(l)$.

В последнем случае поступим с σ_1, как мы поступили с σ во втором случае, и по п. 2 получим перестановку $(i\,l)\,(j\,k)$.

Итак, во всех случаях в G_0 содержится перестановка вида $(i\,l)\,(j\,k)$.

Докажем, что G_0 содержит любую перестановку вида $(k_1k_2)\times(k_3k_4)$, где все числа k_1, k_2, k_3, k_4 попарно различны. Для этого

образуем перестановку $\sigma_2=\begin{pmatrix} k_1 & k_2 & k_3 & k_4 \\ i & l & j & k \end{pmatrix}$ и подсчитаем перестановку вида $\sigma_2((i\,l)\,(j\,k))\,\sigma_2^{-1}=\begin{pmatrix} k_1 & k_2 & k_3 & k_4 \\ k_2 & k_1 & k_4 & k_3 \end{pmatrix}=(k_1k_2)\,(k_3k_4)$. Кроме того, обозначим $\sigma_0\rightleftarrows\sigma_2\cdot(i\,l)$ и подсчитаем:

$$\sigma_0((i\;\;l)\,(j\;\;k))\,\sigma_0^{-1}=\sigma_2((i\,l)\,((i\,l)\,(j\,k))\,(i\,l))\,\sigma_2^{-1}=\sigma_2\begin{pmatrix} i & j & k & l \\ l & k & j & i \end{pmatrix}\sigma_2^{-1}=$$

$$=\sigma_2(i\,l)\,(j\,k)\,\sigma_2^{-1}=(k_1k_2)\,(k_3k_4).$$

Перестановки σ_0 и σ_2 отличаются на транспозицию $(i\,l)$. Поэтому одна из них четная, а другая нечетная. Четную из них обозначим σ_3. Поэтому $\sigma_3\in A_5$ и $(i\,l)\,(j\,k)\in G_0$ и

$$(k_1k_2)\,(k_3k_4)=\sigma_3(i\;\;l)\,(j\;\;k)\sigma_3^{-1}.$$

Поскольку G_0 — нормальный делитель, то $(k_1k_2)\,(k_3k_4)\in G_0$.

Наконец, рассмотрим произвольную четную перестановку σ из группы A_5. По теореме 4а она разлагается в произведение четного числа транспозиций. Разобьем это произведение на соседние пары транспозиций. Если каждая такая пара содержится в G_0, то и произведение всех пар содержится в G_0.

Пусть $(k_1k_2)\,(k_3k_4)$ — одна из таких пар. В ней числа k_1 и k_2, k_3 и k_4 по определению транспозиции различны. Если множества $\{k_1, k_2\}$ и $\{k_3, k_4\}$ равны, то эти две транспозиции совпадают, а в представлении $(*)$ из теоремы 4, очевидно, отсутствуют одинаковые соседние транспозиции. Если эти два множества дизъюнктны, то все числа k_1, k_2, k_3, k_4 попарно не равны. Уже доказано, что любое такое произведение двух транспозиций содержится в G_0.

Итак, осталось рассмотреть случай, когда эти два множества имеют один общий элемент. Например, пусть $k_1=k_3$. Тогда получится перестановка $(k_1k_2)\cdot(k_1k_4)$. В множестве из пяти элементов (ведь $n=5$), кроме трех различных чисел k_1, k_2, k_4, найдутся еще два различных числа l_1 и l_2, отличных от этих трех чисел. По доказанному перестановки $(k_1k_2)\cdot(l_1l_2)$ и $(l_1l_2)\cdot(k_1k_4)$ принадлежат G_0 и $(k_1k_2)\,(k_1k_4)=(k_1k_2)\,(l_1l_2)\cdot(l_1l_2)\,(k_1k_4)$. Поэтому $(k_1k_2)\,(k_1k_4)\in G_0$. Итак, доказано, что $G_0=A_5$. ■

Следствие. Группа A_5 является простой группой, а группа S_5 не является простой группой.

Замечание. По существу повторяя доказательство теоремы 5, б, легко проверить, что для $n\geqslant 5$ все группы вида A_n суть простые и неразрешимые группы, а все группы вида S_n — не простые и неразрешимые группы.

5. Понятие группы Галуа. Формулировка теоремы Галуа. Пусть фиксирован многочлен $f(x)=x^n+a_1x^{n-1}+\ldots+a_n$ и все его коэффициенты 1, a_1, ..., a_n принадлежат некоторому числовому полю P. Тогда $\boldsymbol{Q\subseteq P\subseteq C}$.

Например, P — поле коэффициентов уравнения $f(x)=0$, т. е. $P\rightleftarrows Q(a_1, \ldots, a_n)$.

Обозначим $x_1, \ldots, x_n$ все (вещественные и комплексные) корни уравнения $f(x)=0$.

Как хорошо известно, над полем $K=P(x_1, \ldots, x_n)$ многочлен f раскладывается в произведение многочленов первой степени, т. е в произведение линейных многочленов:

$$\forall x \in K(f(x)=(x-x_1)\ldots(x-x_n)).$$

Наоборот, если над некоторым полем K, где $P \subseteq K$, многочлен f раскладывается в произведение линейных многочленов, то, очевидно, K содержит все корни уравнения $f(x)=0$.

Определение 5. **а) Полем разложения многочлена $f(\cdot)$ из кольца всех многочленов $P[x]$ называется поле вида $P(x_1, \ldots, x_n)$, где $x_1, \ldots, x_n$ — все корни уравнения $f(x)=0$. Поле разложения такого многочлена f будем обозначать P_f.**

б) Пусть P и K — два произвольных поля, для которых $P \subseteq K$. Группой Галуа поля K над полем P называется множество всех автоморфизмов поля K, т. е. множество всех изоморфизмов вида $W: K \leftrightarrow K$, для которых любая точка из P неподвижна:

$$\forall k \in P(W(k)=k). \qquad (*)$$

Как всегда, в множестве таких автоморфизмов в качестве групповой операции рассматривается операция композиции двух автоморфизмов.

Условие $(*)$ можно записать в эквивалентном виде $W \restriction P = \mathrm{id}_P$. Группа Галуа поля K над полем P обозначается $G(K, P)$. Очевидно, в ней симметричным элементом к функции f является функция f^{-1} и нейтральным элементом — тождественная функция id вида $id: K \leftrightarrow K$.

в) Группой Галуа уравнения $f(x)=0$ называется группа Галуа поля разложения P_f многочлена f над полем коэффициентов P_0 этого уравнения, т. е. группа Галуа вида $G(P_f, P_0)$.

Группой Галуа многочлена f из кольца $P[x]$ и группой Галуа соответствующего уравнения $f(x)=0$ называется группа $G(P_f, P)$. При этом поле P может быть шире поля коэффициентов P_0 уравнения $f(x)=0$.

Напомним, что многочлен f из кольца $P[x]$ называется неприводимым над полем P, если он не представим в виде произведения двух многочленов из $P[x]$, степени которых больше или равны 1. В противном случае многочлен f называется приводимым над P. Мы будем рассматривать только многочлены степени, большей или равной 1, не оговаривая в дальнейшем это обстоятельство.

Если какой-то многочлен неприводим над полем P, то он не имеет корней в P. Обратное неверно, например, для многочлена $(x^2+x+1)^2$ над полем $\boldsymbol{R}$.

В курсе алгебры доказывается критерий неприводимости для произвольного многочлена над полем $\boldsymbol{Q}$. Он называется критерием Эйзенштейна и формулируется следующим образом: пусть $f \in \boldsymbol{Q}[x]$ и многочлен f_1 получается из f умножением f на наименьшее общее кратное знаменателей всех коэффициентов в f. Тогда $f_1 \in \boldsymbol{Z}[x]$. Если все коэффициенты многочлена f_1, кроме старшего, делятся на какое-то простое число q и свободный член в f_1 не делится на число q^2, то исходный многочлен f неприводим над полем $\boldsymbol{Q}$.

Например, по этому критерию любой многочлен вида $x^2+p\cdot x+p$, где p — какое-то фиксированное простое число, неприводим над $\boldsymbol{Q}$, так как в качестве числа q можно выбрать исходное число p.

Для произвольного поля P даже в случае $\boldsymbol{Q} \subset P \subseteq \boldsymbol{R}$ неизвестно каких-либо явных критериев неприводимости над ним произвольного многочлена. Уравнения над такими полями гораздо труднее исследовать.

Неприводимым уравнением над полем P назовем уравнение вида $f(x)=0$, где многочлен f неприводим над P.

Понятие неприводимого над полем P уравнения интересно для нас тем, что в случае, когда многочлен f приводим над P, любой вопрос о корнях уравнения $f(x)=0$ сводится к вопросу о корнях сомножителей f и, таким образом, сводится к вопросу о корнях нескольких, уже неприводимых над P уравнений.

Поэтому в дальнейшем будем считать, что все рассматриваемые уравнения и многочлены неприводимы над исходным полем P (если специально не оговорено противное).

Важность группы Галуа видна из следующей фундаментальной теоремы и того замечания, которое за ней следует.

Теорема Э. Галуа. ***Неприводимое уравнение $f(x)=0$, где $f \in P[x]$, разрешимо в радикалах в том и только в том случае, если его группа Галуа $G(P_f, P)$ разрешима.***

Замечания. **1.** Эта теорема не только сводит вопрос о выразимости с помощью радикалов корней уравнения $f(x)=0$ над полем P к вопросу о свойствах одной определенной группы. Она дает больше: группа Галуа для многих конкретных уравнений сравнительно легко и явно вычисляется и для нее сравнительно легко выясняется, разрешима ли она. Для таких уравнений теорема Галуа становится практическим критерием разрешимости уравнения в радикалах.

2. Теорема Галуа остается верной и для произвольных, не обязательно неприводимых уравнений. Мы не рассматриваем этот более сложный случай хотя бы потому, что любое уравнение сводится к неприводимым уравнениям.

В этом параграфе мы докажем первую часть теоремы Галуа (теорему 6), а в следующем параграфе — ее вторую часть.

Теорема 6. ***Если неприводимое уравнение $f(x)=0$, где $f \in P[x]$, разрешимо в радикалах, то его группа Галуа $G(P_f, P)$ разрешима.***

Чтобы лучше уяснить формулировку этой теоремы, сначала рассмотрим ее применения, а только потом, в пункте 7 этого параграфа разберем ее доказательство. Для применения этой теоремы, в первую очередь, нужно научиться вычислять группу Галуа данного уравнения. Даже для многочлена второй степени $f(x)=x^2+a\cdot x+b$ пока неясно, как ее фактически находить. Ниже излагаются предложения 4—8 и теорема 7, которые позволяют явно вычислять группу Галуа для многих конкретных много-

членов. При этом в доказательствах используется следующее вспомогательное.

Предложение 4. ***а) Любое поле вида $K(a_1, ..., a_s)$ состоит в точности из значений в точке $(a_1, ..., a_s)$ всех таких рациональных функций f из поля $K(x_1, ..., x_s)$, которые определены в этой точке.***

б) Если a — корень неприводимого над полем K многочлена g степени n, то поле $K(a)$ состоит в точности из значений в точке a всех многочленов степени меньшей или равной $n-1$ из кольца $K[x]$. Размерность векторного пространства с носителем $K(a)$ над полем K равна числу n.

в) Пусть числа $a_1, ..., a_n$ являются корнями каких-то многочленов над полем P и $K \rightleftarrows P(a_1, ..., a_n)$. Тогда любой элемент из поля K является значением в точке $(a_1, ..., a_n)$ некоторого многочлена из кольца $K[x_1, ..., x_n]$. (Числа $a_1, ..., a_n$ называются алгебраическими над P.)

г) Любое неприводимое уравнение не имеет кратных корней.

▷ а) Если $f(x_1, ..., x_s) \in K(x_1, ..., x_s)$, то, конечно, $f(a_1, ..., a_s) \in K(a_1, ..., a_s)$. Обозначим $K_1 \rightleftarrows K(a_1, ..., a_s)$ и $X \rightleftarrows \{f(a_1, ..., a_s) | f \in K(x_1, ..., x_s)\}$ и получим $X \subseteq K_1$. Множество X содержится в $\boldsymbol{C}$, и относительно обычных операций над числами это множество, очевидно, является полем, содержащим поле K и числа $a_1, ..., a_s$. Поэтому поле K_1 содержится в X. Следовательно, $K_1 = X$.

б) Очевидно, выполняется $X \rightleftarrows \{f(a) | f \in K[x]\} \subseteq K(a)$ и $K \subseteq X$, $a \in X$. Если X — поле, то $K(a) \subseteq X$, и, следовательно, $K(a) = X$. Проверим, что множество X является полем. Для этого достаточно установить, что $(f(a))^{-1} \in X$, если $f(a) \neq 0$.

Многочлены f и g взаимно просты над K, так как если бы они имели общий делитель степени $\geqslant 1$, то он, являясь нетривиальным делителем неприводимого многочлена g, совпадал бы с $k \cdot g$, где $k \in P$. Тогда $f(x) = k \cdot g(x) \cdot h(x)$ и $f(a) = 0$. Получили противоречие.

Поскольку g и f взаимно просты, то по известной теореме[1] из курса алгебры существуют многочлены h_1 и h_2 над полем K, для которых

$$f(x) \cdot h_1(x) + g(x) \cdot h_2(x) \equiv 1.$$

Тогда $f(a) \cdot h_1(a) = 1$ и $(f(a))^{-1} = h_1(a)$ и $h_1(a) \in X$.

Итак, $X = K(a)$. Любой многочлен f из кольца $K[x]$ можно поделить с остатком на многочлен g. Получится $f = g \cdot g_1 + r$, где g_1, r — многочлены из $K[x]$ и r из степени $\leqslant n-1$. Тогда $f(a) = r(a)$.

Поэтому $X = \{r(a) | r \in K[x]$, степень $r \leqslant n-1\}$.

Отсюда следует, что любой элемент k из $K(a)$ представим в виде $k = k_0 + k_1 \cdot a + ... + k_{n-1} \cdot a^{n-1}$. Поэтому линейные комбинации над полем K элементов (чисел) $1, a, ..., a^{n-1}$ образуют все множество $K(a)$. Убедимся, что эти элементы линейно независимы. Если они линейно зависимы, то существуют такие скаляры (числа) $k_0, ..., k_{n-1}$ из

[1] См.: *Куликов Л. Я.*, с. 527.

K, не все равные нулю, что $k_0 + k_1 a + ... + k_{n-1} a^{n-1} = 0$. Обозначим $g_1(x) \rightleftarrows k_0 + k_1 \cdot x + ... + k_{n-1} \cdot x^{n-1}$.

Число a является корнем этого многочлена над K. Поэтому многочлены g_1 и g имеют над полем $K(a)$ наибольший общий делитель степени $\geqslant 1$. По замечательному свойству алгоритма Евклида этот наибольший общий делитель является многочленом над полем ^1K. Поэтому многочлен g имеет делитель над полем K степени от 1 до $n-1$, что противоречит его неприводимости над этим полем.

Итак, множество

$$\{1, a, ..., a^{n-1}\}$$

является базисом в векторном пространстве $K(a)$ над полем скаляров K. (Вспомним, что любое поле является векторным пространством над любым своим подполем!) Следовательно, размерность $K(a)$ над K равна n.

в) Доказательство проведем индукцией по n. Если $n = 1$, то $K = P(a_1)$ и утверждение 4,в является частным случаем утверждения 4, б.

Допустим, что утверждение доказано для всех натуральных чисел, строго меньших n.

Пусть $L \rightleftarrows P(a_1, ..., a_{n-1})$ и $k \in K \rightleftarrows P(a_1, ..., a_n) = L(a_n)$. По предложению 4,б найдется такой многочлен $h(x)$ с коэффициентами из поля L, что $k = \gamma_0 + \gamma_1 \cdot a_n + ... + \gamma_m \cdot a_n^m$, где $\gamma_i \in L$.

По предположению индукции все коэффициенты γ_i многочлена $h(x)$ равны значениям в точке $(a_1, ..., a_{n-1})$ каких-то многочленов h_i над полем P, зависящих от $n-1$ переменной. Положим по определению

$$g(x_1, ..., x_n) \rightleftarrows h_0(x_1, ..., x_{n-1}) + h_1(x_1, ..., x_{n-1}) \cdot x_n + ... + h_m(x_1, ..., x_{n-1}) \cdot x_n^m.$$

Тогда $g(x_1, ..., x_n)$ — многочлен над P и

$$k = h(a_n) = g(a_1, a_2, ..., a_{n-1}, a_n).$$

г) Допустим, уравнение $f(x) = 0$ над полем P степени n и имеет корень x_0 кратности $\geqslant 2$. Тогда этот корень является и корнем уравнения $f'(x) = 0$. Следовательно, наибольший общий делитель g многочленов $f(x)$ и $f'(x)$ над полем $P(x_0)$ имеет степень $\geqslant 1$. При этом сами $f(x)$ и $f'(x)$ являются многочленами именно над полем P. Следствием алгоритма Евклида является утверждение о том, что g является многочленом именно над полем ^1P. Поэтому g — делитель f над P. Степень многочлена g, с одной стороны, $\geqslant 1$ и, с другой стороны, $\leqslant (n-1)$, так как g — делитель и многочлена f'. Это противоречит условно о неприводимости многочлена f над полем P. ■

[1] См.: *Куликов Л. Я.*, с. 454, 471.

Предложение 5. ***Для любого алгебраического уравнения***

$$f(x)=x^n+a_1\cdot x^{n-1}+\ldots+a_n=0$$

его группа Галуа $G(P_f, P)$ изоморфна подгруппе группы S_n над полем P.

▷ Поле разложения P_f многочлена f обозначим K, т. е. $K\rightleftarrows P_f\rightleftarrows P(x_1, \ldots, x_n)$, где $x_1, \ldots, x_n$ — все корни исходного уравнения $f(x)=0$. Разумеется, $x_i\in K$.

Для доказательства этого принципиально важного утверждения определим инъективный гомоморфизм (вложение) ψ вида $\psi: G(K, P)\to S_n$. Тогда ψ является изоморфизмом вида $\psi: G(K, P)\leftrightarrow R(\psi)$ и $R(\psi)$ является подгруппой группы S_n.

Поскольку изоморфные объекты, в частности изоморфные группы, обычно не различают, то наличие такого изоморфизма означает, что группу $G(K, P)$ можно отождествить с группой $R(\psi)$ и, следовательно, отождествить с подгруппой группы S_n.

Рассмотрим сначала случай (основной для нас), когда многочлен f неприводим над P и, следовательно, по предложению 4,г все его корни $x_1, \ldots, x_n$ различны. В этом случае не играет роли, писать ли сам корень x_i или его номер i, так как корни $x_1, \ldots, x_n$ и их номера $1, \ldots, n$ находятся в фиксированном биективном соответствии.

Итак, положим:

$$\psi(W)\rightleftarrows\begin{pmatrix} x_1 & \ldots & x_n \\ W\cdot x_1 & \ldots & W\cdot x_n\end{pmatrix}, \text{ где } W\in G(K, P).$$

Числа вида $W\cdot x_i$ являются корнями уравнения $f(x)=0$, так как $W(f(x_i))=W(0)=0$ и $W(f(x_i))=f(W\cdot x_i)$, т. е. $f(W\cdot x_i)=0$. Здесь используется тот факт, что $W(x^n+a_1\cdot x^{n-1}+\ldots+a_n)=(W\cdot x)^n+a_1\times\times(W\cdot x)^{n-1}+\ldots+a_n$ для любого элемента $x\in K$.

Если $x_i\neq x_j$, то $W\cdot x_i\neq W\cdot x_j$, так как $W: K\to K$ — инъективная функция. Поэтому в нижней строке таблицы $\begin{pmatrix} x_1 & \ldots & x_n \\ W\cdot x_1 & \ldots & W\cdot x_n\end{pmatrix}$ находятся n различных чисел. Все они корни уравнения $f(x)=0$, а всего таких корней не более n. Следовательно, в нижней строке содержатся все n корней этого уравнения, т. е. эта таблица определяет биекцию верхней строки на себя. Поэтому она является элементом группы S_n, т. е. $\psi: G(K, P)\to S_n$. Функция $\psi(\cdot)$ инъективна. Если $W\neq W'$ и $(W\cdot x_1, \ldots, W\cdot x_n)=(W'\cdot x_1, \ldots, W'\cdot x_n)$, то для любого числа k из поля K по предложению 4,а найдется представление в виде

$$k=\frac{h_1(x_1, \ldots, x_n)}{h_2(x_1, \ldots, x_n)}, \text{ где } h_1, h_2\in P[u_1, \ldots, u_n].$$

Тогда $W\cdot k=\frac{h_1(W\cdot x_1, \ldots, W\cdot x_n)}{h_2(W\cdot x_1, \ldots, W\cdot x_n)}=\frac{h_1(W'\cdot x_1, \ldots, W'\cdot x_n)}{h_2(W'\cdot x_1, \ldots, W'\cdot x_n)}=W'\cdot k.$

Получили противоречие. (Конечно, вместо предложения 4,а можно было сослаться на предложение 4,в.)

Осталось проверить, что ψ — гомоморфизм. Действительно, $\psi(W_1\circ W_2)=\begin{pmatrix} x_1 & \ldots \\ (W_1\cdot W_2)\cdot x_1 & \ldots\end{pmatrix}=\begin{pmatrix} x_1 & \ldots \\ W_1(W_2x_1)\ldots\end{pmatrix}=\begin{pmatrix} x_1 & \ldots \\ W_2\cdot x_1 & \ldots\end{pmatrix}\times$

$\times\begin{pmatrix} \ldots W_2x_1 & \ldots \\ \ldots W_1W_2x_1 & \ldots\end{pmatrix}=\psi(W_1)\circ\psi(W_2).$

Если исходное уравнение $f(x)=0$ не предполагается неприводимым, то среди его корней $x_1, \ldots, x_n$ могут быть равные.

Пусть среди них попарно различных корней ровно m, где $m\leqslant n$. Тогда множество чисел $\{1, \ldots, m\}$ биективно с множеством попарно различных корней $x_1, \ldots, x_m$, и, следовательно, для функции $\psi(W)\rightleftarrows\rightleftarrows\begin{pmatrix} x_1 & \ldots x_m \\ W\cdot x_1 & \ldots W\cdot x_m\end{pmatrix}$ проходит то же самое рассуждение, что и выше. В результате группа $G(K, P)$ вкладывается в группу S_m, где $m\leqslant n$. Но любая группа S_m изоморфна подгруппе в S_n с помощью инъективного гомоморфизма вида

$$\psi_1: \begin{pmatrix} 1 & \ldots & m \\ k_1 & \ldots & k_m\end{pmatrix}\mapsto\begin{pmatrix} 1 & \ldots & m & m+1 & \ldots & n \\ k_1 & \ldots & k_m & m+1 & \ldots & n\end{pmatrix}.$$

Поэтому и в случае кратных корней группа $G(K, P)$ также изоморфна подгруппе группы S_n. ■

Замечание. Таким образом, группе Галуа $G(P_f, P)$ раз и навсегда фиксированным способом соответствует некоторая подгруппа группы S_n всех перестановок корней уравнения n-ой степени $f(x)=0$ или, что то же самое, всех чисел (номеров этих корней) $1, \ldots, n$. Группа Галуа $G(P_f, P)$ называется транзитивной, если соответствующая ей подгруппа группы S_n транзитивна. Как уже отмечалось, любая подгруппа группы S_n действует на множестве $\{1, \ldots, n\}$. Поэтому группа Галуа $G(P_f, P)$ действует не только на множестве P_f, но и на его части — множестве всех корней исходного уравнения.

Предложение 5 имеет принципиальное значение, потому что оно позволяет искать группу Галуа исходного уравнения среди подгрупп одной, конкретной группы S_n. В частности, из него следует, что группа Галуа всегда конечна. Предложение 5, теорема 5 и упражнение 8 немедленно дают первое глубокое следствие теоремы Галуа (которое, правда, другим способом в основном было получено в п. I).

Теорема 7. ***Любое уравнение степени $\leqslant 4$ имеет в качестве группы Галуа какую-то подгруппу группы S_4 и, следовательно, разрешимую группу. Поэтому исходное уравнение разрешимо в радикалах.***

В следующем пункте мы продолжим исследование вопроса о том, какой конкретный вид имеет группа Галуа исходного (неприводимого) уравнения.

6. Пример алгебраического уравнения, группа Галуа которого совпадает с симметрической группой 5-й степени.

Предложение 6. ***Пусть K — поле разложения какого-то***

многочлена f над полем P, т. е. $K=P_f$, и g — какой-то неприводимый над P многочлен. Тогда для любых двух корней a и b уравнения $g(x)=0$, содержащихся в поле K, найдется такой элемент W_0 из группы Галуа $G(K, P)$, что

$$W_0(a)=b.$$

▷ Пусть в множестве $G(K, P)$ ровно m элементов. Перенумеруем все элементы группы $G(K, P)$, начиная, как обычно, с единичного элемента: $W_1 \rightleftarrows id$, W_2, ..., W_m.

Рассмотрим все числа вида

$$W_1(a), \ldots, W_m(a) \qquad (*)$$

и докажем, что в этой последовательности чисел содержится число b. Тогда b вида $b=W_l(a)$, где l — какое-то число от 1 до m, и, следовательно, искомое W_0 равно W_l.

Все числа из последовательности $(*)$ содержатся в поле K, так как $W_i : K \to K$. Если $W \in G(K, P)$, то для любого i, где $1 \leqslant i \leqslant m$, число $W \cdot W_i(a)=(W \cdot W_i)(a)$ находится в последовательности чисел $(*)$, так как произведение биекций $W \cdot W_i$ совпадает с каким-то элементом W_j, где $1 \leqslant j \leqslant m$, группы $G(K, P)$. Образуем многочлен над полем K вида

$$h(x) \rightleftarrows \prod_{i=1}^{m} (x - W_i(a)).$$

Очевидно, он степени m и множество его корней в точности совпадает с множеством чисел в последовательности $(*)$, и, главное, все его коэффициенты α не меняются под действием на них любых биекций W из $G(K, P)$, т. е. $W \cdot \alpha = \alpha$. Последнее легко проверить, явно выписывая коэффициенты этого многочлена (сделайте это, например, для $m=3$).

В приложении 3 в конце книги доказывается следующая

Лемма 9. ***Любое число k из поля $K \rightleftarrows P_f$ неподвижное под действием всех изоморфизмов из группы Галуа $G(K, P)$, т. е. $\forall W \in G(K, P)\ (W(k)=k)$, принадлежит полю P.***

Поэтому все коэффициенты многочлена h принадлежат полю P, и, следовательно, h многочлен над полем P. Число $a=W_1(a)$ является корнем многочленов и g, и h. Поэтому g и h имеют наибольший общий делитель степени, большей или равной 1 над полем K и, следовательно, над полем P (см. доказательство предложений 4,б, г). Поскольку многочлен g неприводим над P, то таким делителем может быть только $k \cdot g$, где $k \in P$.

Итак, g делит h над полем P, и, следовательно, каждый корень g является корнем h. В частности, корень b многочлена g является корнем многочлена h, т. е. одним из чисел в последовательности $(*)$. Поэтому b совпадает с каким-то числом $W_l(a)$, и тогда искомое W_0 равно W_l.

Именно в упомянутой лемме 9 из приложения 3 используется условие о том, что поле $K \rightleftarrows P_f$ является полем разложения какого-то

многочлена f. При первом чтении разумно принять эту лемму без доказательства. ■

Замечания. **1.** Многочлены f и g могут совпадать, но могут и отличаться.

2. В дальнейшем мы будем часто использовать предложение 6 для группы Галуа $G(K, P)$ и неприводимого многочлена g над полем P. При этом, конечно, нужно проверять условие этого предложения, касающееся вида поля K, т. е. проверять существование такого многочлена f, что $K=P_f$. Это не всегда легко, однако во всех рассматриваемых далее случаях это так, т. е. возможен выбор такого многочлена f, что $K=P_f$. Такую проверку иногда будем оставлять для самостоятельной работы.

3. Если числа a и b являются корнями приводимого уравнения $g(x)=0$ над полем P, то в общем случае неверно, что эти числа являются корнями какого-то одного неприводимого над P уравнения (в последнем случае числа a и b называются сопряженными над полем P).

4. Пусть K — поле разложения некоторого многочлена над полем P и g неприводимый над P многочлен. Если X множество всех корней многочлена g, то группа $G(K, P)$ транзитивно действует на множестве X.

Из предложения 6 уже вытекает ответ на вопрос о конкретном виде группы Галуа для неприводимого уравнения степени, меньшей или равной 3, и после некоторых дополнительных усилий — для неприводимого уравнения степени, меньшей или равной 4. А именно для неприводимого уравнения 2-й степени $f(x)=0$ над полем P группа Галуа $G(P_f, P)$ является подгруппой группы S_2. Таких подгрупп ровно две: $\{id\}$ и S_2. По предложению 6 группа Галуа транзитивна, т. е. в ней содержится биекция, переводящая различные корни исходного уравнения друг в друга. Такая подгруппа ровно одна: S_2. Итак, $G(P_f, P)=S_2$.

Упражнение 10. *Докажите, что транзитивными подгруппами группы S_3 являются ровно две подгруппы: сама группа S_3 и ее знакопеременная подгруппа A_3.*

Поэтому группа Галуа неприводимого уравнения 3-й степени совпадает с S_3 или A_3.

Как узнать для конкретных уравнений, какой из этих двух случаев имеет место? Дискриминантом кубического уравнения $f(x)=x^3+a \cdot x^2+b \cdot x+c=0$ называется число

$$D_f=a^2 \cdot b^2-4 \cdot b^3-4 \cdot a^3 \cdot c-18a \cdot b \cdot c-27 \cdot c^2.$$

Очевидно, $D_f \in P$.

Упражнение 11*. Докажите, что $G(P_f, P)=S_3$, если $\sqrt{D_f} \notin P$, и $G(P_f, P)=A_3$, если $\sqrt{D_f} \in P$.

Указание. Проверьте, что $D_f=\prod_{n \geqslant i > j \geqslant 1} (x_i-x_j)^2$, где x_1, ..., x_n — все корни уравнения n-й степени (здесь $n=3$).

Для неприводимого уравнения 4-й степени можно, хотя и не-

сколько более длинно, чем в случае уравнения 3-й степени, подсчитать все транзитивные подгруппы группы S_4.

Упражнение 12*. *Докажите, что транзитивными подгруппами группы S_4 являются ровно 6 подгрупп: сама группа S_4, знакопеременная группа A_4, клейновская группа B_4, группа $B_4' \neq \{e, t_1, t_2, t_3, \sigma_{12}, t_1\cdot\sigma_{12}, t_2\cdot\sigma_{12}, t_3\cdot\sigma_{12}\}$ и еще две группы вида $(23)\cdot B_4'\cdot(23)$ и $(24)\cdot B_4'\cdot(24)$, а также циклическая группа 4-го порядка с образующим элементом* (1 2 3 4).

Поэтому группа Галуа неприводимого уравнения 4-й степени совпадает с одной из этих 7 групп. Например, если исходное уравнение не имеет корней в поле P (это наиболее содержательный случай, так как иначе можно перейти к уравнению над полем P более низкой степени), то его группа Галуа либо S_4, либо A_4. Какой из этих двух случаев имеет место, можно узнать по дискриминанту исходного уравнения, он определяется с помощью произведения, которое для случая уравнения 3-й степени приведено в упражнении 11.

Однако, не так легко явно найти группу Галуа неприводимого уравнения 4-й и, тем более, 5-й степени, основываясь только на предложении 6. На помощь приходит то обстоятельство, что можно и далее конкретизировать ее вид.

Предложение 7. ***Если уравнение $f(x)=0$ рассматривается над полем P, содержащимся в R, и уравнение имеет ровно два невещественных корня, то его группа Галуа $G(P_f, P)$ содержит хотя бы одну транспозицию.***

▷ Поскольку коэффициенты многочлена f — вещественные числа, то его невещественные корни сопряжены, т. е. имеют вид $a_{1,2}=a\pm bi$. Остальные корни $a_3, \ldots, a_n$ этого многочлена по предположению вещественные числа. Поле разложения многочлена f определяется как $P_f=P(a_1, \ldots, a_n)$, а любой его элемент t по предложению 4,в имеет вид:

$$t=g(a_1, \ldots, a_n),$$

где $g(x_1, \ldots, x_n)\in P[x_1, \ldots, x_n]$. Поскольку все коэффициенты многочлена g — вещественные числа, то

$$\bar{t}=g(\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n)=g(a_2, a_1, a_3, \ldots, a_n)\in P_f.$$

Положим $W(t)\neq\bar{t}$, где $(\bar{\cdot})$ — функция сопряжения комплексного числа. Тогда $W: P_f\to P_f$ и $\forall k\in P(W(k)=k)$.

Легко проверить, что W — изоморфизм, поэтому $W\in G(P_f, P)$.

В соответствии с предложением 5 этому изоморфизму соответствует перестановка $\sigma_{12}=\begin{pmatrix}1&2&3&\ldots&n\\2&1&3&\ldots&n\end{pmatrix}$, которая является транспозицией. ■

Из предложений 6 и 7 получаем следствие: *неприводимое уравнение, удовлетворяющее условиям предложения 7, имеет транзитивную группу Галуа, содержащую транспозицию.*

Это сочетание двух свойств — транзитивности и наличия транспозиции — позволяет по-новому подойти к отысканию группы Галуа для неприводимого уравнения с вещественными коэффициентами 3-й степени, которое имеет ровно один вещественный корень. А именно его группа Галуа содержит транспозицию, например, $\sigma_2=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$ (см. список всех элементов группы S_3 на с. 249). Группа Галуа транзитивна, т. е. содержит перестановки, переводящие, в частности, 1 в 2 и 1 в 3. Пусть, например, это будут перестановки $\sigma_4=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$ и $\sigma_3=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$. Тогда она содержит перестановку $\sigma_2\cdot\sigma_3=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}=\sigma$ и перестановку $\sigma_2\cdot\sigma_4=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}=\sigma_5$. Поэтому $G(P_f, P)=S_3$. Продолжите это рассуждение, разобрав в нем все возможные варианты значений σ_2 и σ_4, σ_3. Аналогичные, но более длинные рассуждения позволяют по-новому найти группу Галуа для неприводимого уравнения 4-й степени с вещественными коэффициентами, имеющего ровно два невещественных корня. Однако главный вывод из этого сочетания свойств группы Галуа содержится в следующем предложении 8.

Отвлечемся сейчас для обсуждения одного вспомогательного вопроса, важного при практическом использовании предложения 7. Каким образом узнать, имеет ли уравнение $f(x)=0$ ровно два невещественных корня? Для этого служит одна из наиболее фундаментальных теорем алгебры — теорема Штурма ([1], с. 524), которую мы сейчас напомним.

Пусть f — многочлен степени, большей или равной 1, с вещественными коэффициентами. Положим:

$f_0\neq f$, $f_1\neq f'$, где f' — производная функция f, и

$$\begin{aligned}
f_0&=g_1\cdot f_1-f_2, \ \deg f_2<\deg f_1,\\
f_1&=g_2\cdot f_2-f_3, \ \deg f_3<\deg f_2,\\
&\vdots\qquad\qquad\vdots\\
f_{m-2}&=g_{m-1}\cdot f_{m-1}-f_m, \ \deg f_m<\deg f_{m-1}.
\end{aligned}$$

Таким образом, конечная последовательность $f_0, f_1, f_2, \ldots, f_m$ получается применением к многочленам f и f' алгоритма Евклида с той разницей, что всякий раз знак остатка меняется на противоположный.

Для определения числа вещественных корней многочлена f необязательно знать саму последовательность $f_0, f_1, f_2, \ldots, f_m$, достаточно знать последовательность знаков старших коэффициентов и последовательность степеней многочленов $f_0, f_1, \ldots, f_m$. Поэтому если каким-либо приемом можно найти эти две последовательности, не вычисляя самих многочленов f_i, то этого достаточно. Последовательность многочленов $f_0, f_1, \ldots, f_m$ называется последовательностью Штурма.

Например, пусть $f(x)=x^5-n\cdot x+n$, где n — любое фиксированное натуральное число. Соответствующая последовательность Штурма находится следующим образом:

$f_1(x)\neq f'(x)=5x^4-n$; $f(x)=f_1(x)\cdot g(x)+r(x)$, где

$$\frac{x^5-nx+n}{-\frac{4}{5}nx+n}\ \Big|\ \frac{5x^4-n}{\frac{1}{5}x}$$

и $r(x)=-\frac{4}{5}nx+n$, отсюда $f_2(x)=\frac{4}{5}nx-n$; $f_1(x)=f_2(x)\cdot q(x)+r(x)$, где $\deg r<$

$<\deg f_2$, т. е. $\deg r=0$

и r — постоянная. Поэтому найдем r, не выполняя фактически деления f_1 на f_2.

Корнем многочлена f_2, очевидно, является число $x_0=\frac{5}{4}$ и $r=f_1(x_0)=\frac{5\cdot 5^4}{4^5}-n$.

Пусть $n\geqslant\frac{5^5}{4^4}=12,\ 21\ \ldots$. Тогда $r\leqslant 0$ и $f_3\geqslant 0$. Выпишем знаки и степени многочленов из последовательности Штурма:

$$\begin{array}{cccc} + & + & + & + \\ 5 & 4 & 1 & 0 \end{array}$$

Теперь по теореме Штурма нужно подсчитать число перемен знаков в верхней последовательности (оно, очевидно, равно нулю) и вычесть его из числа перемен знаков в следующей последовательности: $-\ +\ -\ +$, которая образуется по правилу $(-1)^5\cdot(+1),\ (-1)^4\cdot(+1),\ (-1)^1\cdot(+1),\ (-1)^0\cdot(+1)$.

В данном случае получится число 3. Это означает, что исходное уравнение $f(x)=0$ имеет ровно три вещественных корня и, следовательно, ровно два невещественных корня при $n\geqslant 13$.

Предложение 8. ***Транзитивная подгруппа группы S_p простой степени p, содержащая транспозицию, совпадает с S_p.***

Прежде чем доказывать это предложение, приведем его важное

Следствие. Неприводимому уравнению простой степени над числовым полем P, содержащемся в R, которое имеет ровно два невещественных корня, соответствует в качестве группы Галуа группа S_p.

▷ Оно непосредственно вытекает из следствия к предложению 7 и предложения 8. ■

▷ Изложим доказательство предложения 8, которое при первом чтении лучше пропустить. Обозначим транзитивную подгруппу, о которой говорится в этом предложении, буквой G. По условию она содержит какую-то транспозицию вида (i_1, i_2) и, кроме нее, может быть, еще какие-то транспозиции, содержащие число i_1.

Выпишем их все:

$$(i_1i_2),\ (i_1,\ i_3),\ \ldots,\ (i_1i_m). \qquad (*)$$

Следовательно, во-первых, G содержит все транспозиции вида (i_ki_l), где $1\leqslant k,\ l\leqslant m$, так как

$$(i_ki_l)=(i_1i_k)\cdot(i_1,\ i_l)\cdot(i_1i_k),$$

и, во-вторых, G не содержит ни одной транспозиции вида (ji_k), где $1\leqslant k\leqslant m$ и $j\neq i_1,\ \ldots,\ i_m$, так как в противном случае она содержала бы транспозицию $(i_1j)=(i_1i_2)\,(ji_2)\,(i_1i_2)$.

Последнее противоречит тому, что список $(*)$ содержит все транспозиции, содержащие i_1. Это второе утверждение о подгруппе G будет в дальнейшем доказательстве часто использоваться; обозначим его $(**)$.

Мы хотим показать, что на самом деле выполняется $m=p$ и, следовательно, среди чисел $i_1,\ \ldots,\ i_m$ находятся все числа от 1 до p.

Тогда G содержит любую транспозицию, так как для любых чисел $i,\ j$ выполняется $(ij)=(i_1i)\,(i_1j)\,(i_1i)$.

Отсюда по теореме 4,а подгруппа G содержит все перестановки, т. е. $G=S_p$, что и требуется доказать.

Допустим, выполняется $m<p$. Тогда существует число j, $1\leqslant j\leqslant p$, отличное от всех чисел $i_1,\ \ldots,\ i_m$. Поскольку G транзитивна, то существует перестановка $\sigma_1\in G$, для которой $\sigma_1(i_1)=j$.

Пусть σ_1 вида

$$\sigma_1=\begin{pmatrix} i_1\, i_2\, \ldots\, i_m\, \ldots \\ j\, j_2\, \ldots\, j_m\, \ldots \end{pmatrix}.$$

Тогда выполняется $\sigma_1^{-1}(i_1i_k)\sigma_1=(jj_k)\in G$ для любого числа k, где $2\leqslant k\leqslant m$.

Действительно, $(\sigma_1^{-1}(i_1i_k)\sigma_1)\,(j)=j_k$, и аналогичным образом j_k переходит в j. Для $l\neq j,\ j_k$ получаем, что $\sigma_1^{-1}(l)\neq i_1,\ i_k$, т. е. $(\sigma_1^{-1}(i_1i_k)\sigma_1)\,(l)=(\sigma_1^{-1}\sigma_1)\,(l)$. Ни одно из чисел $j_1\neq j,\ j_2,\ \ldots,\ j_m$ не совпадает ни с одним из чисел $i_1,\ i_2,\ \ldots,\ i_m$. Действительно, число j отличается от всех чисел $i_1,\ \ldots,\ i_m$ по его выбору. Допустим, например, $j_2=i_m$, тогда, как только что доказано, транспозиция $(jj_2)=(ji_m)$ принадлежит G, что противоречит доказанному выше утверждению $(**)$.

Итак, мы нашли $2m$ попарно различных чисел $\{i_1,\ \ldots,\ i_m,\ j_1,\ j_2,\ \ldots,\ j_m\}$, и, следовательно, $2m\leqslant p$.

Если выполняется $2m<p$, то совершенно аналогично существует число k, $1\leqslant k\leqslant p$, отличное от всех этих $2m$ чисел. По транзитивности подгруппы G существует перестановка $\sigma_2\in G$, для которой $\sigma_2(i_1)=k$. Пусть σ_2 вида

$$\sigma_2=\begin{pmatrix} i_1\, i_2\, \ldots\, i_m\, \ldots \\ k\, k_2\, \ldots\, k_m\, \ldots \end{pmatrix}.$$

Как и в случае перестановки σ_1, доказывается, что ни одно из чисел $k_1\neq k,\ k_2,\ \ldots,\ k_m$ не равно ни одному из чисел $i_1,\ \ldots,\ i_m$. И, кроме того, ни одно из чисел $k_1,\ \ldots,\ k_m$ не равно ни одному из чисел $j_1,\ j_2,\ \ldots,\ j_m$. Действительно, если, например, $k_2=j_m$, то G содержит транспозицию

$$\sigma_1\sigma_2^{-1}(i_1i_2)\sigma_2\sigma_1^{-1}=(\sigma_1^{-1}(k)i_m).$$

Поскольку $k\neq j_1,\ j_2,\ \ldots,\ j_m$, то $\sigma_1^{-1}(k)\neq i_1,\ i_2,\ \ldots,\ i_m$, что противоречит утверждению $(**)$.

Итак, мы нашли множество попарно различных чисел $\{i_1,\ i_2,\ \ldots,\ i_m,\ j_1,\ j_2,\ \ldots,\ j_m,\ k_1,\ k_2,\ \ldots,\ k_m\}$, состоящее из $3m$ элементов. Поэтому $3m\leqslant p$. Ясно, что когда-то такое построение оборвется и будет выполняться равенство $g\cdot m=p$, поэтому число m делит число p, но по условию теоремы p — простое число и $m\geqslant 2$. Поэтому $m=p$. ■

По следствию к предложению 8 еще раз по-новому получаем, что неприводимые уравнения степени 2, 3, 5 над полем P, где $P \subseteq R$, у которых ровно два невещественных корня, имеют в качестве группы Галуа соответственно группы S_2, S_3, S_5. К уравнению 4-й степени это следствие неприменимо.

Теперь непосредственно из следствия к предложению 8 получаем второе глубокое следствие теоремы Галуа.

Теорема 8. *Любое уравнение вида*

$$x^5 - p \cdot x + p = 0,$$

где p — фиксированное простое число, большее или равное 13, имеет группой Галуа неразрешимую группу S_5, и, следовательно, оно неразрешимо в радикалах.

Замечание **1.** Теоремы 6, 7, 8 являются основными применениями теоремы Галуа к вопросу о решении в радикалах уравнений степени, меньшей или равной пяти.

2. Как мы увидим в § 3, теорема 7 не только утверждает разрешимость в радикалах любых уравнений с разрешимой группой Галуа, но и позволяет по этой группе фактически найти в радикалах корни таких уравнений.

7. Доказательство необходимого условия в теореме Галуа. Докажем теорему 6, приведенную на с. 243, которая, очевидно, является необходимым условием в теореме Галуа (сравните эти теоремы). Доказательство теоремы 6 довольно сложно и длинно. Оно состоит из нескольких частей, концы которых будут отмечаться знаком ■. Читатель с пользой для себя может сформулировать эти части как самостоятельные утверждения. В процессе доказательства мы используем несколько лемм, доказательства которых содержатся в приложении 3 в конце книги. При первом чтении разумно принять эти леммы на веру.

Общий план доказательства приводится ниже на с. 258.

Пусть неприводимое над полем P уравнение $f(x)=0$ разрешимо в радикалах. Это по определению означает, что существует радикальная цепочка полей вида

$$P \subset P'_1 \subset P'_2 \subset \ldots \subset P'_{s'}, \qquad (*)$$

в которой поле $P'_{s'}$ содержит все корни исходного уравнения. Тем самым $P_f \subseteq P'_{s'}$. Существенную роль в доказательстве играет следующее определение.

Поле K называется нормальным расширением поля P, если $P \subseteq K$, а K конечномерно над P, и любое неприводимое над P уравнение $g(x)=0$, имеющее в K хотя бы один корень, содержит в K все свои комплексные корни.

Лемма 10. ***Для того чтобы поле K являлось нормальным расширением поля P, необходимо и достаточно, чтобы K являлось полем разложения некоторого многочлена над P.***

Следствие. Если поле K является полем разложения какого-то многочлена над P, то число элементов в группе Галуа $G(K, P)$ совпадает с размерностью $\dim_P K$ поля K над полем P.

Лемма 11. ***Любое радикальное расширение $P'_{s'}$ поля P содержится в некотором нормальном радикальном расширении K поля P.***

По лемме 11 вместо цепочки $(*)$ можно рассмотреть радикальную цепочку

$$P \rightleftarrows P_0 \subset P_1 \subset \ldots \subset P_s \rightleftarrows K, \qquad (**)$$

в которой новые поля P_1, ..., P_s, может быть, отличаются от полей P'_1, ..., $P'_{s'}$, входивших в исходную радикальную цепочку, но поле K содержит прежнее радикальное расширение $P'_{s'}$ и потому содержит поле P_f. Следовательно, K содержит все корни исходного уравнения и — ради чего мы перешли от цепочки $(*)$ к цепочке $(**)$ — поле K есть нормальное расширение исходного поля P.

По определению радикальной цепочки каждое поле P_i является простым радикальным расширением предыдущего поля P_{i-1}.

Напомним, что простым радикальным расширением поля P_{i-1} мы назвали поле вида $P_i \rightleftarrows P_{i-1}(\zeta, \sqrt[m]{a_{i-1}})$, где a_{i-1} — какой-то фиксированный элемент из P_{i-1} и ζ — первообразный корень m-й степени из единицы. Поле P_i является полем разложения уравнения над полем P_{i-1} вида $g(x) \rightleftarrows (x^m - a_{i-1}) = 0$.

Действительно, пусть θ — какое-то одно значение корня m-й степени из числа a_{i-1}, т. е. $\theta \rightleftarrows \sqrt[m]{a_{i-1}}$. Тогда P_i содержит P_{i-1} и все m корней θ, $\theta \cdot \zeta$, ..., $\theta \cdot \zeta^j$, ..., $\theta \cdot \zeta^{m-1}$ уравнения $g(x)=0$ (см. п. 1). Поэтому $P_g \subseteq P_i$. Наоборот, P_g содержит P_{i-1} и числа θ и $\theta \cdot \zeta$ как корни уравнения $g(x)=0$. Поэтому P_g содержит P_{i-1} и числа θ и $\zeta = \frac{\theta \cdot \zeta}{\theta}$, т. е. $P_i \subseteq P_g$. (Здесь существенно используется то, что уравнение $g(x)=0$ и соответственно его поле разложения P_g рассматриваются именно над полем P_{i-1}.)

По лемме 10 получаем, что каждое P_i является нормальным расширением поля P_{i-1}.

Образуем цепочку групп Галуа вида

$$G(K, P) = G(K, P_0), \ldots, G(K, P_i), \ldots, G(K, P_s) = \{id\}. \qquad (***)$$

Оказывается, эти группы включены друг в друга и при этом в обратную сторону по сравнению с исходной цепочкой $(**)$! А именно

$$G(K, P) \supset G(K, P_1) \supset \ldots \supset G(K, P_i) \supset \ldots \supset G(K, P_s).$$

Действительно, если $W \in G(K, P_i)$, то $W : K \to K$ — изоморфизм и $\forall t \in P_i (W(t) = t)$. Тогда тем более $\forall k \in P_{i-1} (W(k) = k)$, так как $P_{i-1} \subset P_i$. Обозначим группы, составляющие цепочку $(***)$, соответственно G, G_1, ..., G_i, ..., $G_s = \{\mathrm{id}\}$. ■

1) Поскольку $P_{i-1} \subset P_i \subset K$ и P_i — нормальное расширение P_{i-1}, то G_i — нормальный делитель G_{i-1}. Докажем это.

Пусть W — какой-то элемент группы $G(K, P_{i-1})$. Положим $W' \rightleftarrows W \restriction P_i$. За счет нормальности P_i над P_{i-1} получаем $\forall t \in P_i (W'(t) = W(t) \in P_i)$. Действительно, P_i — конечномерное векторное пространство над полем P_{i-1} какой-то размерности e. Поэтому для любого фиксированного t из P_i выполняется $1, t^1, t^2, \ldots, t^e$ — линейно зависимая система элементов над P_{i-1}, так как иначе в P_i над P_{i-1} существовал бы базис из $(e+1)$-го элемента.

Итак, существуют числа (скаляры) $k_0, \ldots, k_e$ из P_{i-1}, для которых выполняется

$$k_0 \cdot 1 + \ldots + k_e \cdot t^e = 0.$$

Поэтому $g(t) = 0$, где $g(x) \rightleftarrows k_0 + \ldots + k_e \cdot x^e$ — многочлен над P_{i-1}.

Если многочлен g приводим над P_{i-1}, то t является корнем одного из сомножителей g. Таким образом, найдется неприводимый над P_{i-1} многочлен, корнем которого является число t. Этот неприводимый над P_{i-1} многочлен также обозначим g. Так как t корень многочлена g, то $W \cdot t$ также его корень.

По нормальности поля P_i под полем P_{i-1} корень $W(t)$ принадлежит P_i.

Итак, $W' : P_i \to P_i$, следовательно, $W' \in G(P_i, P_{i-1})$. Образно говоря, под действием W множество P_{i-1} остается на месте, а множество P_i перемешивается в себе.

Получилось отображение $h : W \mapsto W'$ вида $G(K, P_{i-1}) \to G(P_i, P_{i-1})$. Легко проверить, что оно гомоморфизм. Вычислим его ядро: $\operatorname{Ker} h = \{W \in G(K, P_{i-1}) \mid W' = \mathrm{id}\} = \{W \in G(K, P_{i-1}) \mid \forall t \in P_i (W(t) = t)\} = G(K, P_i)$. По упражнению 7 получаем: $G(K, P_i)$ нормальный делитель группы $G(K, P_{i-1})$, т. е. G_i нормальный делитель G_{i-1}. ■

По тому же упражнению 7 : $G(K, P_{i-1})/\operatorname{Ker} h = G(K, P_{i-1})/G(K, P_i) \cong R(h)$, где $R(h)$ подгруппа группы $G(P_i, P_{i-1})$.

2) Покажем, что $R(h) = G(P_i, P_{i-1})$. Тогда группа G_{i-1}/G_i изоморфна группе $G(P_i, P_{i-1})$ и, следовательно, их можно не различать.

Для определенности рассмотрим случай, когда $i = 1$. Здесь в свою очередь возможны два случая $\zeta \in P$ и $\zeta \notin P$.

Рассмотрим первый из них, когда $\zeta \in P$. Тогда $P_1 \rightleftarrows P(\zeta, \theta) = P(\theta)$, где θ один из корней $\sqrt[m]{a_0}$ и $a_0 \in R$. Фиксируем это число θ. Уравнение $z^m = a_0$, может быть, приводимо над P, однако для его корня θ найдется неприводимое над P уравнение $g(x) = 0$, для которого $g(\theta) = 0$. Оно имеет какую-то степень l, где $l \leqslant m$. По предложению 4б поле P_1 как векторное пространство l-мерно над полем P, и это обстоятельство будет использовано чуть дальше.

Фактор-множество $G(K, P)/G(K, P_1)$ содержит ровно l элементов. Действительно, два изоморфизма W и W' из $G(K, P)$ считаются эквивалентными и попадают в один класс, если $W' \cdot W^{-1} \in G(K, P_1)$. По лемме 10 из нормальности K над P вытекает, что поле K — поле разложения какого-то многочлена над P. Для уравнения $g(x) = 0$ и любого его корня θ_i, где $1 \leqslant i \leqslant l$, по предложению 6 найдется такой автоморфизм $W \in G(K, P)$, что $W(\theta) = \theta_i$. Этот

автоморфизм обозначим через W_i. Таким образом получилось l автоморфизмов $W_1, W_2, W_3, \ldots, W_l$. Они попарно не эквивалентны, так как если бы $W_i \cdot W_j^{-1} \in G(K, P_1)$, то $W_i \cdot W_j^{-1}(\theta) = \theta$ (ведь $\theta \in P_1$) и $W_i(\theta) = W_j(\theta)$, $\theta_j = \theta_i$, $W_i = W_j$. Поэтому в рассматриваемом фактор-множестве содержится l разных классов эквивалентности $[W_1], \ldots, [W_l]$.

С другой стороны, никаких иных классов эквивалентности в нем нет.

Действительно, произвольное W из $G(K, P)$ эквивалентно одному из изоморфизмов W_i, т. е. $[W] = [W_i]$, так как число $W(\theta)$ равно одному из чисел θ_i; поэтому $W(\theta) = \theta_i = W_i(\theta)$, $W_i^{-1} \circ W(\theta) = \theta$, $(W_i^{-1} \circ W) \restriction P_1 = \mathrm{id}_{P_1}$, $W_i^{-1} \circ W \in G(K, P_1)$. Это и означает, что автоморфизмы W и W_i эквивалентны. Итак, в фактор-множестве $G(K, P)/G(K, P_1)$ ровно l элементов. Следовательно, и в группе $R(h)$, биективной с этим фактор-множеством, ровно l элементов. Поскольку поле P_1 l-мерно над полем P, то по следствию к лемме 10 группа $G(P_1, P)$ состоит ровно из l элементов, и, следовательно, $R(h) = G(P_1, P)$.

Итак, случай $\zeta \in P$ полностью разобран.

Отметим, что доказательство последнего равенства не использовало никаких свойств числа θ, кроме одного: это число является корнем неприводимого уравнения $g(x) = 0$ над полем P. Воспользуемся этим, чтобы доказать такое же равенство во втором случае, когда $\zeta \notin P$. Тогда $P_1 \rightleftarrows P(\zeta, \theta)$.

Найдется такое число a, что $P(\zeta, \theta) = P(a)$ и a корень уравнения неприводимого над полем P. Поэтому второй случай сводится к первому. Убедитесь в этом, повторив все рассуждения сначала. Чтобы указать такое a, можно использовать соответствующий результат из курса алгебры или решить следующее

Упражнение 13. *а) Пусть a_1 — корень неприводимого над полем P уравнения $f_1(x) = 0$ и a_2 — корень другого неприводимого над полем $P(a_1)$ уравнения $f_2(x) = 0$. Пусть $\beta_1 = a_1, \beta_2, \ldots, \beta_n$ — все корни уравнения $f_1(x) = 0$ и $\gamma_1 = a_2, \gamma_2, \ldots, \gamma_m$ — все корни уравнения $f_2(x) = 0$.*

Наконец, пусть рациональное число c отлично от всех чисел вида $\frac{\beta_i - \beta_1}{\gamma_1 - \gamma_j}$, где $i = 1, 2, \ldots, n$ и $j = 2, \ldots, m$ и $a \rightleftarrows a_1 + c \cdot a_2$. Докажите, что $P(a_1, a_2) = P(a)$.

б) Построенное число a является корнем некоторого неприводимого над полем P уравнения.

в) Далее по индукции получаем аналогичную «свертку»

$$P(a_1, a_2, a_3) = P(a_1, a_2)(a_3) = P(a)(a_3) = P(a, a_3) = P(a'),$$

где число a' образуется по числам a и a_3 точно так же, как выше число a было образовано по числам a_1 и a_2. Таким же образом поступаем для всех полей вида $P(a_1, \ldots, a_n)$.

Указание к б. $P(a_1, a_2) = P(a)$ конечномерно над полем P.

Итак, можно образовать фактор-группы G_{i-1}/G_i для всех $=1, \ldots, s$ и они отождествляются соответственно с группами

$$G(P_i, P_{i-1}). \quad \blacksquare$$

3) Следующий принципиально важный шаг нашего доказательства состоит в том, что все группы $G(P_i, P_{i-1})$ и, следовательно, все группы G_{i-1}/G_i разрешимы.

После этого получается почти то самое, что нужно доказать в теореме 6. Действительно, цепочка групп $(***)$ состоит из нормальных делителей и фактор-группы G_{i-1}/G_i разрешимы, что «почти то же самое», что они цикличны. Поэтому $(***)$ разрешимая цепочка для группы $G \rightleftarrows G(K, P)$, т. е. $G(K, P)$ разрешима. Группа $G(K, P)$ «почти то же самое», что группа $G(P_f, K)$. Поэтому $G(P_f, P)$ разрешима. А это и требуется доказать в теореме 6.

Разрешимость всех групп $G(P_i, P_{i-1})$ проверяется, конечно, одинаково. Для определенности мы будем рассматривать случай, когда $i=1$, т. е. докажем, что группа $G(P_1, P)$ разрешима. Для этого мы собираемся вложить группу $G(P_1, P)$ в одну конкретную разрешимую группу, которую ниже определим и обозначим G_n. Иными словами, $G(P_1, P)$ окажется подгруппой разрешимой группы G_n, и тогда можно будет воспользоваться следующей леммой.

Лемма 12. ***Любая подгруппа и любая фактор-группа разрешимой группы сама является разрешимой группой.***

Сейчас, когда доказательство теоремы 6 должно в целом просматриваться, полезно еще раз обратить внимание на следующий общий план этого доказательства:

1) Даны многочлен f над полем P и радикальная цепочка $P \subset P_1' \subset \ldots \subset P_{s'}'$, где $P_f \subseteq P_{s'}'$.

2) Получаем радикальную цепочку $P \subset P_1 \subset \ldots \subset K$, где K есть нормальное расширение поля P и $P_f \subseteq K$.

3) Поле $P_i \rightleftarrows P_{i-1}(\zeta, \sqrt[m]{a_{i-1}})$ есть поле разложения многочлена $g(x) \rightleftarrows x^m - a_{i-1}$, нормальное над полем P_{i-1}.

4) Получаем цепочку групп $G \supset G_1 \supset \ldots \supset G_{i-1} \supset G_i \supset \ldots \supset \{\mathrm{id}\}$, где G_i — нормальный делитель группы G_{i-1}.

5) $G_{i-1}/G_i \underset{h}{\cong} R(h) = G(P_i, P_{i-1})$.

6) Группа $G(P_i, P_{i-1})$ разрешима, так как существует ее вложение h_1 в разрешимую группу G_n (лемма 12).

7) Разрешимость групп G_{i-1}/G_i «все равно», что их цикличность.

8) Разрешимость группы $G(K, P)$ влечет разрешимость группы $G(P_f, P)$.

Поскольку оставшаяся часть доказательства теоремы 6, с одной стороны, уже просматривается, а с другой стороны, носит более технический характер, чем раньше, то при первом чтении ее разумно пропустить. $\blacksquare$

4) Определим группу G_n и, главное, докажем, что она разрешима.

Пусть F_n одновременно множество и кольцо, состоящее из всех

вычетов по модулю натурального числа n, т. е. $F_n \rightleftarrows \mathbf{Z}/\sim_1$, где отношение эквивалентности $k_1 \sim_1 k_2$ определяется условием, что остатки от деления чисел k_1 и k_2 на число n (в кольце $\mathbf{Z}$) равны. Вместо $k_1 \sim k_2$ часто пишут $k_1 = k_2 (\mathrm{mod}\, n)$. Элементы множества F_n обозначаются как обычно $[k]$ или иногда k. Кольцевые операции в множестве F_n определяются как обычно:

$$[k_1]+[k_2] \rightleftarrows [k_1+k_2],\ [k_1]\cdot[k_2] \rightleftarrows [k_1 \cdot k_2],\ 0 \rightleftarrows [0].$$

Легко проверить, что такая система действительно является кольцом.

Упражнение 14. *Докажите, что элемент $[k]$ из кольца F_n обратим в нем тогда и только тогда, когда число k взаимно просто с числом n (в кольце $\mathbf{Z}$).*

Обозначим K' множество, состоящее из всех обратимых элементов кольца K, и определим в нем бинарную операцию, индуцированную умножением в K. Если K тело, то $K^* = K'$. В частности, множество F_n', определяемое по кольцу F_n вида $F_n' = \{[k] \in F_n \mid k$, взаимно просто с $n\}$.

Упражнение 15. *Проверьте: если K — кольцо, то K' — группа. В частности, F_n' есть конечная абелева группа.*

Обозначим $G_n \rightleftarrows (\{k \in \mathbf{Z} \mid k$ взаимно просто с $n\} \times \mathbf{Z})/\sim_2$, где $(k, l) \sim_2 (k_1, l_1)$ означает $k \sim_1 k_1$ и $l \sim_1 l_1$. Класс эквивалентности относительно отношения эквивалентности $\sim_2$ обозначим $[k, l]$. В множестве G_n определим операции: $[k, l]\cdot[t, r] \rightleftarrows [k\cdot t, l\cdot t + r]$ и $[k, l]^{-1} \rightleftarrows \rightleftarrows [k^{-1}, k^{-1}\cdot l]$, где k^{-1} любой представитель из класса эквивалентности $[k]^{-1}$, содержащегося в F_n', а также положим $1_{G_n} \rightleftarrows [1, 0]$.

Упражнение 16. *Проверьте, что эти операции корректно определены и относительно них множество G_n является группой.*

При $n \geqslant 3$ группа G_n не абелева, так как, например, $[1, 2]\cdot[a, 0] = [a; 2a]$ и $[a, 0]\cdot[1, 2] = [a, 2]$.

Положим $\psi : G_n \to F_n'$, $\psi([k, l]) \rightleftarrows [k]$.

Упражнение 17. *Докажите, что функция ψ корректно определена и является сюръективным гомоморфизмом.*

Отсюда по упражнению 7 получаем

$$(G_n/\mathrm{Ker}\,\psi) \cong F_n'.$$

Ядро функции ψ равно $\mathrm{Ker}\,\psi = \{[k, l] \mid \psi([k, l]) = [1]\} = \{[k, l] \mid [k] = [1]\} = \{[k, l] \mid k \sim_1 1\} = \{[1, l] \mid l \in \mathbf{Z}\}$.

Положим $\psi_1 : \mathrm{Ker}\,\psi \to F_n$, $\psi_1([1, l]) \rightleftarrows [l]$. Функция ψ_1 является изоморфизмом между подгруппой $\mathrm{Ker}\,\psi$ группы G_n и аддитивной группой кольца F_n, т. е. $\mathrm{Ker}\,\psi \cong F_n^+$. Группа F_n^+ циклическая с образующей $[1]$. Поэтому $\mathrm{Ker}\,\psi$ также циклическая группа.

Итак, мы нашли цепочку групп $G_n \supset \mathrm{Ker}\,\psi \supset \{[1, 0]\}$ $(*)$ для группы G_n. При этом $\mathrm{Ker}\,\psi$, нормальный делитель в группе G_n и $\mathrm{Ker}\,\psi/\{[1, 0]\} = \mathrm{Ker}\,\psi \cong F_n^+$ — циклическая группа, и также $G_n/\mathrm{Ker}\,\psi \cong F_n'$ — конечная абелева группа, и, следовательно, к ней

применима ниже указанная лемма 13. Поэтому обе фактор-группы $G_n/\mathrm{Ker}\,\psi$ и $(\mathrm{Ker}\,\psi)/\{[1, 0]\}$ в цепочке групп $(*)$ разрешимы.

Лемма 13. ***Любая конечная абелева группа разрешима.***

Теперь воспользуемся следующей леммой.

Лемма 14. ***Если группа G обладает цепочкой групп***

$$G \rightleftarrows G_0 \supset G_1 \supset G_2 \supset \ldots \supset G_s = \{\mathrm{id}\},$$

в которой все G_i являются нормальными делителями предыдущих групп G_{i-1} и все фактор-группы G_{i-1}/G_i разрешимы, то G обладает разрешимой цепочкой групп и, следовательно, является разрешимой группой.

Мы приходим к тому, что группа G_n разрешима. ■

5) Следующий шаг состоит в том, чтобы вложить группу $G(P_1, P)$ в группу G_n и тем самым доказать, что группа $G(P_1, P)$ изоморфна подгруппе группы G_n. Отсюда по лемме 12 немедленно получится, что группа $G(P_1, P)$ разрешима. Тогда по той же лемме 12 и группа G/G_1 разрешима.

Аналогично разрешимы все группы вида G_{i-1}/G_i. Поэтому цепочка $(***)$ состоит из нормальных делителей (все G_i — нормальные делители G_{i-1}) и все фактор-группы G_{i-1}/G_i разрешимы. По лемме 14 это означает, что исходная группа $G(K, P)$, ради которой и проводились все эти построения, разрешима. Это почти то самое, что требуется доказать в теореме 6.

Итак, вложим группу $G(P_1, P)$ в группу вида G_n при некотором значении числа n. Напомним, что $P_1 \rightleftarrows P(\zeta, \theta)$, где $\theta = \sqrt[m]{a_0}$.

Если $W \in G(P_1, P)$, то число $W(\zeta)$ — корень уравнения $x^m = 1$, и, следовательно, существует натуральное число k, для которого $W(\zeta) = \zeta^k$ (почему?). Это число k взаимно просто с числом m. Действительно, если это не так, то существует число $t > 1$, для которого $k = t \cdot r_1$ и $m = t \cdot r_2$. Тогда $(\zeta^k)^{r_2} = \zeta^{k \cdot r_2} = \zeta_1^{r_1 m}$, так как $kr_2 = t \cdot r_1 r_2 = r_1 m$, и далее $= (\zeta^m)^{r_1} = 1^{r_1} = 1$. Поэтому число ζ^k является корнем степени r_2 из 1, где $r_2 < m$. Отсюда $\zeta = W^{-1}(\zeta^k)$ и $\zeta^{r_2} = W^{-1}((\zeta^k)^{r_2}) = W^{-1}(1) = 1$, т. е. $\zeta^{r_2} = 1$, где $r_2 < m$. Согласно определению 3 в п. 1 получаем $\zeta^{r_2} = e^{\frac{2\pi r_2}{m} i}$, т. е. $\cos \frac{2\pi}{m} r_2 = 1$, $\frac{2\pi}{m} r_2 = 2q\pi$, $\frac{r_2}{m} = q$, где $0 < \frac{r_2}{m} < 1$ и q — целое число. Получили противоречие.

Число $W(\theta)$ — корень другого уравнения $x^m = a_0$. Поэтому существует натуральное число l, для которого $W(\theta) = \zeta^l \cdot \theta$. Положим $h_1(W) \rightleftarrows [k, l]$. Ясно, что функция h_1 вида $h_1 : G(P_1, P) \to G_m$.

Упражнение 16. *Проверьте, что функция h_1 корректно определена.*

Функция h_1 — гомоморфизм: если $h_1(W_1) = [k, l]$ и $h_1(W_2) = [t, r]$, т. е. $W_1(\zeta) = \zeta^k$, $W_1(\theta) = \zeta^l \cdot \theta$, $W_2(\zeta) = \zeta^t$, $W_2(\theta) = \zeta^r \cdot \theta$, то $(W_1 \cdot W_2)(\zeta) = W_2(\zeta^k) = \zeta^{k \cdot t}$ и $(W_1 W_2)(\theta) = W_2(\zeta^l \cdot \theta) = (W_2(\zeta))^l \cdot \zeta^r \cdot \theta = \zeta^{lt+r} \cdot \theta$. Итак, $h_1(W_1 \cdot W_2) = [k \cdot t, l \cdot t + r] = h(W_1) \cdot h(W_2)$. Функция h_1

инъективна, так как $\mathrm{Ker}\, h_1 = \{\mathrm{id}\}$. Действительно, если $h_1(W) = [1, 0]$, то $W(\zeta) = \zeta^1 = \zeta$ и $W(\theta) = \zeta^0 \cdot \theta = \theta$. По предложению, 4,б отсюда вытекает, что любой элемент поля $P_1 = P(\zeta, \theta)$ неподвижен относительно W, т. е. $W = \mathrm{id}_{P_1}$. Итак, h_1 — инъективный гомоморфизм, и, следовательно, функция h_1 — изоморфизм вида $h_1 : G(P_1, P) \leftrightarrow R(h_1)$, где $R(h_1)$ — подгруппа G_m. Поэтому группа $G(P_1, P)$ изоморфна подгруппе $R(h_1)$ группы G_m с помощью изоморфизма h_1. Это и нужно было доказать.

Итак, группа $G(K, P)$ разрешима. ■

6) Нас интересует группа $G(P_f, P)$, где $P_f \subseteq K$. Как перейти от разрешимости группы $G(K, P)$ к разрешимости группы $G(P_f, P)$? Для этого применим следующую лемму.

Лемма 15. ***Если поле K — нормальное расширение поля P и поле разложения P_f некоторого многочлена над P содержится в K, то группа Галуа $G(P_f, P)$ изоморфна фактор-группе $G(K, P)/(G(K, P_f)$.***

Отсюда по лемме 12 независимо от разрешимости группы $G(K, P_f)$ получаем, что группа $G(P_f, P)$ разрешима. Итак, теорема 6 доказана. ■

Следствие. Неприводимое над полем P уравнение $f(x) = 0$ имеет корень, выражающийся в радикалах, в том и только в том случае, если оно разрешимо в радикалах.

▷ Пусть это уравнение имеет корень, выражающийся в радикалах. По определению это означает, что существует радикальное расширение P_s исходного поля P, содержащее этот корень. По лемме 2 существует нормальное радикальное расширение K поля P, содержащее P_s и, следовательно, содержащее один из корней исходного уравнения. В силу нормальности поля K над полем P оно содержит и все остальные корни неприводимого над P уравнения $f(x) = 0$. ■

§ 3. РЕШЕНИЕ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ СТЕПЕНИ, МЕНЬШЕЙ ИЛИ РАВНОЙ 4, В РАДИКАЛАХ

1. План решения в радикалах алгебраических уравнений с разрешимой группой Галуа. В этом параграфе мы изложим способ, в сущности алгоритм, решения в радикалах тех алгебраических уравнений, которые имеют корни, выражающиеся в радикалах, т. е. по теореме Галуа таких уравнений, которые имеют разрешимую группу Галуа. Конечно, этот способ не будет столь простым, как подстановка коэффициентов квадратного уравнения в формулу $\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. Но этого и нельзя ожидать; даже аналогичные формулы для уравнений 3-й и 4-й степени, т. е. формулы Кардано и Феррари, настолько не просты, что при численном решении этих уравнений на ЭВМ они обычно не используются. Наш же способ

решения алгебраических уравнений в принципе годится для любых уравнений, а не только для уравнений 3-й и 4-й степени.

Кроме того, в этом параграфе мы докажем теорему Галуа в полном объеме, а именно докажем утверждение, обратное к теореме 6.

План этого способа решения в радикалах алгебраического уравнения состоит в следующем. Пусть

$$f(x)=x^n+a_1x^{n-1}+\ldots+a_n=0 -$$

любое уравнение n-й степени над полем P с разрешимой группой Галуа. Последнее означает, что группа $G(P_f, P)$ разрешима.

На протяжении всего этого параграфа будем предполагать, что исходное поле P содержит первообразные корни из 1 всех степеней. Если в каком-то конкретном примере это не так, то всегда можно расширить поле P, добавив к нему все корни из единицы всевозможных степеней.

Разрешимость группы Галуа означает существование разрешимого ряда

$$G(P_f, P)\supset G_1\supset\ldots\supset G_s=\{id\},$$

в котором все G_i — нормальные делители групп G_{i-1} и все фактор-группы $H_{i-1}\rightleftarrows G_{i-1}/G_i$ циклические.

Будем строить корни уравнения $f(x)=0$ индукцией по переменной s. Начальный шаг соответствует случаю $s=1$, т. е. случаю, когда сама группа Галуа $G(P_f, P)$ циклическая. Индуктивный шаг состоит в переходе от $s-1$ к s, т. е. предполагая, что мы умеем по любой разрешимой группе вида $G(P_f, P)$ с разрешимым рядом длины $s-1$ построить все корни уравнения $f(x)=0$ (точнее — построить радикальное расширение поля P, содержащее поле P_f и, в частности, содержащее все корни исходного уравнения), мы должны указать способ построения для разрешимой группы вида $G(P_f, P)$ с разрешимым рядом длиной s всех корней уравнения $f(x)=0$, т. е. опять-таки построить радикальное расширение K поля P, содержащее поле P_f.

2. Разрешимость в радикалах алгебраических уравнений с циклической группой Галуа.

Теорема 9. ***Если группа Галуа уравнения $f(x)=0$ над полем P циклическая, т. е. группа $G(P_f, P)$ циклическая, то само поле P_f является радикальным расширением поля P и даже его простым радикальным расширением.***

▷ Пусть $W\in G(P_f, P)$ и W — образующая группы $G(P_f, P)$. Тогда все элементы этой группы вида $W^0, W^1, W^2, \ldots, W^{n-1}$, где n — число элементов в ней. Обозначим ζ первообразный корень из 1 степени n, т. е. комплексное число $\zeta\rightleftarrows e^{\frac{2\pi}{n}i}$.

По предположению, сделанному в п. 1, это число ζ принадлежит P.

Обозначим $K\rightleftarrows P_f$. Определим семейство функций, зависящее от числового параметра t: $(\zeta^t, k)\rightleftarrows k+\zeta^t\cdot W^1(k)+\zeta^{2t}\cdot W^2(k)+\ldots+\zeta^{(n-1)\cdot t}\cdot W^{n-1}(k)=\sum_{l=0}^{n-1}\zeta^{l\cdot t}\cdot W^l(k)$, где t — фиксированное целое число и переменная k пробегает множество K. Все функции из этого семейства функций вида

$$(\zeta^t, k): K\to K.$$

Функция (ζ^t, k) называется *t-й резольвентой Лагранжа.* Рассмотрим значение (ζ^1, k) 1-й резольвенты Лагранжа в какой-то точке k из поля K. Ясно, что $P(k)\subseteq K$. Подполю $P(k)$ соответствует подгруппа $G(K, P(k))$ группы $G(K, P)$. Так как по условию группа $G(K, P)$ циклическая, то любая ее подгруппа циклическая.

Упражнение 19. *Если G_1 — подгруппа циклической группы G и g — образующая группы G, то элемент g^d, где d — число элементов в группе G/G_1, является образующей в группе G_1. Число d называется индексом подгруппы G_1 в группе G.*

Согласно упражнению 19 элемент W^d, где d — индекс подгруппы $G(K, P(k))$ в группе $G(K, P)$, является образующей в группе $G(K, P(k))$. В частности, $W^d\in G(K, P(k))$. Поэтому $W^d(k)=k$ и $W^{i\cdot d+j}(k)=W^i\circ(W^d\circ\ldots\circ W^d)(k)=W^j(k)$ для любых чисел i и j.

Отметим, что число d классов эквивалентности в фактор-множестве $G(K, P)/(G(K, P(k))$ делит общее число n элементов во множестве $G(K, P)$, так как $m\rightleftarrows\frac{n}{d}$ есть число элементов в любом классе эквивалентности. Любое натуральное число l представимо в виде $l=i\cdot d+j$, где $0\leqslant j\leqslant d-1$ и i, j — натуральные числа.

Тогда $(\zeta, k)=\sum_{l=0}^{n-1}\zeta^l\cdot W^l(k)=\sum_{i=0}^{m-1}\sum_{j=0}^{d-1}\zeta^{i\cdot d+j}\cdot W^{i\cdot d+j}(k)=$

$=\sum_{i=0}^{m-1}\sum_{j=0}^{d-1}\zeta^{i\cdot d+j}\cdot W^j(k)=\sum_{j=0}^{d-1}\zeta^j\cdot W^j(k)\cdot\sum_{i=0}^{m-1}\zeta^{i\cdot d}$. Если $d<n$, то $\zeta^d\neq 1$, и поэтому $\sum_{i=0}^{m-1}\zeta^{i\cdot d}=1+\zeta^d+\ldots+\zeta^{(m-1)\cdot d}=\frac{1-\zeta^{m\cdot d}}{1-\zeta^d}=\frac{1-\zeta^n}{1-\zeta^d}=$ $=0$, так как $\zeta^n=1$.

Следовательно, условие $d<n$ влечет $(\zeta^1, k)=0$, поэтому условие $(\zeta, k)\neq 0$ влечет $d=n$, так как $d\leqslant n$. Но равенство $d=n$ означает, что каждый класс эквивалентности состоит из одного элемента, и, в частности, группа $G(K, P(k))$ состоит из одного элемента, т. е. $G(K, P(k))=\{\mathrm{id}\}$.

Отсюда вытекает, что $P(k)=K$. Действительно, по ниже указанной лемме 16 поле K является полем разложения какого-то неприводимого многочлена g над полем $P(k)$. Поэтому $K=P(k)(y_1, \ldots, y_s)$, где $y_1, \ldots, y_s$ — различные корни уравнения $g(x)=0$.

Допустим, $s\geqslant 2$. По предложению 6 найдется автоморфизм W из $G(K, P(k))$, для которого $W(y_1)=y_2$. Такое W, конечно, не может быть равно id. Получили противоречие. Поэтому $s=1$, т. е. многочлен g первой степени (почему?), и, следовательно, $k_1\cdot y_1+k_2=$

$=0$, где k_1, $k_2 \in P(k)$. Отсюда $y_1 = -k_1^{-1} \cdot k_2$ и $P(k)(y_1) = P(k)$, $P(k) = K$.

Лемма 16. ***Если поле K является полем разложения какого-то многочлена над полем P и $P \subset L \subset K$, то поле K является полем разложения и над полем L.***

▷ Поле K имеет вид $K = P(x_1, \ldots, x_n)$, где $x_1, \ldots, x_n$ — все корни какого-то многочлена f над полем P. По упражнению 13 поле K можно представить в виде $K = P(a)$, где a — корень какого-то неприводимого многочлена $g(x)$ над полем P. Это $g(x)$ является многочленом и над полем L. Очевидно, выполняется $K = P(a) \subseteq L(a) \subseteq K$, поэтому $L(a) = K$.

По лемме 10 из § 2 п. 7 поле K нормально над полем P и, следовательно, содержит все корни уравнения $g(x) = 0$. Так как $L(a)$ — наименьшее поле над L, содержащее один из корней уравнения $g(x) = 0$, а именно корень a, то $L(a)$ — поле разложения многочлена $g(x)$ над полем L. ■

Итак,

$$\forall k \in K((\zeta^1, k) \neq 0 \Rightarrow K \rightleftarrows P_f = P(k)). \qquad (*)$$

Дальнейший план доказательства таков: найдем число k_0 из поля P_f, для которого

а) $k_1 \rightleftarrows (\zeta^1, k_0) \neq 0$,

и затем убедимся, что

б) $P_f = P(k_0)$ совпадает с $P(k_1)$ и k_1 — значение радикала какой-то степени из какого-то элемента поля P.

Отсюда будет следовать, что P_f — простое радикальное расширение P, что и требуется доказать.

Начнем с п. а. Поле P_f по определению вида $P_f \rightleftarrows P(x_1, \ldots, x_s)$, где $x_1, \ldots, x_s$ — корни уравнения $f(x) = 0$. По упражнению 13 найдется такое число a, что $P_f \rightleftarrows P(x_1, \ldots, x_s) = P(a)$.

Докажем, что значение 1-й резольвенты Лагранжа хотя бы на одном из чисел $a, a^2, \ldots, a^{n-1}$ отлично от нуля, т. е. хотя бы одно из чисел $(\zeta, a), \ldots, (\zeta, a^{n-1})$ не равно нулю. Допустим противное.

Тогда

$$a + \zeta \cdot W(a) + \ldots + \zeta^{n-1} \cdot W^{n-1}(a) = 0,$$
$$a^2 + \zeta \cdot W(a^2) + \ldots + \zeta^{n-1} \cdot W^{n-1}(a^2) = 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a^{n-1} + \zeta \cdot W(a^{n-1}) + \ldots + \zeta^{n-1} \cdot W^{n-1}(a^{n-1}) = 0.$$

Так как $(\zeta, 1) = 1 + \zeta + \ldots + \zeta^{n-1} = 0$, то столбцы следующего определителя линейно зависимы:

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ a & W(a) & \ldots & W^{n-1}(a) \\ a^2 & W(a^2) & \ldots & W^{n-1}(a^2) \\ \ldots & \ldots & \ldots & \ldots \\ a^{n-1} & W(a^{n-1}) & \ldots & W^{n-1}(a^{n-1}) \end{vmatrix}.$$

Следовательно, он равен нулю. С другой стороны, мы сейчас

покажем, что он не равен нулю, и это приведет к противоречию. Так как W — гомоморфизм, то указанный выше определитель равен следующему определителю размером $n \times n$:

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ a & W(a) & \ldots & W^{n-1}(a) \\ a^2 & (W(a))^2 & \ldots & (W^{n-1}(a))^2 \\ \ldots & \ldots & \ldots & \ldots \\ a^{n-1} & (W(a))^{n-1} & \ldots & (W^{n-1}(a))^{n-1} \end{vmatrix}$$

Как известно из курса алгебры, этот определитель (называемый определителем Вандермонда)[1] равен произведению всевозможных разностей чисел $a, W(a), \ldots, W^{n-1}(a)$. Но никакие два числа из этой последовательности чисел не равны между собой.

Действительно, если $W^i(a) = W^j(a)$, где $0 \leqslant i < j \leqslant n-1$, то $W^{j-i}(a) = a$.

Поле K конечномерно над полем P. Любой элемент поля K, в частности элемент a, является корнем некоторого алгебраического уравнения над P просто потому, что числа $a^0, a^1, \ldots, a^m$ не могут быть линейно независимы при всех m (иначе K над P было бы бесконечномерно). По предложению 4,б любой элемент k из поля $P(a)$ является значением многочлена над P в точке a, поэтому $W^{j-i}(k) = k$.

Следовательно,

$$\forall k \in K(W^{j-i}(k) = k),\ W^{j-i} = \mathrm{id},$$

где $j - i \leqslant n-1$, поэтому число элементов в группе $G(P_f, P) = \{W^0 \rightleftarrows E, W^1, \ldots, W^{n-1}\}$ меньше или равно $n-1$, что противоречит определению числа n.

Итак, п. а доказан, т. е. найдется такое число k_0, для которого число $k_1 \rightleftarrows (\zeta^1, k_0) \neq 0$ (и следовательно, $P_f = P(k_0)$).

Перейдем к п. б и проверим, что $P_f = P(k_0)$ совпадает с $P(k_1)$ и, кроме того, число k_1^n принадлежит P. Тогда $a_0 \circ k_1^n \in P$ и $P_f = P(\sqrt[n]{a_0})$. Это завершит доказательство теоремы 9. Отметим, что k_0 любое число, для которого $(\zeta^1, k_0) \neq 0$.

Пусть t — любое целое число. Рассмотрим t-ю резольвенту Лагранжа (ζ^t, x) в точке k_0 и вычислим:

$$W((\zeta^t, k_0)) = W(k_0 + \zeta^t \cdot W^1(k_0) + \ldots + \zeta^{(n-1)t} \cdot W^{n-1}(k_0)) = W(k_0) + \zeta^t \cdot W^2(k_0) + \ldots + \zeta^{(n-2)t} \cdot W^{n-1}(k_0) + \zeta^{(n-1)t} W^n(k_0) = W(k_0) + \zeta^t \cdot W^2(k_0) + \ldots + \zeta^{(n-2)t} \cdot W^{n-1}(k_0) + \zeta^{(n-1)t} \cdot k_0 = \zeta^{-t} \cdot (\zeta^t, k_0),$$

так как $W(\zeta^i) = \zeta^i$ (ибо $\zeta \in P$) и $W^n = \mathrm{id}$ и $\zeta^{n-1} = \zeta^{-1}$.

Итак, $W((\zeta^t, k_0)) = \zeta^{-t} \cdot (\zeta^t, k_0)$. В частности, $W((\zeta, k_0)) = \zeta^{-1} \cdot (\zeta, k_0)$. Возведем обе части этого равенства в t-ю степень и получим:

$$W((\zeta, k_0)^t) = \zeta^{-t} \cdot (\zeta, k_0)^t, \qquad (*)$$

[1] См.: *Куликов Л. Я.*, с. 471.

так как W — гомоморфизм.

Разделим предыдущее равенство на это:

$$W\left(\frac{(\zeta^t,\ k_0)}{(\zeta,\ k_0)^t}\right)=\frac{(\zeta^t,\ k_0)}{(\zeta,\ k_0)^t}.$$

Функция W и, следовательно, любая ее степень оставляют неподвижным число $\frac{(\zeta^t,\ k_0)}{(\zeta,\ k_0)^t}$.

По лемме 9 из § 2 п. 6 это означает, что число $\frac{(\zeta^t,\ k_0)}{(\zeta,\ k_0)^t}$ принадлежит полю P, так как $G(K,\ P)=\{W^0,\ W^1,\ \ldots,\ W^{n-1}\}$. Обозначим это число c_t. Итак, $\forall t\in\mathbf{Z}\ (c_t\in P\wedge(\zeta^t,\ k_0)=c_t\cdot(\zeta,\ k_0)^t)$ и все числа вида $(\zeta^t,\ k_0)$, где $t\in\mathbf{Z}$, принадлежат полю $P(k_1)$. Вычислим сумму резольвент Лагранжа:

$$\sum_{t=0}^{n-1}(\zeta^t,\ k_0)=\sum_{t=0}^{n-1}\sum_{l=0}^{n-1}\zeta^{l\cdot t}\cdot W^l(k_0)=\sum_{l=0}^{n-1}W^l(k_0)\cdot\sum_{t=0}^{n-1}\zeta^{l\cdot t}.$$

Если $l\neq 0$, то $\zeta^l\neq 1$ и $\sum_{t=0}^{n-1}\zeta^{l\cdot t}=\frac{\zeta^{l\cdot n}-1}{\zeta^l-1}=0$, так как $\zeta^n=1$, и

$\sum_{t=0}^{n-1}(\zeta^t,\ k_0)=n\cdot k_0$.

Замечательная формула! Поскольку все числа вида $(\zeta^t,\ k_0)$ принадлежат $P(k_1)$, то k_0 из $P(k_1)$ и $P(k_0)\subseteq P(k_1)$. Но $K=P(k_0)$, поэтому $K\subseteq P(k_1)$. Так как $k_1\in K$, то $P(k_1)\subseteq K$ и $K\rightleftarrows P_f=P(k_1)$.

В формулу $(*)$ подставим $t=n$ и получим $W(k_1^n)=k_1^n$. Поэтому любая степень W и, следовательно, все элементы группы $G(K,\ P)$ оставляют число k_1^n неподвижным. По лемме 9 из § 2 получаем, что $k_1^n\in P$. Следовательно, число k_1 является корнем уравнения над полем P вида $x^n=a_0$, где $a_0\rightleftarrows k_1^n$. ■

3. Разрешимость в радикалах квадратного уравнения. Прежде чем двигаться дальше по плану, намеченному в п. 1 (прочтите еще раз), применим теорему 9 к случаю квадратного уравнения.

Пусть $f(x)=x^2+a_1\cdot x+a_2$ и $a_1,\ a_2\in P$. Поле разложения этого многочлена вида $P(x_1,\ x_2)$, где $x_1,\ x_2$ — два корня уравнения $f(x)=0$.

Пусть многочлен f неприводим над полем P, и, следовательно, $x_1\neq x_2$ (иначе уравнение $f(x)=0$ сводится к линейному уравнению над P и решается тривиально). Группа Галуа этого уравнения $G(P_f,\ P)$ совпадает с симметрической группой второго порядка S_2 (см. с. 233), т. е. $G(P_f,\ P)=\{E,\ W\}$, где $W(x_1)=x_2$. Эта группа циклическая. Поэтому можно применить предыдущую теорему 9. Будем строить корни этого уравнения в соответствии с ходом доказательства теоремы 9. Обозначим $K\rightleftarrows P_f$. Поскольку $G(K,\ P)\cong S_2$, то $n=2!=2$. Первообразный корень 2-й степени из единицы является числом (-1), т. е. $\zeta=-1$. Число ζ принадлежит P.

Рассмотрим t-ю резольвенту Лагранжа:

$$((-1)^t,\ k)=k+(-1)^t\cdot W(k).$$

Пусть $k=x_1$ и $t=1$ или $t=2$. Тогда

$$\begin{cases}((-1)^1,\ x_1)=x_1-x_2\in K,\\ ((-1)^2,\ x_1)=x_1+x_2\in K.\end{cases}$$

Единственным неявным шагом в доказательстве теоремы 9 является нахождение числа k_0: для этого в теореме применялось упражнение 13. Конечно, в конкретных примерах можно найти число k_0 подбором. Например, в нашем случае за k_0 можно принять $k_0'=x_1$, и тогда $(\zeta^1,\ k_0')=x_1-x_2\neq 0$. Следовательно, за k_1' нужно принять число $k_1'=(\zeta^1,\ k_0')=x_1-x_2$. Однако для лучшего понимания доказательства теоремы 9 найдем число k_0 таким же путем, как и в этом доказательстве, т. е. используем упражнение 13. Число x_1, как и число x_2, имеет неприводимым многочленом над P сам исходный неприводимый многочлен $f(x)$. Нужно выписать все его корни, один раз начиная с x_1, а другой раз начиная с x_2, образовать числа $\frac{\beta_i-\beta_1}{\gamma_1-\gamma_j}$, где $i=1,2$ и $j=2$. А именно

$$\begin{cases}x_1,\ x_2\\ x_2,\ x_1\end{cases}\text{ и }\frac{x_1-x_1}{x_2-x_1}=0,\ \frac{x_2-x_1}{x_2-x_1}=1.$$

Выберем рациональное число c, отличное от чисел 0 и 1. Например, $c=-1$. Тогда $a=x_1+c\cdot x_2=x_1-x_2$. Случайно получилось значение 1-й резольвенты $((-1)^1,\ x_1)$. Найдем все числа вида $(\zeta,\ a)$, $(\zeta,\ a^2)$, ..., ..., $(\zeta,\ a^{n-1})$, но $n=2$, и в этой последовательности чисел ровно одно число $(\zeta,\ a)=(-1,\ x_1-x_2)=x_1-x_2+(-1)^1\cdot W(x_1-x_2)=x_1-x_2-(x_2-x_1)=2(x_1-x_2)\neq 0$, как и должно быть по общей теории. Итак, $k_0\rightleftarrows x_1-x_2$ и $k_1=2(x_1-x_2)$.

В соответствии с доказательством теоремы 9 (впрочем, это ясно и непосредственно):

$$K\rightleftarrows P_f=P(a)=P(k_0)=P(k_1)=P(2(x_1-x_2)).$$

Далее $k_1^n=4(x_1-x_2)^2$.

Обозначим $c\rightleftarrows 4(x_1-x_2)^2$.

По общей теории (что легко проверить и непосредственно) $c\in P$. Итак, $2(x_1-x_2)=\frac{1}{2}\sqrt{c}$ и $K\rightleftarrows P_f=P\sqrt[2]{c})$, где $c\in P$, что и означает построение с помощью рациональных действий и радикалов корней $x_1,\ x_2$ исходного уравнения.

Конечно, мы не получим обычной формулы для решения квадратного уравнения, но теория Галуа не ориентируется на это. Впрочем, эти формулы легко получить. Из теоремы о симметрических многочленах[1] вытекает, что $x_1+x_2=-a_1$ и $x_1\cdot x_2=a_2$ (теорема Виета). Обозначим $x_1-x_2=k_1'$. Отсюда $x_{1,\,2}=\frac{-a_1\pm k_1'}{2}$

[1] См.: *Куликов Л. Я.*, с. 232.

$$= x_1^2 + x_2^2 - 2x_1 \cdot x_2 = (x_1 + x^2)^2 - 4x_1x_2 = a_1^2 - 4a_2 \quad \text{и} \quad x_{1,2} =$$
$$= \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2}}{2} = -\frac{a_1}{2} \pm \sqrt{\left(\frac{a_1}{2}\right)^2 - a_2}.$$

4. Разрешимость в радикалах алгебраических уравнений с разрешимой группой Галуа. В этом пункте рассмотрим индуктивный шаг из того плана, который был изложен в п. 1.

Теорема 10. ***Если группа Галуа уравнения $f(x)=0$ над полем P разрешима, т. е. группа $G(P_f, P)$ разрешима, то существует радикальное расширение поля P, содержащее поле разложения P_f.***

Теоремы 6, 10 вместе составляют теорему Галуа, которая является основой теории Галуа.

▷ В соответствии с п. 1 и учитывая теорему 7, нам осталось проверить индуктивный шаг.

Пусть $K \rightleftarrows P_f$ и группа $G(K, P)$ имеет разрешимый ряд вида $G(K, P) \supset G_1 \supset \ldots \supset G_s \rightleftarrows \{E\}$. Рассмотрим подполе поля K (соответствующее группе G_1): $L_{G_1} \rightleftarrows \{k \in K \mid \forall W \in G_1 (W(k) = k)\}$ (проверьте, что это действительно подполе).

Лемма 17. ***Пусть K — поле разложения какого-то многочлена над полем P и G_1 — нормальный делитель группы Галуа $G(K, P)$. Пусть $L \rightleftarrows L_{G_1} \rightleftarrows \{k \in K \mid \forall W \in G_1 (W(k) = k)\}$. Тогда L_{G_1} является полем разложения некоторого многочлена над полем P.***

▷ Если $W \in G_1$ и $W_0 \in G(K, P)$, то по определению нормального делителя выполняется $W_0^{-1} W W_0 \in G_1$. Поэтому $\forall k \in l_{G_1} (W_0^{-1} W W_0(k) = k,\ W W_0(k) = W_0(k))$, т. е. число $W_0(k)$ неподвижно относительно всех автоморфизмов W из группы G_1. По определению поля L_{G_1} получаем $W_0(k) \in L_{G_1}$. Если k корень какого-то неприводимого над P уравнения, принадлежащий L_{G_1}, то для любого другого корня k_1 этого уравнения по лемме 10 из пункта 7 § 2 получаем $k_1 \in K$ и по предложению 6 найдется автоморфизм W_0 из группы $G(K, P)$, для которого $W_0(k) = k_1$. Поэтому $k_1 \in L_{G_1}$. Поле K над полем P конечномерно (по лемме 10 из п. 7), тем более поле L_{G_1} над полем P конечномерно (почему?).

Итак, мы доказали, что:

1) для любого неприводимого над полем P уравнения, если поле L_{G_1} содержит один его корень, то L_{G_1} содержит все его корни;

2) поле L_{G_1} над полем P конечномерно.

Эти два свойства означают, что поле L_{G_1} — нормальное расширение поля P (см. определение на с. 254).

По лемме 10 получаем, что L_{G_1} — поле разложения некоторого многочлена над полем P. ■ По леммам 10 и 15 выполняется

$$G(L_{G_1}, P) \cong (G(K, P)/G(K, L_{G_1})).$$

По лемме 18 из приложения 3 $G(K, L_{G_1}) = G_1$. Поэтому $G(L_{G_1}, P) = G(K, P)/G_1$, по условию данной теоремы 10 это означает, что группа $G(L_{G_1}, P)$ циклическая. Следовательно, по теореме 9 поле L_{G_1} является простым радикальным расширением поля P.

Группа $G(K, L_{G_1}) = G_1$ не только разрешимая, в отличие от исходной группы $G(K, P)$ она имеет разрешимый ряд длины $s-1$; а именно, таковым является часть исходного ряда от группы G_1 до группы G_s. Поэтому к полю K над полем L_{G_1} применим индуктивный шаг нашего доказательства. Это означает, что поле K содержится в радикальном расширении поля L_{G_1}, которое мы обозначим $\bar{K}$; поле $\bar{K}$ и является искомым.

Действительно, L_{G_1} — радикальное расширение P и $\bar{K}$ — радикальное расширение L_{G_1}, поэтому поле $\bar{K}$ — радикальное расширение поля P, при этом $\bar{K}$ содержит $K \rightleftarrows P_f$, что и требовалось доказать. ■

5. Разрешимость в радикалах кубического уравнения. Применим теорему 10 к уравнению 3-й степени. Рассмотрим неприводимое уравнение 3-й степени

$$f(x) = x^3 + px + q = 0$$

над полем P. Конечно, $p, q \in P$. Легко проверить прямым подсчетом, что произвольное уравнение 3-й степени $y^3 + a_1y^2 + a_2y + a_3 = 0$ заменой переменной $x = y + \frac{a_1}{3}$ приводится к виду $f(x) = 0$. Корни x_1, x_2, x_3 уравнения $f(x) = 0$ в силу его неприводимости различны, и его поле разложения P_f вида $P_f \rightleftarrows P(x_1, x_2, x_3)$.

По упражнению 11 группа Галуа этого уравнения совпадает с S_3, если $\sqrt{-4p^3 - 27q^2} \notin P$, и совпадает с A_3, если $\sqrt{-4p^3 - 27q^2} \in P$.

Для определенности рассмотрим первый случай. Тогда, в частности, $G(P_f, P)$ состоит из $3! = 6$ элементов. Как мы знаем, эта группа разрешима, но не циклична. Ее разрешающий ряд вида

$$S_3 \supset A_3 \supset \{e\}$$

(см. с. 238). Будем строить корни этого уравнения в соответствии с ходом доказательства теоремы 10. Обозначим $K \rightleftarrows P_f$. В нашем случае $G_1 = A_3$ и группе G_1 соответствует промежуточное поле $L \rightleftarrows L_{G_1} \rightleftarrows \{k \in K \mid \forall W \in G_1 (W(k) = k)\}$. Группа G_1 конечна, циклична с каким-то образующим элементом W, ее легко описать:

$$G_1 = A_3 = \left\{ W^0 = e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix},\ W \rightleftarrows \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},\ W^2 \rightleftarrows \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}.$$

Поле L бесконечно, и его нельзя описать столь явным образом. Однако по предложению 4,в $K = \{g(x_1, x_2, x_3) \mid g(u, v, w) \in P[u, v, w]\}$, и, следовательно, $L = \{g(x_1, x_2, x_3) \mid g \in P[u, v, w] \wedge g(x_1, x_2, x_3) = g(x_2, x_3, x_1) = g(x_3, x_1, x_2)\}$.

Ясно, что $P \subseteq L \subseteq K$. Более тонким моментом являются неравенства: $P \neq L$ и $L \neq K$. Действительно, число $a_0 \rightleftarrows (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$ принадлежит L, так как $W \cdot a_0 = a_0$, $W^2 \cdot a_0 = a_0$, но не принадлежит P, так как, например, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \cdot a_0 \neq a_0$. Число x_1

принадлежит K, но не принадлежит L, так как, например, $W \cdot x_1 \neq x_1$.

Итак, $P \subset L \subset K$.

Ясно, что $G_1 \subseteq G(K, L) \subseteq G(K, P)$. Поэтому $G(K, L) = G_1$ или $G(K, L) = G(K, P)$. Второй случай невозможен, так как, например, число a_0 не неподвижно относительно биекции $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ из $G(K, P)$. Итак, $G(K, L) = G_1$. То же самое, т. е. $G(K, L) = G_1$, получается по лемме 18.

По лемме 15 получаем $G(L, P) = G(K, P)/G(K, L) = S_3/A_3 = \left\{E, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}\right\} = Z_2$, т. е. группа Галуа поля L над полем P циклична.

Так как $G(K, L)$ также циклическая группа, то ее разрешимый ряд длины 1, т. е. индуктивный спуск, о котором говорится в п. 1 и доказательстве теоремы 10, закончился. Иными словами, к полю K над полем L и к полю L над полем P применима теорема 9, соответствующая начальному шагу индукции.

Сначала рассмотрим поле K над полем L. При этом $n = 3$, так как группа $G(K, L)$ содержит три элемента, и ξ первообразный корень третьей степени из 1, а именно, $\xi = e^{\frac{2\pi}{3} i} = \cos\frac{2\pi}{3} + i \sin\frac{2\pi}{3} = -\frac{1}{2} + i \cdot \frac{\sqrt{3}}{2}$.

По нашему общему предположению в п. 1 число ξ принадлежит P и, следовательно, L.

Согласно доказательству теоремы 9 нужно представить K в виде $K = L(a)$, где a, конечно, можно найти в точном соответствии с доказательством теоремы 9 по лемме 16 и упражнению 13. Но можно найти a, основываясь на том, что это число должно быть не инвариантно относительно действия группы $G(K, L) = A_3$; действительно, если a инвариантно относительно $G(K, L)$, то по лемме 9 выполняется $a \in L$ и, значит, не выполняется $L(a) = K$. Именно так, на основе такой неинвариантности чуть выше проверялось, что число $a \neq x_1$ не принадлежит L. Далее $L(a) = K$, так как если $L(a) \subset K$, то $\dim_{L(a)} K \geqslant 2$. Значит, по следствию к лемме 10 выполняется $|G(K, L)| > |G(K, L(a)| \geqslant 2$, и так как $\{E\} \subseteq G(K, L(a)) \subset G(K, L)$, то получаем $\{E\} \neq G(K, L(a)) \neq A_3$ — противоречие.

Найдем числа k_0 и k_1. Для этого нужно вычислить 1-ю резольвенту Лагранжа в точках a и a^2:

$$(\zeta, a) = x_1 + \zeta^1 \cdot W^1(x_1) + \zeta^2 \cdot W^2(x_1) = x_1 + \zeta \cdot x_2 + \zeta^2 \cdot x_3 = x_1 + \zeta x_2 + \bar{\zeta} x_3,$$
$$(\zeta, a^2) = x_1^2 + \zeta^1 \cdot W^1(x_1^2) + \zeta^2 \cdot W^2(x_1^2) = x_1^2 + \zeta x_2^2 + \bar{\zeta} \cdot x_3^2,$$

так как $\zeta \cdot \zeta^2 = \zeta^3 = 1$ и, следовательно, $\bar{\zeta} = \bar{\zeta}^{-1} = \zeta$. Как проверялось в доказательстве теоремы 9, одно из этих чисел не равно 0.

Для определенности рассмотрим первую из этих двух возможностей. Тогда $k_0 = a$ и $k_1 = x_1 + \zeta \cdot x_2 + \bar{\zeta} \cdot x_3$ и по той же теореме 9:

$$K \neq P_1 = P(x_1, x_2, x_3) = L(a) = L(x_1 + \zeta \cdot x_2 + \bar{\zeta} \cdot x_3) \text{ и } c \neq k_1^3 \in L.$$

Последнее соотношение можно проверить и непосредственным возведением числа k_1 в куб. Итак, P_1 — простое радикальное расширение поля L значением радикала вида $k_1 = \sqrt[3]{c}$, где $c \in L$.

Иными словами, $K = P_1 = L(\sqrt[3]{c})$, $c \in L$. Теперь рассмотрим поле L над полем P. Его группа Галуа $G(L, P) = Z_2$ циклична, поэтому и к ней применима все та же теорема 9. Опять нужно найти число a, для которого $L = P(a)$. Его можно найти, как и в доказательстве теоремы 9, на основе леммы 16 и упражнения 13.

Можно облегчить себе задачу, найдя подбором такое число, которое инвариантно относительно G_1, но не инвариантно относительно $G(K, P) = S_3$. Чуть выше проверялось, что таким числом является a_0. Тогда $P \subset P(a_0) \subseteq L$ и, как и выше, проверяется, что $P(a_0) = L$.

В соответствии с теоремой 9 нужно найти $n \neq |G(L, P)| = 2$ и $\zeta = e^{\frac{2\pi}{2} i} = -1$ (отметим, что $\zeta \in P$ и $W = \left(\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\right)$ и вычислить 1-ю резольвенту Лагранжа в точке a_0, получим $(\zeta^1, a_0) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3) - (x_2 - x_1)(x_2 - x_3)(x_1 - x_3) = 2a_0 \neq 0$, что и должно быть по теореме 9. Поэтому $k_0 = a$ и $k_1 = 2a_0$ и $c_0 = k_1^2 \in P$.

Последнее можно проверить и непосредственным возведением числа k_1 в квадрат. Следовательно, $L = P(\sqrt{c_0})$.

Итак, получили радикальную цепочку $P \subset P(\sqrt[2]{c_0}) \subset P(\sqrt{c_0})(\sqrt[3]{c})$, где $c_0 \in P$ и $c \in P(\sqrt{c_0})$.

Эта радикальная цепочка объясняет, почему в формуле Кардано внутри находится квадратный радикал, а снаружи — кубический. Сравните структуру этой формулы (см. с. 231) с этой радикальной цепочкой.

Конечно, мы не получили на этом пути и не должны получить саму формулу Кардано.

Сделаем это сейчас. Для этого вычислим 1, 2, 3-ю резольвенты Лагранжа в точке a:

$$(\zeta, a) = x_1 + \zeta \cdot x_2 + \bar{\zeta} \cdot x_3,$$
$$(\zeta^2, a) = x_1 + \zeta^{2 \cdot 1} \cdot W^1(a) + \zeta^{2 \cdot 2} \cdot W^2(a) = x_1 + \bar{\zeta} \cdot x_2 + \zeta \cdot x_3,$$
$$(\zeta^3, a) = x_1 + \zeta^{3 \cdot 1} \cdot x_2 + \zeta^{3 \cdot 2} \cdot x_3 = x_1 + x_2 + x_3 = 0,$$

где последнее равенство получено по той же теореме о симметрических многочленах, что и в п. 3 (иначе говоря, по теореме Виета). Это равенство соответствует теореме 7. Складывая эти три равенства, получим $(\zeta, a) + (\zeta^2, a) + (\zeta^3, a) = 3x_1 + (\zeta + \bar{\zeta} + 1) \cdot x_2 + (\zeta + \bar{\zeta} + 1) \cdot x_3 = \zeta \cdot x_1$, так как $\zeta + \bar{\zeta} + 1 = -\frac{1}{2} + i\frac{\sqrt{3}}{2} - \frac{1}{2} - i\frac{\sqrt{3}}{2} + 1 = 0$. Отсюда $x_1 = \frac{1}{3}(\zeta, a) + \frac{1}{3}(\zeta^2, a)$. Итак, уже найдено значение одного корня x_1 и осталось только найти значения резольвент Лагранжа (ζ, a) и (ζ^2, a). Для этого возведем число $k_1 = (\zeta, a) = (\zeta, x_1)$ в куб и получим:

$$k_1^3 = x_1^3 + x_2^3 + x_3^3 + 3\zeta x_1^2 x_2 + 3\bar{\zeta} x_1^2 x_3 + 3\bar{\zeta} x_1 x_2^2 + 3\zeta x_2^2 x_3 + 3\zeta x_1 x_3^2 + 3\bar{\zeta} x_2 x_3^2 +$$
$$+ 6x_1x_2x_3 = (x_1^3 + x_2^3 + x_3^3) - \frac{3}{2}\left(x_1^2x_2 + x_2^2x_3 + x_1x_3^2 + x_1x_2^2 + x_2x_3^2 + \right.$$
$$\left. + x_1^2x_3\right) + 6\left(x_1x_2x_3\right) + i\frac{3\sqrt{3}}{2}(x_1^2x_2 + x_2^2x_3 + x_1x_3^2 - x_1x_2^2 - x_2x_3^2 - x_1^2x_3).$$

В скобках выделены симметрические многочлены, т. е. алгебраические инварианты группы S_3. Выразим эти симметрические многочлены через элементарные симметрические многочлены[1].

Учтем, что по теореме Виета $x_1 + x_2 + x_3 = 0$, $x_1x_2 + x_1x_3 + x_2x_3 = p$, $x_1x_2x_3 = -q$.

Получим

$$\begin{cases} x_1^3 + x_2^3 + x_3^3 = -9q, \\ x_1^2x_2 + x_2^2x_3 + x_1x_3^2 + x_1x_2^2 + x_2x_3^2 + x_1^2x_3 = 3q, \\ x_1x_2x_3 = -q. \end{cases}$$

Очевидно, $a_0 = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3) = x_1^2x_2 + x_2^2x_3 + x_1x_3 - x_1x_2^2 - x_2x_3^2 - x_1^2x_3$. Поэтому $(\zeta,\ a)^3 = (\zeta,\ x_1)^3 = -\frac{27}{2}\cdot q + i\frac{3\sqrt{3}}{2}\cdot a_0$. Аналогичные вычисления показывают, что $(\zeta^2,\ a)^3 = (\zeta^2,\ x_1)^3 = -\frac{27}{2}\cdot q - i\frac{3\sqrt{3}}{2}\cdot a_0$. Аналогичным образом и в соответствии с общей теорией получаем $a_0^2 = -4p^3 - 27q^2$.

Итак,

$$x_2 = \frac{1}{\zeta}\sqrt[3]{-\frac{27}{q} + i\frac{3\sqrt{3}}{2}\sqrt{-4p^3 - 27q^2}} + \ldots -$$

один из корней исходного уравнения. Поделив исходное уравнение на $(x - x_1)$, мы перейдем к квадратному уравнению, для которого уже известна формула решения в радикалах.

[1] См.: *Куликов Л. Я.*, с. 499.

Глава V

ЛОГИКО-МАТЕМАТИЧЕСКИЕ ОСНОВАНИЯ ПОНЯТИЯ ЧИСЛА

§ 1. ПОНЯТИЕ НАТУРАЛЬНОГО РЯДА

Термин «число» может пониматься как натуральное, целое, рациональное, вещественное и комплексное число. Понятие числа неразрывно связано с определенными операциями и отношениями в множестве чисел: сложением, вычитанием, умножением, делением чисел и порядком, метрикой, топологией в множестве чисел. Определения соответствующих числовых структур (числовых систем) начинаются с определения исходного множества натуральных чисел и исходной структуры — структуры натурального ряда.

Существуют принципиально разные подходы к определению натурального числа и натурального ряда $\boldsymbol{N}$ и соответственно разные подходы к введению операций и отношений в $\boldsymbol{N}$. Рассмотрим три принципиально разных подхода к определению натурального ряда, или, как еще говорят, три точки зрения на натуральный ряд. Каждую из них можно распространить и на другие математические объекты и таким образом на всю математику или, по крайней мере, на ту ее часть, которая тесно связана с приложениями. Это финитная, теоретико-множественная и аксиоматическая точки зрения.

1. Финитный подход к определению натурального ряда. Рассмотрим два алфавита $Al_1 = \{|\}$ и $Al_2 = \{0,\ 1\}$. Натуральным числом называется любое непустое слово в алфавите Al_1 или любое непустое слово, не начинающиеся со знака 0, в алфавите Al_2. Интуитивно слова $|||||$ и 101 соответствуют тому, что в арабской записи обозначают знаком 5.

Связь между этими двумя определениями натуральных чисел задается вычислимой функцией вида f: $\underbrace{|\ldots|}_{n} \mapsto \sigma_1 \ldots \sigma_l$, где

$$n = \sigma_1 \cdot 2^{l-1} + \ldots + \sigma_l \cdot 2^0. \qquad (*)$$

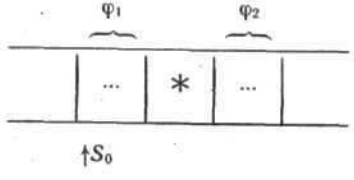
Конечно, при таком определении функции f нужно обосновать ее вычислимость, т. е. нужно построить алгоритм $\mathfrak{A}$, для которого $\forall\varphi(\varphi$ — слово в алфавите $Al_1) \Rightarrow f(\varphi) = \mathfrak{A}(\varphi)$. Обычно это делается в курсе теории алгоритмов.

Поскольку понятие машины Тьюринга и описание процесса ее работы никак не опираются на представление о натуральном числе и не используют операций над натуральными числами, то алгоритм $\mathfrak{A}$ позволяет определить функцию f без всякого упоминания об n, l, 2^l, $+$, $\cdot$, которые фигурируют в формуле $(*)$.

Итак, возможно определение функции f как алгоритма $\mathfrak{A}$, и такое определение использует только слова и машины Тьюринга.

Обозначим $l(\varphi)$ длину слова φ. Отметим, что $l(f(\varphi))\leqslant(\log_2(l(\varphi)+1)$, т. е. запись натурального числа в алфавите Al_2 существенно короче записи того же натурального числа в алфавите Al_1. Поэтому с практической точки зрения определение натурального числа с помощью алфавита Al_2 удобнее. (Действительно, $2^{l-1}+\ldots+2^0=2^l-1$, и если $n=2^l-1$, то $l=\log_2(n+1)$; если в представлении n по формуле $(*)$ встречаются нули, то l еще меньше.)

Пусть φ_1, φ_2 — слова в алфавите Al_1. Определим предикат $\varphi_1\leqslant\varphi_2$. А именно: выполняется $\varphi_1\leqslant\varphi_2$, если следующий алгоритм $\mathfrak{A}_{\leqslant}$ на входном слове вида $\varphi_1*\varphi_2$ выдает выходное слово 1, и не выполняется $\varphi_1\leqslant\varphi_2$, если тот же алгоритм $\mathfrak{A}_{\leqslant}$ на том же входном слове выдает выходное слово 0. Таким образом, алгоритм $\mathfrak{A}_{\leqslant}$ по существу является определением предиката $\leqslant$. Этот алгоритм задается машиной Тьюринга, которая перерабатывает слово вида $\varphi_1*\varphi_2$ в слово вида 1 или 0. Условимся, что все наши алгоритмы начинают работу в конфигурации:



Программа, определяющая алгоритм $\mathfrak{A}_{\leqslant}$ и, следовательно, определяющая предикат $\leqslant$, такова:

$$
\begin{array}{ll}
|S_0\to\varnothing S_1+1 & \varnothing S_3\to\varnothing S_0+1\\
|S_1\to|S_1+1 & \varnothing S_2\to 0S_4-1\\
*S_1\to *S_2+1 & *S_0\to|S_5+1\\
|S_2\to *S_3-1 & *,|S_4\to\varnothing S_4-1\\
*S_2\to *S_2+1 & *,|S_5\to\varnothing S_5+1\\
*,|S_3\to *,|S_3-1 &
\end{array}
$$

Аналогично определяется операция сложения. А именно выполняется соотношение $\varphi_1+\varphi_2=\varphi_3$, если следующий алгоритм $\mathfrak{A}_+$ на входном слове $\varphi_1*\varphi_2$ выдает выходное слово φ_3. Алгоритм $\mathfrak{A}_+$ и, следовательно, операция сложения определяются следующей программой:

$$
\begin{array}{l}
|S_0\to|S_0+1\\
*S_0\to|S_0+1\\
\varnothing S_0\to\varnothing S_1-1\\
|S_1\to\varnothing S_10
\end{array}
$$

Аналогично определяются операции $\varphi_1-\varphi_2$ (в случае, когда $\varphi_2<\varphi_1$), $\varphi_1\cdot\varphi_2$ и $\varphi_1:\varphi_2$ (в случае, когда φ_2 делит φ_1). Предикаты $\varphi_2<\varphi_1$ и «φ_2 делит φ_1» также определяются программами машин Тьюринга.

Упражнение 1. *Укажите программы для машин Тьюринга, определяющие перечисленные операции и предикаты.*

Такое понимание натурального ряда очень просто и наглядно. В сущности оно является взглядом на натуральный ряд с точки зрения ЭВМ. Такая точка зрения на натуральный ряд и другие математические объекты называется *финитной*. Ее недостатком является то, что с финитной точки зрения трудно и, может быть, даже невозможно последовательно развить всю элементарную и тем более всю вузовскую математику. Однако значительные части элементарной математики можно построить, оставаясь на финитной точке зрения.

Итак, финитная точка зрения состоит в рассмотрении математического объекта как слова и алгоритма.

2. Теоретико-множественный и аксиоматический подходы к определению натурального ряда. Второе возможное понимание натуральных чисел — теоретико-множественное. Теоретико-множественный подход в математике состоит в рассмотрении математического объекта как множества. Иными словами, в рамках этого подхода все изучаемые понятия сводятся к двум исходным понятиям: «множество» и «принадлежность», которые считаются интуитивно ясными и не определяемыми. Одноместный предикат «x — множество» иногда обозначается $M(x)$, двухместный предикат «множество x является элементом (принадлежит) множества y» обозначается $x\in y$.

Часто думают, что элементами множества являются какие-то особые, отличные от множеств объекты. С теоретико-множественной точки зрения это не так: любой элемент любого множества сам является множеством, а его элементы в свою очередь являются множествами и т. д.

Вернемся к натуральным числам. В рамках второго подхода число нуль определяется как пустое множество $\varnothing$, т. е. $0\rightleftarrows\varnothing$. Число 1 определяется как множество, состоящее из одного элемента 0, т. е. $1\rightleftarrows\{0\}$. Число 2 определяется как множество, состоящее из двух элементов 1 и 0, т. е. $2\rightleftarrows\{0, 1\}$. И аналогично дальше: $3\rightleftarrows\{0, 1, 2\}$, $4\rightleftarrows\{0, 1, 2, 3\}$, $5\rightleftarrows\{0, 1, 2, 3, 4\}$. Заметим, что при таком понимании натуральных чисел для всякого натурального числа n выполняются свойства $\forall x(x\in n\Rightarrow(x$ — натуральное число$)\wedge x<n)$;

$$n<m\Leftrightarrow n\in m;\ n\leqslant m\Leftrightarrow n\subseteq m.$$

Теперь перейдем к строгим определениям.

Определим предикат $\mathrm{Tr}(x)\rightleftarrows\forall y,\ z(y\in z\wedge z\in x\Rightarrow y\in x)$, т. е.

$$\mathrm{Tr}(x)\Leftrightarrow\forall y(y\in x\Rightarrow y\subseteq x).$$

Предикат $\mathrm{Tr}(\cdot)$ определен на семействе всех множеств, так как в качестве x можно подставить любое множество. Интуитивный смысл суждения $\mathrm{Tr}(x)$ в том, что любой элемент y из x сам состоит из элементов множества x. Например, $\mathrm{Tr}(0)$, $\mathrm{Tr}(1)$, $\mathrm{Tr}(2)$ и т. д., где множества 0, 1, 2... определены выше.

Ординалом называется такое множество x, для которого выполняется условие $(\mathrm{Tr}(x) \wedge \forall y \in x(\mathrm{Tr}(y)))$.

Интуитивный смысл этого условия в том, что элементы y множества x сами состоят из элементов множества x и, кроме того, элементы z элементов y множества x также состоят из элементов множества x. (К этому не сразу привыкаешь.)

Семейство всех ординалов называется ординальным рядом и обозначается On. Как мы увидим, ординальный ряд в значительной мере соответствует интуитивным представлениям о натуральном ряде. Чтобы в этом убедиться, рассмотрим свойства ординального ряда. Очевидно, пустое множество является ординалом. Назовем его нулем и обозначим 0. Очевидно, множество $\{\varnothing\}$ также является ординалом. Назовем его единицей и обозначим 1. Столь же легко проверить, что все указанные выше множества 2, 3, 4, ... являются ординалами.

Произвольный ординал обозначим α или β. Определим отношение $\alpha < \beta$ как $\alpha \in \beta$ и отношение $\alpha \leqslant \beta$ как

$$(\alpha < \beta) \vee (\alpha = \beta).$$

Отношение $\leqslant$ транзитивно по определению ординала.

Ординалы обладают важным свойством: $\neg(\alpha < \alpha)$. Последнее вытекает из особой аксиомы, принятой в теории множеств и называемой «аксиома фундирования»: $\forall x \neq \varnothing \exists y(y \in x \wedge \neg \exists z \in x(z \in y))$. Из нее вытекает $\neg \exists u(u \in u)$. Действительно, если такое u существует, то рассмотрим множество $x = \{u\}$ и увидим, что для него не выполняется аксиома фундирования.

Следующее свойство ординалов записывается формулой: $(\alpha \leqslant \beta) \wedge (\beta \leqslant \alpha) \Rightarrow \alpha = \beta$, т. е. структура $\langle \mathrm{On}, \leqslant \rangle$ является порядком. (Напомним, что порядком называется любое частично упорядоченное множество. Порядок называется полным, если любое непустое подмножество носителя порядка содержит наименьший элемент. Порядок называется линейным, если $\forall x, y(x \leqslant y \vee y \leqslant x)$.)

Из аксиомы фундирования тривиальным образом заключаем, что порядок $\langle \mathrm{On}, \leqslant \rangle$ является полным. Сложнее доказать, что этот порядок является линейным. Проверим последнее. Для этого обозначим $\varphi(\alpha, \beta) \rightleftharpoons (\alpha < \beta \vee \alpha = \beta \vee \beta < \alpha)$. Допустим, $\exists \alpha, \beta(\Gamma\ \varphi(\alpha, \beta))$. Пусть α_0 наименьший ординал, для которого $\exists \beta \neg\ \varphi(\alpha_0, \beta)$, и пусть β_0 наименьший ординал, для которого $\Gamma\ \varphi(\alpha_0, \beta_0)$. Проверим, что $\beta_0 \subseteq \alpha_0$. Если $\beta \in \beta_0$, то $\varphi(\alpha_0, \beta)$, т. е. $\alpha_0 < \beta \vee \alpha_0 = \beta \vee \beta < \alpha_0$. Если $\alpha_0 \leqslant \beta$, то $\alpha_0 \leqslant \beta_0$, т. е. $\varphi(\alpha_0, \beta_0)$. Поэтому $\beta < \alpha_0$, т. е. $\beta \in \alpha_0$. Если $\beta_0 = \alpha_0$, то опять-таки $\varphi(\alpha_0, \beta_0)$, поэтому $\beta_0 \subset \alpha_0$. Пусть $\beta \in (\alpha_0 \setminus \beta_0)$. Поскольку $\beta \in \alpha_0$, то $\varphi(\beta, \beta_0)$, т. е. $\beta < \beta_0 \vee \beta = \beta_0 \vee \beta_0 < \beta$. Поскольку $\beta \notin \beta_0$, то $\beta = \beta_0 \vee \beta_0 < \beta$. Так как $\beta < \alpha_0$, то $\beta_0 < \alpha_0$, т. е. $\varphi(\alpha_0, \beta_0)$. Противоречие.

Итак, доказано

Предложение 1. ***Система*** $\langle \mathrm{On}, \leqslant \rangle$ ***является полным линейным порядком с наименьшим элементом*** 0.

Предложение 1 показывает, что система ординального ряда обладает основными свойствами системы натуральных чисел.

Положим $\alpha' \rightleftharpoons (\alpha \cup \{\alpha\})$. (Иногда вместо α' пишут $\alpha + 1$).

Предложение 2. ***Для любого ординала*** α ***множество*** α' ***является ординалом и*** α' ***наименьший ординал, строго больший*** α ***(т. е.*** α' ***ординал, «непосредственно следующий» за*** α***). Кроме того, выполняется свойство:*** $\alpha' = \beta' \Rightarrow \alpha = \beta$.

▷ Очевидно, что α' ординал и $\alpha < \alpha'$. Если $\beta < \alpha'$, то $\beta \in \alpha$ или $\beta = \alpha$, т. е. $\beta \leqslant \alpha$. Поэтому предположение о существовании такого β, для которого $\beta > \alpha$ и $\beta < \alpha'$, влечет противоречие. Если $\alpha' = \beta'$ и $\alpha \neq \beta$, то $\alpha \in \beta$ и $\beta \in \alpha$, т. е. $\alpha \in \alpha$, что невозможно. ■

Для ординального ряда легко доказывается и следующее свойство, обычно называемое трансфинитной индукцией:

$$((\forall \alpha < \beta(\varphi(\alpha))) \Rightarrow \varphi(\beta)) \Rightarrow \forall \beta(\varphi(\beta)),$$

где φ любая формула.

Иногда считают, что φ произвольное свойство, не уточняя что такое «свойство». Мы будем подразумевать, что φ формула языка теории множеств, которая определяется следующим индуктивным образом: слово вида $x \in y$ называется формулой, где x и y любые непустые слова в алфавите $\{|\}$; если φ и ψ — формулы и x любое непустое слово в алфавите $\{|\}$, то слова вида $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\neg \varphi)$, $(\varphi \Rightarrow \psi)$, $(\exists x \varphi)$, $(\forall x \varphi)$ называются формулами. Таким образом, любая формула языка теории множеств (кратко говоря — формула) является словом определенного вида в алфавите $\{\wedge, \vee, \neg, \Rightarrow, \exists, \forall, \in, |, (,)\}$.

Множество всех таких слов (формул) называется языком Цермело — Френкеля (языком ZF). Непустые слова в алфавите $\{|\}$ называются переменными. Интуитивно переменную x представляют себе как произвольное множество, а запись $\varphi(x)$ — как утверждение, что множество x обладает свойством φ.

Докажем свойство трансфинитной индукции. Образуем семейство $\{\beta | \neg \varphi(\beta)\}$ и допустим, что оно непусто. Поскольку

$$\{\beta | \neg \varphi(\beta)\} = \bigcup_{\alpha} \{\beta < \alpha | \neg \varphi(\beta)\},$$

то одно из слагаемых — непустое множество и в силу полноты порядка в семействе On это множество имеет наименьший элемент. Обозначим его β_0. Итак, $\neg \varphi(\beta_0)$ и $\forall \alpha < \beta_0(\varphi(\alpha))$. По условию отсюда вытекает $\varphi(\beta_0)$ — противоречие. ■

Итак, доказано

Предложение 3. ***Для структуры*** $\langle \mathrm{On}, \leqslant \rangle$ ***выполняется свойство трансфинитной индукции.***

Аксиомами Пеано называются следующие три аксиомы, характеризующие структуру вида $\langle X, f(\cdot), x_0 \rangle$, где $f: X \to X$, $x_0 \in X$:

1) $\forall x \in X(f(x) \neq x_0)$;

2) $f(x) = f(y) \Rightarrow x = y$;

3) $(\varphi(x_0) \wedge \forall x \in X(\varphi(x) \Rightarrow \varphi(f(x)) \forall x \in X(\varphi(x))$, где $\varphi(\cdot)$ — любая формула.

Третья аксиома называется аксиомой математической индукции.

В аксиоме математической индукции фигурирует произвольная формула $\varphi(\cdot)$, например, вида $\exists x_1 \exists x_2 \forall x_3 \exists x_4 \varphi_0(x, x_1, x_2, x_3, x_4)$. Поскольку интуитивно переменные x_1, x_2, x_3, x_4 суть произвольные множества и переменная x — произвольное множество из X, а $\varphi_0(\cdot)$ выражает некоторое свойство множеств x_1, x_2, x_3, x_4, x, то, чтобы пользоваться аксиомами Пеано, нужно постулировать какие-то свойства множеств. В противном случае нельзя доказать ничего содержательного об этих $x_1, \ldots, x_4, x$. В этом смысле аксиоматика Пеано не полна, т. е. предполагает, что к ней должны быть добавлены какие-то аксиомы, характеризующие свойства переменных, иными словами, характеризующие свойства множеств, входящих в формулу $\varphi(\cdot)$. Например, в качестве таких аксиом можно выбрать все аксиомы теории Цермело — Френкеля (теории ZF).

Однако в этой главе мы не будем фактически развивать аксиоматическую систему Пеано (фактически получать следствия из ее аксиом), и поэтому для нас не важен конкретный вид этих дополнительных аксиом.

При аксиоматическом подходе натуральным рядом называется любая система вида $\langle X, f(\cdot), x_0 \rangle$, для которой выполняются аксиомы Пеано. Такую систему называют моделью аксиоматики Пеано. При аксиоматическом подходе главный вопрос состоит в том, как привести хотя бы один пример натурального ряда. Иными словами, как построить хотя бы одну модель натурального ряда.

Итак, при аксиоматическом подходе натуральный ряд понимается как произвольная модель аксиоматики Пеано. Мы не касаемся здесь формально-аксиоматического подхода, когда с аксиоматикой Пеано вообще не связывается представление о каких-либо моделях; вместо этого из аксиом Пеано и аксиом теории ZF по законам логики только лишь выводятся формальные следствия.

Является ли система $\langle On, (\cdot)', 0 \rangle$ моделью аксиом Пеано? Как проверялось выше, эта система имеет нужный вид и для нее выполняются аксиомы 1—2.

Попробуем из свойства трансфинитной индукции, которое выполняется для этой системы, вывести аксиому математической индукции. Если это удастся, то система $\langle On, (\cdot)', 0 \rangle$ окажется моделью аксиоматики Пеано. Для этого достаточно из посылки математической индукции вывести посылку трансфинитной индукции. Допустим, что $\forall \alpha < \beta(\varphi(\alpha))$, и тогда проверим $\varphi(\beta)$. Если $\beta = 0$, то свойство $\varphi(\beta)$ дано в посылке математической индукции. Если $\beta \neq 0$ и существует такой ординал $x \in On$, что $x' = \beta$, то $x < \beta$. Поэтому выполняется $\varphi(x)$, и по посылке математической индукции получаем $\varphi(x')$, т. е. $\varphi(\beta)$. Таким образом, системе On для того, чтобы являться моделью аксиоматики Пеано, не хватает ровно одного свойства:

$$\forall \beta \in \mathrm{On}(\beta \neq 0 \Rightarrow \exists x \in \mathrm{On}(x' = \beta)). \qquad (*)$$

Такое x называется ординалом, непосредственно предшествующим ординалу β, и обозначается $\beta - 1$.

Ординал α называется предельным, если он отличен от ординала 0 и $\Gamma \exists \beta(\beta' = \alpha)$. В противном случае он называется непредельным.

Предложение 4. ***Предельный ординал существует, и, следовательно, система*** $\langle \mathrm{On}, (\cdot)', 0 \rangle$ ***не является моделью аксиоматики Пеано.***

▷ Одна из аксиом теории множеств (аксиома бесконечности) утверждает $\exists X(0 \in X \wedge \forall y \in X(y' \in X))$.

Пусть Z — множество всех ординалов, принадлежащих множеству X, которое существует по аксиоме бесконечности. Пусть $\alpha = \bigcup Z$. Множество α является ординалом, так как $\forall u \in \alpha \exists v \in Z(u \in v)$, но v ординал, поэтому $u \subseteq v$ и $u \subseteq \alpha$. Кроме того, если $v \in u \in \alpha$, то $v \in \alpha$ и $v \subseteq \alpha$.

Лемма 1. ***Выполняется соотношение*** $(\beta \subseteq \alpha) \Leftrightarrow (\beta \leqslant \alpha)$.

▷ Если $\beta < \alpha$, то $\beta \in \alpha$ и по определению ординала $\beta \subseteq \alpha$. Допустим, что $\Gamma(\beta \leqslant \alpha)$, тогда $\alpha < \beta$, т. е. $\alpha \in \beta$. Но $\alpha \notin \alpha$, поэтому $\alpha \in (\beta \setminus \alpha)$ и получаем $\Gamma(\beta \subseteq \alpha)$. ■

Продолжим доказательство предложения 4. Ординал α больше всех ординалов, входящих в множество Z, так как $\beta \in Z \Rightarrow \beta \subseteq \alpha$, и можно применить лемму 1. Проверим, что α — наименьший ординал среди тех, которые $\geqslant$ всех элементов из Z. Действительно, если бы выполнялось $\forall u \in Z(u \leqslant \beta)$, то $\forall u \in Z(u \subseteq \beta)$, $(\bigcup Z) \subseteq \beta$ и $\alpha = (\bigcup Z) \leqslant \beta$.

Поскольку $0 \in X$, то $0' \leqslant \alpha$ и соотношение $0 < 0' \leqslant \alpha$ влечет $\alpha \neq 0$.

Предположим, что $\exists \beta(\beta' = \alpha)$. Тогда $\beta < \alpha$ и β не может быть $\leqslant$ всех ординалов из Z, т. е.

$$\exists \gamma \in Z(\beta < \gamma).$$

Поэтому $\alpha = \beta' \leqslant \gamma < \gamma'$, но $\gamma' \in Z$ (по определению множеств Z и X); поэтому $\gamma' \leqslant \alpha$, т. е. $\alpha < \alpha$. Получили противоречие. ■

Итак, семейство всех предельных ординалов непусто. Поэтому в нем существует наименьший элемент. Обозначим его $\boldsymbol{N}$. Следовательно, все элементы $\boldsymbol{N}$ непредельные ординалы. Поэтому для множества $\boldsymbol{N}$ выполняется свойство $(*)$.

Сузим операцию $(\cdot)'$ на такое множество $\boldsymbol{N}$. Конечно, $0 \in \boldsymbol{N}$ и условие $\alpha \in \boldsymbol{N}$ влечет $\alpha' \in \boldsymbol{N}$. Аксиома математической индукции выполняется в системе $\langle \boldsymbol{N}, (\cdot)', 0 \rangle$, а все остальные аксиомы Пеано тем более выполняются для нее. Итак, доказана следующая важная теорема.

Теорема 1. ***Система*** $\langle \boldsymbol{N}, (\cdot)', 0 \rangle$ ***является моделью аксиом Пеано.***

С точки зрения аксиоматического подхода эта теорема устанавливает, что хотя бы один натуральный ряд существует, а именно таковым является система $\langle \boldsymbol{N}, (\cdot)', 0 \rangle$. Эта конкретная модель аксиоматики Пеано называется стандартным натуральным рядом.

Возможны и другие модели аксиоматики Пеано, которые даже не биективны стандартному натуральному ряду.

Аксиоматический подход предполагает (имеет своей целью) получение всевозможных следствий по законам логики из некоторой системы аксиом. Любое такое следствие истинно (выполняется) в произвольной модели исходной системы аксиом. Поэтому аксиоматический подход позволяет изучать сразу и одновременно все модели исходной системы аксиом (всю совокупность моделей исходной системы аксиом). В нашем случае он позволяет изучать сразу все модели аксиом Пеано. Если мы захотим изучать свойства, присущие только какой-то одной из этих моделей, например, стандартному натуральному ряду $\boldsymbol{N}$, то аксиоматический подход не позволит этого: все, что по законам логики вытекает из аксиом Пеано, выполняется сразу во всех моделях этих аксиом.

Интуитивно мы представляем себе натуральный ряд как какой-то один определенный объект, обычно как стандартный натуральный ряд; поэтому аксиоматический подход, быть может, не дает вполне адекватного описания идеи натурального ряда.

Бывают системы аксиом, которые имеют всего одну модель, точнее — у которых любые две модели изоморфны. Такие системы аксиом называются категоричными.

Изучение всех моделей категоричной системы аксиом равносильно изучению ее единственной модели. К сожалению, интересные и категоричные системы аксиом — большая редкость. Некатегорична и аксиоматика Пеано.

Определим операции в натуральном ряду $\boldsymbol{N}$.

Пусть x — какое-то фиксированное натуральное число из $\boldsymbol{N}$. Рассмотрим множество $X \rightleftarrows \{n \in \boldsymbol{N} \mid \exists q(l(q)=n \wedge q(1)=x \wedge \forall i<n, q(i+1)=q(i)')\}$, где $q(\cdot)$ — произвольная конечная последовательность, $l(q)$ — ее длина и $q(1)$, $q(2)$, ... $q(n)$ — все ее значения. По свойству вполне упорядоченности множества $\boldsymbol{N}$ множество X совпадает с $\boldsymbol{N}$, т. е. $X=\boldsymbol{N}$ (докажите).

Интуитивно множество X состоит из всех таких натуральных чисел n, для которых можно сделать n шагов по правилу $q(i+1)=q(i)'$.

Рассмотрим множество $Y \rightleftarrows \{n \in \boldsymbol{N} \mid \forall q,\ r(l(q))=n=l(r) \wedge q(1)=r(1) \wedge \forall i<n(q(i+1)=q(i)' \wedge r(i+1)=r(i)') \Rightarrow q(n)=r(n)\}$, где $q(\cdot)$, $r(\cdot)$ — произвольные конечные последовательности и $q(1)$, ..., $q(n)$ и $r(1)$, ..., $r(n)$ — их значения. По свойству вполне упорядоченности множества $\boldsymbol{N}$ множество Y совпадает с $\boldsymbol{N}$ (докажите). Интуитивно множество Y состоит из всех таких натуральных чисел n, для которых любые две последовательности $q(\cdot)$ и $r(\cdot)$, совпадающие вначале и образующиеся по одинаковому правилу $q(i+1)=q(i)'$ и $r(i+1)=r(i)'$, за n шагов не могут разойтись, т. е. не может случиться, что $q(n) \neq r(n)$.

Может показаться, что равенства $X=\boldsymbol{N}$ и $Y=\boldsymbol{N}$ выполняются тривиально и не зависят от аксиом натурального ряда. Это не так. Представим себе, что те же самые множества X и Y определены

для множества N_1, похожего на множество $\boldsymbol{N}$, а именно $N_1 \rightleftarrows \{0, 1, 2, 3, ..., n, ..., \omega_0\}$. Иными словами, в множестве N_1 после всех обычных натуральных чисел находится еще одно «число», которое мы обозначим ω_0. Рассмотрим две последовательности $q(\cdot)$ и $r(\cdot)$ длиной ω_0, определяемые равенствами $q(1)=2=r(1)$, $q(n+1)=q(n)'$ и $r(n+1)=r(n)'$ и $q(\omega_0)=2$, $r(\omega_0)=3$.

Последовательности $q(\cdot)$ и $r(\cdot)$ длиной ω_0, они одинаково начинаются, строятся по одинаковому правилу, а в конце расходятся, т. е. $q(\omega_0) \neq r(\omega_0)$. Поэтому $\omega_0 \in N_1$ и $\omega \notin Y$ и $Y \neq N_1$.

Представим себе, что множество N_2 имеет вид $0 \to 1 \to 2$, $\nwarrow 3 \swarrow$ где стрелка показывает непосредственно следующее число. Тогда $4 \rightleftarrows 0'''' = 1,5 \rightleftarrows 0''''' = 2$, и таким образом определяются все «числа» 6, 7, 8, ... Для такого «натурального ряда» N_2, например, $4 \notin X$, если $x=0$, так как $q(1)=x$, $q(2)=x'$, $q(3)=x''$, $q(4)=x'''$, но $4=1$ и $x \neq x'''$, что невозможно.

Положим $(x+n) \rightleftarrows z$, если

$$\exists q(l(q)=n \wedge$$

$$\wedge q(1)=x \wedge \forall i<n(q(i+1)=q(i)') \wedge q(n)=z), \qquad (*)$$

где x, n, $z \in \boldsymbol{N}$. Рассматривая упомянутые выше множества X и Y, легко показать, что для любых чисел x, $n \in \boldsymbol{N}$ такое число z существует и единственно. Поэтому условие $(*)$ действительно определяет двуместную операцию вида $\boldsymbol{N}^2 \to \boldsymbol{N}$. Аналогично определяется операция умножения.

Для этих операций легко проверяются все свойства обычных операций сложения и умножения натуральных чисел; например все свойства, выражаемые аксиомами, сформулированными в книге В. И. Нечаева «Числовые системы» (с. 53).

Будем рассматривать совокупность всех этих аксиом как аксиоматическое определение системы натуральных чисел вида $\langle X, \leqslant, +, \cdot, x_0, x_1 \rangle$, где $\leqslant$ бинарное отношение на X и $+$, $\cdot$ двуместные операции в X, а x_0, x_1 элементы X. Тогда нами доказана следующая теорема.

Теорема 2. ***Система $\langle \boldsymbol{N}, \leqslant, \cdot, 0, 1 \rangle$ является моделью аксиоматического определения системы натуральных чисел.***

Конкретная модель, о которой говорится в теореме 2, также называется стандартным натуральным рядом. Существуют модели аксиоматики системы натуральных чисел, которые не изоморфны модели из теоремы 2.

3. Сравнение определений целых чисел. Совсем кратко рассмотрим определения целых чисел. Финитное определение целого числа аналогично финитному определению натурального числа. А именно, целым числом называется пустое слово или любое слово вида $\sigma\varphi$, где φ — любое непустое слово в алфавите A_1 и σ — один

из двух знаков $+$ или $-$. Как и раньше, алгоритмически определяются операции $+$, $-$, $\cdot$, $:$ и отношение порядка $\leqslant$.

Упражнение 2. *Напишите программы для машин Тьюринга, определяющие эти операции и отношение порядка.*

Теоретико-множественное определение целого числа опирается на понятие пары. Парой называется множество X, обладающее свойством $\exists y,\ z \forall u (u \in X \Leftrightarrow u = y \vee u = z)$.

Такое множество существует и единственно, его легко указать:

$$X = \{y, z\}.$$

Упорядоченной парой называется множество X, обладающее свойством $\exists y,\ z\ \forall u\ (u \in X \Leftrightarrow u = \{y\} \vee u = \{z, y\})$. Такое множество существует и единственно, его легко указать: $X = \{\{y\}, \{y, z\}\}$. Оно обозначается $\langle y, z \rangle$.

Теперь целое число можно определить как пустое множество или упорядоченную пару вида $\langle \sigma, n \rangle$, где $\sigma \in \{+, -\}$ и $n \in \boldsymbol{N}$. Операции и отношение порядка в $\boldsymbol{N}$ легко индуцируют одноименные операции и отношение порядка в множестве всех целых чисел. Определите их самостоятельно.

Аксиоматическое определение целых чисел содержится, например, у В. И. Нечаева в книге «Числовые системы» (с. 95).

§ 2. ОПРЕДЕЛЕНИЕ РАЦИОНАЛЬНОГО ЧИСЛА КАК ЛИНЕЙНОЙ ФУНКЦИИ

Читатель может по аналогии с определениями предыдущего параграфа самостоятельно сформулировать финитное, теоретико-множественное и аксиоматическое определения рационального числа. Поэтому больше не будем обращать внимание на эту сторону дела.

Сравним два по существу разных определения рационального числа. Оба они, как и все последующее изложение, находятся в рамках теоретико-множественного подхода.

В соответствии с традицией будем в дальнейшем считать, что натуральный ряд $\boldsymbol{N}$ не содержит числа 0.

Первое определение рационального числа хорошо известно. Рациональным числом называется любой класс эквивалентности в множестве $\boldsymbol{Z} \times \boldsymbol{N}$ относительно следующего отношения эквивалентности:

$$\langle x, n \rangle \sim_1 \langle y, m \rangle \rightleftarrows x \cdot m = y \cdot n.$$

Иными словами, множеством рациональных чисел называется фактор-множество вида $(\boldsymbol{Z} \times \boldsymbol{N}) / \sim_1$. Элементы этого фактор-множества называются рациональными числами. Проверьте, что $\sim_1$ является отношением эквивалентности в множестве $\boldsymbol{Z} \times \boldsymbol{N}$. Вместо $\langle x, n \rangle$ по традиции пишут $\frac{x}{n}$. Множество рациональных чисел обозначается $\boldsymbol{Q}$. Положим:

$$\left[\frac{x}{n}\right] + \left[\frac{y}{m}\right] \rightleftarrows \left[\frac{x \cdot m + y \cdot n}{n \cdot m}\right] \text{ и } \left[\frac{x}{n}\right] \leqslant \left[\frac{y}{m}\right] \rightleftarrows (x \cdot m \leqslant y \cdot n).$$

Проверьте, что эти определения корректны, т. е. не зависят от выбора представителей в классах эквивалентности. В общем случае далеко не всегда возможно опустить операции на фактор-множество именно потому, что возникает зависимость от выбора представителя.

Легко проверить, что система $\boldsymbol{Q} = \langle \boldsymbol{Q}, +, -, \cdot, ^{-1}, 0, 1, \leqslant \rangle$ является упорядоченным полем, где

$$-\left[\frac{x}{n}\right] \rightleftarrows \left[\frac{-x}{n}\right], \left[\frac{x}{n}\right]^{-1} \rightleftarrows \left[\frac{\operatorname{sign}(x) \cdot n}{|x|}\right]$$

и

$$\operatorname{sign}(x) \rightleftarrows \begin{cases} +1, \text{ если } x \geqslant 0, \\ -1, \text{ если } x < 0. \end{cases}$$

В этой системе может вызвать недоумение сложный вид формулы, определяющей операцию сложения, причем сложение определяется сложнее, чем умножение. Причина такого вида этой формулы станет ясна из второго определения рационального числа.

При втором определении рациональным числом будем называть класс эквивалентности в множестве всех частично определенных функций вида $\boldsymbol{Z} \to \boldsymbol{Z}$. Например, число 3/2 при первом определении понимается как множество пар вида $\left\{\frac{3 \cdot n}{2 \cdot n} \,\middle|\, n \in \boldsymbol{N}\right\}$, а при втором определении — как множество функций, одна из которых вида $\frac{3}{2}(x) : \boldsymbol{Z} \to \boldsymbol{Z}$, где $\frac{3}{2}(x) \rightleftarrows \left(\frac{3}{2} \cdot x\right)$ и x кратно числу 2. Что означает здесь операция умножения $\cdot$ (обратите внимание, что запись $\frac{3}{2} \cdot x$ нельзя понимать как умножение рационального числа $\frac{3}{2}$ на целое число x, так как еще не определено не только умножение рациональных чисел, но и само понятие рационального числа — мы только собираемся определить, что такое рациональное число)? Можно ли вычислить, например, $\frac{3}{2}(5) = \frac{3}{2} \cdot 5$ и получить в результате число, принадлежащее $\boldsymbol{Z}$? Если сузить область определения функции $\frac{3}{2}(x)$ до подгруппы $\boldsymbol{Z}$ вида $D\left(\frac{3}{2}\right) \rightleftarrows \{2 \cdot k \mid k \in \boldsymbol{Z}\}$, то, во-первых, ясно, как определить операцию умножения: $\frac{3}{2}(2k) \rightleftarrows \frac{3}{2} \cdot 2k \rightleftarrows$ $\rightleftarrows 3 \cdot k$, где $3 \cdot k$ есть обычное умножение двух натуральных чисел. И во-вторых, $\frac{3}{2}(\cdot)$ имеет желательный вид $\boldsymbol{Z} \to \boldsymbol{Z}$. Ясно, что функции $\frac{3}{2}(\cdot)$ и $\frac{6}{4}(\cdot)$ с областями определения соответственно $\{2k\}$ и $\{4k\}$, где $k \in \boldsymbol{Z}$, разумно отождествить.

Перейдем к точным определениям.

Рассмотрим множество

$$X \rightleftarrows \{f : G \to Z \mid \forall g_1,\ g_2 \in G (f(g_1 + g_2) = f(g_1) + f(g_2))\},$$

где G — произвольная нетривиальная подгруппа в группе $\boldsymbol{Z}^+$. Напомним, что нетривиальной называется любая подгруппа, отличная от нулевой подгруппы $\{0\}$. В G определены не только групповые операции $+$ и $-$, но и операции умножения на элементы из $\boldsymbol{Z}$, а именно $n \cdot g \rightleftarrows \underbrace{g + \ldots + g}_{n} : G \to G$. Поэтому G — модуль над кольцом $\boldsymbol{Z}$. Если $f \in X$, то, очевидно, выполняется $\forall n \in \boldsymbol{Z}\ \forall g \in G (f(n \cdot g) = n \cdot f(g))$. Поэтому f не только гомоморфизм группы G в себя, но и $\boldsymbol{Z}$ — гомоморфизм ($\boldsymbol{Z}$ — линейная функция) модуля G в себя.

В дальнейшем используем следующее свойство произвольной подгруппы группы $\boldsymbol{Z}^+$: такая подгруппа либо тривиальна, либо имеет вид $\{n \cdot l \mid n \in \boldsymbol{Z}\}$, где l — фиксированное число > 0. Докажем это свойство.

Допустим $\exists g \in G (g \neq 0)$. Если $g < 0$, то $(-g) > 0$ и $(-g) \in G$, т. е. $\exists g \in G (g > 0)$. В множестве $\{g \in G \mid g > 0\}$ найдется наименьший элемент l. Очевидно, $\{n \cdot l \mid n \in \boldsymbol{Z}\} \subseteq G$. Если $g \in G$, то по архимедовости группы Z^+ (пусть, например, $g > 0$) найдется натуральное число n, для которого $n \cdot l \leqslant g < (n+1) \cdot l$. Если $n \cdot l \neq g$, то $(g - nl) \in G$ и $0 < g - nl < l$, что противоречит определению числа l.

(Далеко не для любого кольца K подгруппы G его аддитивной группы K^+ обладают таким свойством.)

Группа $\boldsymbol{Z}^+$ обладает еще одним важным свойством: пересечение любых двух нетривиальных подгрупп группы $\boldsymbol{Z}^+$ является снова нетривиальной подгруппой в $\boldsymbol{Z}^+$. Действительно, если G_1 и G_2 — такие подгруппы, то по доказанному свойству они имеют вид $\{n \cdot l_1 \mid n \in \boldsymbol{Z}\}$ и $\{n \cdot l_2 \mid n \in \boldsymbol{Z}\}$. Тогда $G_1 \cap G_2 = \{nl_3 \mid n \in \boldsymbol{Z}\}$, где l_3 — наименьшее общее кратное чисел l_1 и l_2.

Положим

$$(f_1 \sim f_2) \rightleftarrows (f_1 \restriction (D(f_1) \cap D(f_2)) = f_2 \restriction (D(f_1) \cap D(f_2))),$$

т. е. функции f_1 и f_2 эквивалентны, если они совпадают на пересечении их областей определения. Определим рациональное число как любой класс эквивалентности в множестве X относительно отношения эквивалентности $\sim_2$. Иными словами, определим множество рациональных чисел как фактор-множество вида $X/\sim_2$ и элементы этого фактор-множества назовем рациональными числами. Итак,

$$\boldsymbol{Q} \rightleftarrows (X/\sim_2).$$

Определим операции и отношение порядка в этом фактор-множестве:

$$[f_1] + [f_2] \rightleftarrows [f_1 \restriction G) + (f_2 \restriction G)],\ [f_1] \cdot [f_2] \rightleftarrows [f_1 \circ f_2],$$

где в правой части $+$ обозначает поточечное сложение функций и $\circ$ обозначает композицию двух функций, и

$$[f_1] \leqslant [f_2] \rightleftarrows \forall k \in G (k \geqslant 0 \Rightarrow f_1(k) \leqslant f_2(k)),$$

а $G \rightleftarrows D(f_1) \cap D(f_2)$.

Проверьте, что эти определения корректны, т. е. не зависят от выбора представителей в классах эквивалентности.

Получилось упорядоченное поле

$$\boldsymbol{Q} = \langle \boldsymbol{Q},\ +,\ -,\ \cdot,\ ^{-1},\ 0,\ 1 \rangle, \text{ где}$$

$$-[f] = [-f],\ 0 \rightleftarrows [0],\ 1 \rightleftarrows [\mathrm{id}]$$

и 0 в квадратных скобках — функция вида $\boldsymbol{Z} \to \boldsymbol{Z}$, тождественно равная 0 и $\mathrm{id}(x) \equiv x$, $x \in \boldsymbol{Z}$.

Если $D(f) = \{n \cdot l \mid n \in \boldsymbol{Z}\}$, где $l > 0$ и $f(l) = k \in \boldsymbol{Z}$, то функция f однозначно определяется парой $\langle k, l \rangle$. Действительно, область определения $D(f)$ задается образующей $l \in \boldsymbol{N}$ и функция $f(n \cdot l) = n \cdot f(l)$ определяется числом $k \in \boldsymbol{Z}$. Поэтому функцию f обозначим $f_{k,l}$. Соответствие $g : \langle k, l \rangle \mapsto f_{k,l}$ является биекцией множества $\boldsymbol{Z} \times \boldsymbol{N}$ и множества X (проверьте).

Очевидно, $\mathrm{id} = f_{1,1}$. Проверим, что $1 = [f_{1,1}]$ и $[f_{k,l}]^{-1} = [f_{\mathrm{sign}(k) \cdot l, |k|}]$. (В последней формуле знак числа k переходит на число l, делая знаменатель натуральным числом.)

Действительно,

$$[f_{k,l}] \cdot 1 = [f_{k,l}] \cdot [\mathrm{id}] = [f_{k,l} \circ \mathrm{id}],$$

функция в квадратных скобках определена только на числа вида $n \cdot l$. Поэтому

$$(f_{k,l} \cdot \mathrm{id})(n \cdot l) = k \cdot n = f_{k,l}(n \cdot l).$$

Аналогично проверим: $[f_{k,l}] \cdot [f_{k,l}]^{-1} = [f_{k,l}] \cdot [f_{l,k}] = 1$, где для простоты записи рассмотрим случай $k \geqslant 0$. Поскольку $[f_{k,l}] \neq 0$, то $k > 0$. Образ функции $f_{l,k}(\cdot)$ содержится в $D(f_{k,l})$, так как $f_{l,k}(n \cdot k) = n \cdot l$; следовательно, область определения функции $f_{k,l} \circ f_{l,k}$ совпадает с областью определения функции $f_{l,k}$. Поэтому $f_{k,l} \circ f_{l,k}(n \cdot k) = f_{k,l}(n \cdot l) = n \cdot k = \mathrm{id}(n \cdot k)$ и $[f_{k,l}] \cdot [f_{l,k}] = [\mathrm{id} \mid \{n \cdot k\}] = 1$, так как $(\mathrm{id} \restriction \{n \cdot k\}) \sim_2 \mathrm{id}$.

Предложение 5. ***Определения упорядоченного поля $\boldsymbol{Q}$ как фактор-множества $(\boldsymbol{Z} \times \boldsymbol{N})/\sim_1$ и как фактор-множества $X/\sim_2$ эквивалентны.***

▷ Нужно установить, что система $(\boldsymbol{Z} \times \boldsymbol{N})/\sim_1$ изоморфна системе $X/\sim_2$. Биекция g, которая определена выше, обладает свойством: $\frac{k_1}{l_1} \sim_1 \frac{k_2}{l_2} \Leftrightarrow f_{k_1,l_1} \sim_2 f_{k_2,l_2}$. Поэтому биекция g опускается до биекции g_1 вида $((\boldsymbol{Z} \times \boldsymbol{N})/\sim_1) \leftrightarrow (X/\sim_2)$, где $g_1\left(\left[\frac{k}{l}\right]\right) \rightleftarrows [g \langle k, l \rangle]$.

Осталось проверить, что функция g_1 — гомоморфизм, который сохраняет отношение порядка.

Для этого вычислим $[f_{k_1,\,l_1}]+[f_{k_2,\,l_2}]$.

Нужно по определению сузить функции $f_{k_1,\,l_1}$ и $f_{k_2,\,l_2}$ на пересечение их областей определения, но мы сузим их еще больше, на множество $\{n\cdot l_1l_2 \mid n\in\mathbf{Z}\}$. При этом перейдем к эквивалентным функциям, которые соответственно равны $f_{k_1l_2,\;l_1l_2}$ и $f_{k_2l_1,\;l_1l_2}$. Они имеют общую область определения. Тогда

$$(f_{k_1l_2,\;l_1l_2}+f_{k_2l_1,\;l_1l_2})(n\cdot l_1l_2)=nk_1l_2+nk_2l_1=n\cdot(k_1l_2+k_2l_1)=$$
$$=f_{k_1l_2+k_2l_1,\;l_1l_2}(n\cdot l_1l_2).$$

Итак,

$$[f_{k_1,\,l_1}]+[f_{k_2,\,l_2}]=[f_{k_1l_2+k_2l_1,\;l_1l_2}]$$

и $\left[\frac{k_1}{l_1}\right]+\left[\frac{k_2}{l_2}\right]=\left[\frac{k_1l_2+k_2l_1}{l_1l_2}\right]\mapsto[f_{k_1l_2+k_2l_1,\;l_1l_2}]=[f_{k_1,\,l_1}]+[f_{k_2,\,l_2}]$.

Наглядно это можно изобразить диаграммой:

$$\begin{array}{ccc}\left[\frac{k_1}{l_1}\right] & + & \left[\frac{k_2}{l_2}\right] & = & \left[\frac{k_1l_2+k_2l_1}{l_1l_2}\right]\\ g_1\downarrow & & g_1\downarrow & & g_1\downarrow\\ \left[f_{k_1,\,l_1}\right] & + & \left[f_{k_2,\,l_2}\right] & = & \left[f_{k_1l_2+k_2l_1,\;l_1l_2}\right]\end{array}.$$

Она означает, что $g_1(x+y)=g_1(x)+g_1(y)$.

Аналогично вычислим $[f_{k_1,\,l_1}]\cdot[f_{k_2,\,l_2}]=[f_{k_1,\,l_1}\circ f_{k_2,\,l_2}]$. Чтобы не разбирать вопрос об области определения композиции $f_{k_1,\,l_1}\circ f_{k_2,\,l_2}$, заменим функцию $f_{k_2,\,l_2}$ на эквивалентную ей функцию $f_{k_2l_1,\;l_1l_2}$.

Получаем

$$f_{k_1,\,l_1}\circ f_{k_2l_1,\;l_1l_2}(n\cdot l_1l_2)=n\cdot k_1\cdot k_2=f_{k_1k_2,\;l_1l_2}(n\cdot l_1l_2)\text{ и}$$
$$[f_{k_1,\,l_1}]\cdot[f_{k_2,\,l_2}]=[f_{k_1k_2,\;l_1l_2}].$$

Соответствующая диаграмма выглядит так:

$$\begin{array}{ccc}\left[\frac{k_1}{l_1}\right] & \cdot & \left[\frac{k_2}{l_2}\right] & = & \left[\frac{k_1\cdot k_2}{l_1\cdot l_2}\right]\\ \downarrow & & \downarrow & & \downarrow\\ \left[f_{k_1,\,l_1}\right] & \circ & \left[f_{k_2,\,l_2}\right] & = & \left[f_{k_1k_2,\;l_1l_2}\right]\end{array}.$$

Итак, функция g_1 — гомоморфизм.

Аналогично вычислим предикат порядка: $[f_{k_1,\,l_1}]\leqslant[f_{k_2,\,l_2}]\Leftrightarrow$
$\Leftrightarrow[f_{k_1l_2,\;l_1l_2}]\leqslant[f_{k_2l_1,\;l_1l_2}]\Leftrightarrow\forall n\in\mathbf{Z}_{>0}(f_{k_1l_2,\;l_1l_2}\;(n\cdot l_1l_2)\leqslant f_{k_2l_1,\;l_1l_2}(n\cdot l_1l_2))\Leftrightarrow$
$\Leftrightarrow\forall n\in\mathbf{Z}_{>0}(n\cdot k_1l_2\leqslant n\cdot k_2l_1)\Leftrightarrow k_1l_2\leqslant k_2l_1$.

Теперь очевидно, что

$$\left[\frac{k_1}{l_1}\right]\leqslant\left[\frac{k_2}{l_2}\right]\Leftrightarrow[f_{k_1,\,l_1}]\leqslant[f_{k_2,\,l_2}].$$

По поводу того, какое из этих двух определений рационального числа лучше, приведем слова Платона (который предпочитал именно второе определение): «там, где вычислитель делит, ученые умножают».

§ 3. ОСНОВНЫЕ ПОДХОДЫ К ОПРЕДЕЛЕНИЮ ВЕЩЕСТВЕННЫХ ЧИСЕЛ

Вещественные и комплексные числа рассматриваются в курсах анализа и алгебры, изучаются в курсе «Числовые системы»; в этом и следующем параграфе мы рассмотрим основные способы их определения.

Обычно пользуются одним из трех способов определения вещественного числа: по Коши, по Дедекинду и по Кантору. Каждый из этих способов имеет важное общематематическое значение.

Для введения комплексных чисел пользуются также одним из трех способов: комплексификацией векторного пространства $\boldsymbol{R}^2$ (над полем $\boldsymbol{R}$), алгебраическим расширением поля $\boldsymbol{R}$ и пополнением метрического пространства $\boldsymbol{Q}^2$.

1. Определение вещественного числа как фундаментальной последовательности. Напомним, что при таком подходе в множестве $\boldsymbol{Q}$ рассматривают функцию $\rho(x,\,y)\rightleftarrows(x-y)$ вида $\rho:\boldsymbol{Q}^2\to\boldsymbol{Q}$. Относительно этой функции определяются понятие фундаментальной последовательности (множество всех таких последовательностей обозначим $\mathfrak{X}$) и понятие эквивалентности $\sim$ двух фундаментальных последовательностей (эти определения ниже напоминаются в общем случае, когда вместо пары $\langle\boldsymbol{Q},\,\rho\rangle$ рассматривается произвольное метрическое пространство $\langle X,\,\rho\rangle$). Затем вещественным числом называется любой класс эквивалентности в множестве $\mathfrak{X}$ относительно отношения эквивалентности $\sim$. Соответственно множество $\boldsymbol{R}$ определяется как фактор-множество $\mathfrak{X}/\sim$. После этого определяется функция

$$\rho([\{x_n\}],\,[\{y_n\}])\rightleftarrows[\{\rho(x_n,\,y_n)\}],$$

где в фигурных скобках пишутся фундаментальные последовательности в множестве $\boldsymbol{Q}$, а в квадратных скобках пишутся классы эквивалентности — элементы множества $\boldsymbol{R}$. Легко проверить, что если $\{x_n\}$ и $\{y_n\}$ фундаментальные последовательности, то $\{\rho(x_n,\,y_n)\}$ фундаментальная последовательность и правая часть в этом определении не зависит от выбора представителей в классах эквивалентности $[\{x_n\}]$ и $[\{y_n\}]$. Эта функция ρ называется метрикой в $\boldsymbol{R}$, и по ней, как обычно, определяется топология в $\boldsymbol{R}$. Кроме того, в $\boldsymbol{R}$ определяются операции: $[\{x_n\}]+[\{y_n\}]\rightleftarrows[\{x_n+y_n\}]$, $[\{x_n\}]\cdot[\{y_n\}]\rightleftarrows[\{x_n\cdot y_n\}]$, $-[\{x_n\}]\rightleftarrows[\{-x_n\}]$, $[\{x_n\}]^{-1}\rightleftarrows[\{x_n^{-1}\}]$, а также выделенные элементы: $0\rightleftarrows[\{0\}]$, $1\rightleftarrows[\{1\}]$, где в правых частях записи $\{0\}$ и $\{1\}$ обозначают тождественно нулевую и тождественно единичную последовательности; наконец, в $\boldsymbol{R}$ определяется порядок:

$$([\{x_n\}] < [\{y_n\}]) \rightleftarrows \exists m \forall n > m (x_n \leqslant y_n).$$

Конечно, нужно проверить, что правые части в этих определениях не зависят от выбора представителей в левых частях, а также то, что в правых частях действительно находятся фундаментальные последовательности. Это легко сделать.

Упражнение 3. *Докажите, что система вида*

$$\langle \boldsymbol{R}, +, -, \cdot, ^{-1}, 0, 1, \leqslant, \rho \rangle$$

является (линейно) упорядоченным метрическим полем (определение такого поля приводится ниже в начале п. 2).

Цель пунктов 1, 2 показать, что этот хорошо известный подход к определению системы вещественных чисел есть проявление общематематических закономерностей.

Пусть $\langle X, \rho \rangle$ метрическое пространство, где $\rho(\cdot, \cdot) X^2 \to \boldsymbol{R}$ метрика.

Например, $X = \boldsymbol{Q}$ и $\rho(x, y) = |x - y|$. Обратите внимание, что в этом примере функция ρ имеет вид: $\rho(\cdot, \cdot) : X^2 \to \boldsymbol{Q}$. Поэтому в этом и в ряде других важных случаев структура метрического пространства не предполагает заранее известным множество вещественных чисел $\boldsymbol{R}$, и, следовательно, оно может быть использовано для определения вещественного числа.

Напомним: фундаментальной последовательностью называется такая последовательность $x = \{x_n\} : \boldsymbol{N} \to \boldsymbol{N}$, для которой выполняется свойство $\forall \varepsilon \in \boldsymbol{Q}_{>0} \exists n \forall m,\ k > n (\rho(x_m, x_n) < \varepsilon)$. Множество всех фундаментальных последовательностей обозначим $\mathfrak{X}$.

В этом множестве определим отношение эквивалентности.

А именно две последовательности $\{x_n\}$ и $\{y_n\}$ назовем эквивалентными, если $\forall \varepsilon \in \boldsymbol{Q}_{>0} \exists n \forall m > n (\rho(x_m, y_m) < \varepsilon)$. Как всегда, это отношение эквивалентности обозначим $\sim$. Фактор-множество $\mathfrak{X}/\sim$ назовем пополнением метрического пространства X и обозначим $\hat{X}$.

В фактор-множестве $\hat{X}$ определим метрику по формуле

$$\hat{\rho}([\{x_n\}], [\{y_n\}]) \rightleftarrows \lim_{n \to \infty} \rho(x_n, y_n). \qquad (*)$$

Легко проверить, что это действительно метрика и, следовательно, пара $\langle \hat{X}, \hat{\rho} \rangle$ образует новое метрическое пространство. Например, проверим существование предела в формуле $(*)$:

$$|\rho(x_n, y_n) - \rho(x_m, y_m)| = |\rho(x_n, y_n) + \rho(y_n, x_m) - (\rho(y_n, x_m) + \rho(x_m, y_m))| \leqslant |\rho(x_n, x_m)| + |\rho(y_n, y_m)| \to 0,$$

т. е. $\{\rho(x_n, y_n)\}$ — фундаментальная последовательность. Из полноты множества $\boldsymbol{R}$ вытекает, что рассматриваемый предел существует.

Каким образом новое метрическое пространство $\langle \hat{X}, \hat{\rho} \rangle$ связано с исходным метрическим пространством $\langle X, \rho \rangle$?

Положим $i(x) \rightleftarrows [\{x_n\}]$, где $x_n \equiv x$. Тогда функция i вида $i : X \to \hat{X}$ и i — инъективная функция. Если отождествить объекты x и $i(x)$, то множество X будет частью множества $\hat{X}$ и функция $\hat{\rho}$, определенная на множестве $(\hat{X})^2$, будет продолжением функции ρ, определенной



Рис. 93

на множестве X^2. Во всяком случае, выполняется формула

$$\forall x, y \in X (\rho(x, y) = \hat{\rho}(i(x), i(y))).$$

Теорема 3. а) ***Множество $R(i)$ плотно в $\hat{X}$. Фундаментальная в X последовательность переходит относительно функции i в фундаментальную в X последовательность, которая, кроме того, имеет предел в $\hat{X}$*** (рис. 93).

б) Метрическое пространство $\langle \hat{X}, \hat{\rho} \rangle$ полно.

▷ а) Пусть α — произвольный элемент множества $\hat{X}$.

Если $\alpha = [\{x_n\}]$, то рассмотрим последовательность $\{x_n\}$. В соответствии с определением пополнения $\hat{X}$ все элементы x_n принадлежат X и сама последовательность $\{x_n\}$ фундаментальна в X. Покажем, что $\lim\limits_{n \to \infty} i(x_n) = \alpha$. Для этого фиксируем какое-то значение переменной n, например равное m. Получим $\widehat{\rho}(i(x_m), \alpha) = \lim\limits_{n \to \infty} \rho(x_m, x_n)$, так как по определению $i(x_m) = [\{x_m, x_m, \ldots, x_m, \ldots\}]$.

Фундаментальность последовательности $\{x_n\}$ означает, что

$$\lim_{m,\, n \to \infty} \rho(x_m, x_n) = \lim_{m \to \infty} \left(\lim_{n \to \infty} \rho(x_m, x_n) \right)$$

равен нулю, поэтому с ростом m величина $\hat{\rho}(i(x_m), \alpha)$ стремится к нулю.

Итак, любая точка α из X является пределом элементов вида $i(x_n)$ из $R(i)$. Это значит, что $R(i)$ плотно в $\hat{X}$.

Если $\{x_n\}$ — фундаментальная в X последовательность, то она принадлежит $\mathfrak{X}$ и $\alpha = [\{x_n\}]$ принадлежит $\hat{X}$. По доказанному $i(x_n) \to \alpha$. В частности, последовательность $\{i(x_n)\}$ фундаментальна в $\hat{X}$.

б) Пусть $\{\alpha_m\}$ — произвольная фундаментальная в $\hat{X}$ последовательность. При любом фиксированном m ее элемент α_m принадлежит $\hat{X}$, поэтому она имеет вид $[\{\hat{x}_n^{(m)}\}]$, где все $x_n^{(m)}$ — элементы X. Например, если $m = 5$, то α_5 вида $[\{x_1^{(5)}, x_2^{(5)}, x_3^{(5)}, x_4^{(5)}, \ldots, x_n^{(5)}, \ldots\}]$. По определению $\mathfrak{X}$ любая последовательность $\{x_n^{(m)}\}$ при фиксированном m фундаментальна в X. Для каждой точки α_m из $\hat{X}$ выберем из соответствующей ей последовательности $x_1^{(m)}, \ldots, x_n^{(m)} \ldots$ такой один элемент y_m, что $\hat{\rho}(i(y_m), \alpha_m) = \lim\limits_{n \to \infty} \rho(y_m, x_n^{(m)}) < \frac{1}{m}$. Это возможно для любого фиксированного m, так как $\{x_1^{(m)}, \ldots, x_n^{(m)}, \ldots\}$ — фундаментальная последовательность.

Найденное y_m совпадает с одним из членов последовательности $x_1^{(m)}, \ldots x_n^{(m)}$. Пусть это будет какой-то n_m-й член этой последовательности, т. е. $y_m = x_{n_m}^{(m)}$.

Осуществив выбор такого y_m для каждого m, мы получим последовательность $\{y_1, \ldots, y_m, \ldots\}$, в которой каждое $y_m = x_{n_m}^{(m)}$.

Последовательность $y_1 = x_{n_1}^{(1)}, \ldots, y_m = x_{n_m}^{(m)} \ldots$ образует что-то вроде диагонали в матрице $\{x_n^{(m)}\}$.

Полученная последовательность $\{y_m\}$ фундаментальна в X, так как

$$\rho(y_m, y_{m_1}) = \hat{\rho}(i(y_m), i(y_{m_1})) \leqslant \hat{\rho}(i(y_m), \alpha_m) + \hat{\rho}(\alpha_m, \alpha_{m_1}) +$$

$$+ \hat{\rho}(\alpha_{m_1}, i(y_{m_1})) \leqslant \left(\frac{1}{m} + \frac{1}{m_1} + \hat{\rho}(\alpha_m, \alpha_{m_1})\right) \to 0 \text{ при } m, m_1 \to \infty.$$

Поэтому $\{y_m\} \in \mathfrak{x}$ и $[\{y_m\}] \in \hat{X}$. Этот класс эквивалентности обозначим α_0. По доказанному в п. а выполняется $i(y_m) \to \alpha_0$ при $m \to \infty$.

С другой стороны, последовательности $\{i(y_m)\}$ и $\{\alpha_m\}$ с переменной m эквивалентны в $\hat{X}$, так как $\rho(i(y_m), \alpha_m) < \frac{1}{m}$. Первая из них имеет пределом точку α_0, поэтому вторая из них также имеет пределом этот элемент α_0. ■

Если Y — метрическое пространство и Y_1 — подмножество Y, то в Y_1 всегда рассматривается индуцированная метрика, т. е. $\rho_{Y_1}(y_1, y_2) \rightleftarrows \rho_Y(y_1, y_2)$, где $y_1, y_2 \in Y_1$.

Следствие. Если X и Y — метрические пространства, Y полно и существует метрический изоморфизм $i: X \leftrightarrow Y_1$, где Y_1 — плотное подмножество множества Y, то Y метрически изоморфно $\hat{X}$.

▷ Пусть y — произвольный фиксированный элемент из Y. Каждой последовательности $\{y_n\} \subseteq Y_1$, сходящейся к точке y, сопоставим последовательность $\{x_n\}$, для которой $y_n = i(x_n)$, т. е. $x_n = i^{-1}(y_n)$. Поскольку биекции i и i^{-1} сохраняют метрики, то $\{x_n\}$ — фундаментальная в X последовательность.

Легко проверить, что множество $\alpha_y \rightleftarrows \{\{i^{-1}(y_n)\} \mid \lim\limits_{n\to\infty} \{y_n\} = y, \{y_n\} \subseteq Y_1\}$ является классом эквивалентности в $\mathfrak{x}$, т. е. является элементом $\hat{X}$.

Положим $f(y) \rightleftarrows \alpha_y$, $f: Y \to \hat{X}$. Нетрудно убедиться и в том, что f — биекция и метрический изоморфизм. Например, $\rho_Y(y', y'') = \lim\limits_{n\to\infty} \rho_Y(y_n', y_n'')$, где $y_n' \to y'$ и $y_n'' \to y''$, $\{y_n'\}, \{y_n''\} \subseteq Y_1$, и продолжаем:

$$\rho_Y(y', y'') = \lim_{n\to\infty} \rho_X(x_n', x_n''), \text{ где } y_n' = i(x_n'), \; y_n'' = i(x_n''), \text{ и далее}$$

$$= \rho_{\hat{X}}([\{x_n'\}], [\{x_n''\}]). \quad ■$$

Напомним, что изоморфные (в данном случае метрически изоморфные) объекты обычно не различаются. Поэтому можно считать, что метрические пространства $\hat{X}$ и Y совпадают. Следствие показывает, что при фиксированном исходном метрическом пространстве X метрическое пространство $\hat{X}$ единственно (с точностью до метрического изоморфизма). Следовательно, пополнение метрического пространства X однозначно определяется как такое Y, которое удовлетворяет двум аксиомам:

1) Y — полное метрическое пространство;

2) X метрически изоморфно некоторому подмножеству Y_1, которое плотно в Y.

Итак, пополнение $\hat{X}$ можно определить аксиоматически.

Вернемся теперь к случаю, когда метрика ρ имеет вид $\rho: X^2 \to \boldsymbol{Q}$. В этом случае $\lim\limits_{n\to\infty} \rho(x_n, y_n)$ может не существовать в $\boldsymbol{Q}$.

Последовательность $\{\rho(x_n, y_n)\}$, конечно, остается фундаментальной, но ее предел может принадлежать не $\boldsymbol{Q}$, а $\hat{\boldsymbol{Q}}$. Поэтому формулу (*) приходится изменить, положив

$$\rho([\{x_n\}], [\{y_n\}]) \rightleftarrows [\{\rho(x_n, y_n)\}].$$

В этой формуле $\{x_n\}$ и $\{y_n\}$ — последовательности в X, а $\{z_n\} \rightleftarrows \{\rho(x_n, y_n)\}$ — последовательность в $\boldsymbol{Q}$. Следовательно, классы эквивалентности в левой части принадлежат $\hat{X}$, а класс эквивалентности в правой части принадлежит $\hat{\boldsymbol{Q}}$.

Если теперь рассмотреть случай $\boldsymbol{X} = \boldsymbol{Q}$ и $\rho(x, y) = (x - y): \boldsymbol{Q}^2 \to \boldsymbol{Q}$, то мы получим то определение множества вещественных чисел, которое было проведено в начале этого пункта. Итак, пополнение $\hat{\boldsymbol{Q}}$ можно принять за определение множества $\boldsymbol{R}$.

2. Продолжение алгебраических операций с поля на его пополнение. Метрическим полем называется система, являющаяся одновременно полем и метрическим пространством, т. е. система вида

$$\langle X, +, -, \cdot, {}^{-1}, 0, 1, \rho \rangle, \text{ где } \langle X, +, -, \cdot, {}^{-1}, 0, 1 \rangle —$$

поле и $\langle X, \rho \rangle$ — метрическое пространство, в которой выполняются аксиомы «согласования»: непрерывны функции $+$ и $\cdot$ (вида $X^2 \to X$) и функции (вида $X \to X$)^3 и $(\cdot)^{-1}$ (вида $X^* \to X^*$). Например, $\boldsymbol{Q}$ и $\boldsymbol{R}$ — метрические поля (проверьте).

В первом определении вещественных чисел использовалась только метрика в множестве $\boldsymbol{Q}$; иными словами, оно основано только на структуре метрического пространства в $\boldsymbol{Q}$. Структура поля и порядок в $\boldsymbol{Q}$ никак не используются в первом определении вещественных чисел.

Пополнение метрического пространства всегда является метрическим пространством, поэтому естественно, что $\boldsymbol{R} \rightleftarrows \hat{\boldsymbol{Q}}$ — метрическое пространство. В какой мере закономерно, что $\boldsymbol{R}$ еще и поле? Хотелось бы алгебраические операции в метрическом пространстве X перенести на его пополнение $\hat{X}$, с тем чтобы $\hat{X}$ превратилось в одноименную с X алгебраическую систему. Например, операции в $\boldsymbol{Q}$ продолжаются в $\boldsymbol{R} \rightleftarrows \hat{\boldsymbol{Q}}$ и $\boldsymbol{R}$ превращается в одноименную с $\boldsymbol{Q}$ алгебраическую систему — поле.

Итак, нас интересует вопрос, когда возможно продолжить алгебраические операции в метрическом поле X на его пополнение $\hat{X}$, с тем чтобы $\hat{X}$ тоже было полем, а X его подполем. Если это возможно, то возникает вопрос, единственно ли такое продолже-

ние операций и зависит ли оно от каких-либо особенностей метрического поля $\boldsymbol{Q}$ или имеет место для любого метрического поля X?

В этом пункте мы увидим, что такое продолжение возможно, единственно и является проявлением общей закономерности.

Удобно считать, что исходное метрическое пространство X не только изоморфно множеству $\boldsymbol{R}(i) \subseteq \widehat{X}$, но является подмножеством метрического пространства $\widehat{X}$. Действительно, всегда можно заменить в $\widehat{X}$ подмножество $\boldsymbol{R}(i)$ на множество X. Поэтому в дальнейшем будем считать, что X — подмножество $\widehat{X}$.

Теорема 4. ***Если X и Y метрические пространства, Y полно и $f: X \to Y$ равномерно непрерывная функция, то существует единственная равномерно непрерывная функция $\hat{f}: \widehat{X} \to Y$, для которой $\hat{f} \restriction X = f$.***

Функция $\hat{f}$ называется продолжением по непрерывности функции f. Если функция f не равномерно непрерывна, то ее продолжение даже до непрерывной функции не всегда возможно. Например, непрерывную функцию $\left(\frac{1}{x}\right) \restriction (0, 1]$ нельзя продолжить до непрерывной функции, определенной на отрезке $[0, 1]$.

$\triangleright$ Если $\{x_n\}$ — фундаментальная последовательность в X, то по равномерной непрерывности функции f последовательность $\{f(x_n)\}$ фундаментальна в Y. В силу полноты Y существует $\lim\limits_{n\to\infty} f(x_n)$. Положим

$$g: \{x_n\} \mapsto \lim_{n\to\infty} f(x_n).$$

Эта функция опускается на фактор-множество $\widehat{X}$, и получается функция вида $g_1: [\{x_n\}] \mapsto \lim\limits_{n\to\infty} f(x_n)$.

Итак, если $\alpha \rightleftarrows [\{x_n\}] \in \widehat{X}$, то $g_1(\alpha) \rightleftarrows g(\{x_n\})$. Ясно, что функция g_1 вида $g_1: X \to Y$ — продолжение функции f.

Осталось проверить, что функция g_1 равномерно непрерывна.

Действительно, пусть $\alpha' = [\{x_n'\}]$ и $\alpha'' = [\{x_n''\}]$ две точки из $\widehat{X}$. Тогда $\rho_Y(g_1(\alpha'), g_1(\alpha'')) = \rho_Y(\lim\limits_n f(x_n'), \lim\limits_n f(x_n'')) = \lim\limits_n \rho_Y(f(x_n'), f(x_n''))$.

Пусть дано любое число $\varepsilon > 0$. Выберем число $\delta > 0$, для которого выполняется $\rho_X(x', x'') < \delta \Rightarrow \rho_Y(f(x'), f(x'')) < \varepsilon$. Тогда условие $\rho_{\widehat{X}}(\alpha', \alpha'') = \lim\limits_n \rho_X(x_n', x_n'') < \delta$ обеспечивает соотношение $\exists n_0 \forall n > n_0 (\rho_Y(f(x_n'), f(x_n'')) < \varepsilon)$, т. е. $\rho_Y(g_1(\alpha'), g_1(\alpha'')) \leqslant \varepsilon$. Итак, $\hat{f} = g_1$.

Единственность такого продолжения сразу следует из того, что две непрерывные функции, совпадающие на плотном множестве, совпадают всюду. Равномерная непрерывность влечет непрерывность функции g_1. ■

Теорема 4 подсказывает идею продолжения операций с множеством X на множестве $\widehat{X}$.

Если f — унарная операция (например, $f(x) = -x$ или $f(x) = x^{-1}: X^* \to X^*$), то можно считать, что эта функция вида $f: X \to \widehat{X}$ или вида $X^* \to \widehat{X}$. Если функция f равномерно непрерывна, то можно непосредственно применить теорему 4. В топологической группе,

топологическом кольце, топологическом поле и т. п. все алгебраические операции по определению непрерывны, но откуда возьмется их равномерная непрерывность?

Если X компакт и f — непрерывная функция на X, то по известной теореме из курса анализа f равномерно непрерывна на X. Однако метрическое поле лишь в редких случаях является компактом.

Например, если $X = \boldsymbol{Q}$ и топология в $\boldsymbol{Q}$ определена обычной метрикой $\rho_Q(x, y) \rightleftarrows (x - y)$, то $\boldsymbol{Q}$ не является компактом.

Если X локально компактное пространство, то для любой точки a из X найдется компактная окрестность X_a этой точки в пространстве X. Тогда функция $f \restriction (X_a)$ равномерно непрерывна и к ней можно применить теорему 4. Но и локальная компактность неполного метрического поля X редко встречается. Например, она не имеет места в интересующем нас случае $X = \boldsymbol{Q}$ (докажите).

Тем более замечательно, что даже без всякого дополнительного предположения вроде локальной компактности можно продолжить операции, имеющиеся в X на пополнение $\widehat{X}$.

Теорема 5. ***Если X метрическая абелева группа, метрическое кольцо или тело, то операции, имеющиеся в X, продолжаются до операций в $\widehat{X}$ и относительно них $\widehat{X}$ является одноименной алгебраической системой. При этом X — соответственно подгруппа, подкольцо, подтело в $\widehat{X}$.***

Напомним несколько определений.

Метрической группой называется система вида $\langle G, \cdot, ^{-1}, e, \rho \rangle$, в которой $\langle G, \cdot, ^{-1}, e \rangle$ — группа и $\langle X, \rho \rangle$ — метрическое пространство и выполняются «аксиомы согласования»: групповые операции $\cdot$ и $^{-1}$ непрерывны и, наконец, метрика обладает свойством

$$\forall g_0, g_1, g_2 \in G (\rho(g_1, g_2) = \rho(g_0 g_1, g_0 g_2))$$

(инвариантность слева).

(Метрика ρ определяет в G топологию, относительно которой G является и топологической группой. Проверьте это.) Определение метрического кольца (тела) дается ниже.

В случае абелевой группы метрику и топологию в G удобнее задавать с помощью функции $|g| \rightleftarrows \rho(0, g)$, которая обладает свойствами $|\cdot|: G \to \boldsymbol{R}_{\geqslant 0}$, $|g| = 0 \Leftrightarrow g = 0$, $|-g| = |g|$, $|g_1 + g_2| \leqslant |g_1| + |g_2|$, $\forall g, g_1, g_2 \in G$.

Упражнение 4. *Проверьте, что в метрической группе действительно выполняются эти свойства. Докажите, что для любой абелевой группы G и для любой функции $|\cdot|: G \to \boldsymbol{R}_+$, обладающей этими четырьмя свойствами, в G возникает структура метрической группы; если положить $\rho(g_1, g_2) \rightleftarrows |g_1 - g_2|$.*

В элементарных группах метрика, как правило, задается именно такой функцией $|\cdot|$.

Лемма 2. ***Для любой абелевой метрической группы $\langle G, +, -, 0, |\cdot| \rangle$ функция $f: g \mapsto -g$ равномерно непрерывна.***

◁ Пусть дано число $\varepsilon>0$. Если $|g_1-g_2|<\varepsilon$, то $|(-g_1)-(-g_2)|=|f(g_1)-f(g_2)|<\varepsilon$ и в качестве δ можно выбрать само ε. ■

В доказательстве даже не используется непрерывность функции f.

Лемма 3. ***Всякий непрерывный гомоморфизм f абелевой метрической группы G в абелеву метрическую группу G_1 равномерно непрерывен.***

▷ Пусть дано число $\varepsilon>0$. По непрерывности функции f в точке нуль найдется число $\delta>0$, для которого

$$|g|<\delta\Rightarrow|f(g)|<\varepsilon.$$ Если $|g_1-g_2|<\delta$, то $|f(g_1)-f(g_2)|=|f(g_1-g_2)|<\varepsilon$. ■

В доказательстве используется непрерывность функции f только в одной точке — в нуле.

Лемма 4. ***Для любой абелевой метрической группы G функция $\cdot: G^2\to G$ равномерно непрерывна.***

▷ Операции в декартовом произведении G^2 определим покоординатно. Легко проверить, что относительно них G^2 — абелева группа. Тогда функция $\cdot$ является гомоморфизмом группы G^2 в группу G и к ней применима лемма 3. ■

Доказательство теоремы 5.

Случай группы. По лемме 2 функция $g\mapsto -g$ равномерно непрерывна, и к ней применима теорема 4. По лемме 4 функция $g_1\cdot g_2$ равномерно непрерывна, и к ней применима теорема 4.

Аксиомы абелевой группы представляют собой тождества, выполняющиеся на множестве G^2, плотном в $(\widehat{G})^2$. По непрерывности левых и правых частей этих тождеств они выполняются и на всем множестве $\widehat{G}$.

Случай кольца. Кольцо $\langle K, +, -, 0, \cdot, 1\rangle$ содержит абелеву группу $\langle K, +, -, 0\rangle$. По уже доказанному случаю теоремы 5 система $\langle\widehat{K}, +, -, 0\rangle$ — абелева группа. Осталось продолжить функцию $k_1\cdot k_2: K^2\to K$ до функции вида $(\widehat{K})^2\to\widehat{K}$. Однако уже в простейшем случае $K=\boldsymbol{R}$ функция $x\cdot y:\boldsymbol{R}^2\to\boldsymbol{R}$ и даже функция $x^2:\boldsymbol{R}\to\boldsymbol{R}$ не равномерно непрерывны. Поэтому леммы 2—4 не применимы. Причина этого в том, что метрика в K, инвариантная относительно всех параллельных переносов вида $k\mapsto k+k_0$, не инвариантна относительно гомотетий вида $k\mapsto k\cdot k_0$.

Лемма 5. ***Пусть X_1, X_2, X_3 полные метрические абелевы группы, а G_1 и G_2 плотные подгруппы соответственно в X_1 и X. Всякая непрерывная билинейная форма $f(\cdot,\cdot): G_1\times G_2\to X_3$ продолжается до непрерывной билинейной формы вида $X_1\times X_2\to X_3$.***

Закончим сначала доказательство второго случая теоремы 5.

Поскольку $k_1\cdot k_2: K^2\to\widehat{K}$ — непрерывная билинейная форма, то по лемме 5 продолжим ее на $(\widehat{K})^2$ до функции вида $(\widehat{K})^2\to\widehat{K}$. Аксиомы кольца являются тождествами, выполняющимися на множестве K^2, плотном в $(\widehat{K})^2$. По непрерывности левых и правых частей этих тождеств они выполняются и на $\widehat{K}$.

▷ Доказательство леммы 5. Если $\{g'_n\}$ и $\{g''_n\}$ — фундаментальные последовательности соответственно в G_1 и G_2, которые

сходятся соответственно к точкам g'_0 и g''_0, то $\{f(g'_n, g''_n)\}$ — фундаментальная последовательность в X_3.

Действительно, начиная с достаточно больших n, m, величины $\rho(g'_n, g'_m)$ и $\rho(g''_n, g''_m)$ сколь угодно малы. Выполняется тождество

$$f(x', y')-f(x, y)=f(x'-x, y_1)+f(x_1, y'-y)+f(x'-x, y'-y_1)+f(x-x_1, y'-y).$$

Мы покажем, что слагаемые в его правой части при достаточно малых $|x'-x|$ и $|y'-y|$ и подходящем выборе x_1 и y_1 сколь угодно малы.

По непрерывности функций $f(\cdot,\cdot)$ в точке $(0, 0)$ для любого числа $\varepsilon>0$ существует такое число $\delta>0$, для которого $|x|, |y|<2\delta\Rightarrow|f(x, y)|<\varepsilon$. Если $|x'-x|, |y'-y|<2\delta$, то $|f(x'-x, y'-y)|<\varepsilon$.

Пусть $x_1\in\mathcal{O}_{g'_0,\delta}$ и $y_1\in\mathcal{O}_{g''_0,\delta}$. Тогда

$$\forall x',\ x\in\mathcal{O}_{g'_0\delta}\forall y',\ y\in\mathcal{O}_{g''_0,\delta}(|f(x'-x, y'-y_1)+f(x-x_1, y'-y)|<2\varepsilon).$$

Функции $x\mapsto f(x, y_1)$ и $y\mapsto f(x_1, y)$ непрерывны. В частности, они непрерывны в точке нуль. Поэтому, уменьшив, если нужно, δ, можно обеспечить: $|x'-x|<2\delta\Rightarrow|f(x'-x, y_1)|<\varepsilon$ и $|y'-y|<2\delta\Rightarrow|f(x_1, y'-y)|<\varepsilon$.

Итак, фундаментальность последовательности $\{f(g'_n, g''_n)\}$ доказана.

По полноте X_3 эта последовательность имеет предел. Положим $\langle\{g'_n\}, \{g''_n\}\rangle\mapsto\lim_n f(g'_n, g''_n)$. Такая функция поднимается до функции $g(\cdot,\cdot): \langle[\{g'_n\}], [\{g''_n\}]\rangle\mapsto\lim_n f(g'_n, g''_n)$, которая является продолжением $f(\cdot)$. Аргумент этой функции как раз пробегает $\widehat{G}_1\times\widehat{G}_2=X_1\times X_2$. Непрерывность функции g проверяется точно так же, как в доказательстве теоремы 4. Эти два доказательства и в целом похожи.

Условие билинейности функции g записывается тождеством, которое выполняется на плотном множестве $G_1\times G_2$. По непрерывности оно сохраняется и на $X_1\times X_2$. ■

Случай тела. Тело $\langle K, +, -, 0, \cdot, {}^{-1}, 1\rangle$ содержит кольцо $\langle K, +, -, 0, \cdot, 1\rangle$. По доказанному случаю нашей теоремы 5 $\langle\widehat{K}, +, -, 0, \cdot, 1\rangle$ — кольцо. Осталось только продолжить функцию $k^{-1}: K^*\to K$ до функции вида $k^{-1}: (K^*)^\wedge\to\widehat{K}$. Тогда аксиомы тела проверяются точно так же, как проверялись аксиомы группы и кольца. Однако для продолжения этой функции нельзя применить леммы 2—4 или лемму 5. Это видно из простейшего примера $K=\boldsymbol{R}$ и $k^{-1}=\frac{1}{k}:\boldsymbol{R}^*\to\boldsymbol{R}$.

Пусть $\{k_n\}$ фундаментальная последовательность в K, не сходящаяся к нулю в K. Если $\{k_n^{-1}\}$ фундаментальная последовательность, то положим $\{k_n\}\mapsto\lim_n(k_n^{-1})$. Эту функцию поднимем до функ-

ции вида $[(k_n)] \mapsto \lim_n k_n^{-1}$ и таким образом осуществим продолжение функции k^{-1}. Осталось проверить фундаментальность последовательности $\{k_n^{-1}\}$. По определению метрического тела в нем метрика ρ определяется с помощью функции вида $|\cdot| : K \to \boldsymbol{R}_{\geqslant 0}$ со свойствами:

$$|k| = 0 \Leftrightarrow k = 0,$$
$$|k_1 \cdot k_2| = |k_1| \cdot |k_2|,$$
$$|k_1 + k_2| \leqslant |k_1| + |k_2|, \ \forall k, k_1, k_2 \in K.$$

А именно, $\rho(k_1, k_2) \rightleftarrows |k_1 - k_2|$.

(Не для любого поля можно определить хотя бы одну такую функцию, но для элементарных полей, в частности, для всех числовых полей она может быть определена.) Функцию $|\cdot|$ с этими свойствами называют абсолютным значением. Итак, метрическое тело определяется как тело с абсолютным значением, в котором все операции непрерывны.

Докажем, что последовательность $\{k_n^{-1}\}$ фундаментальна.

По условию о том, что $\lim k_n \neq 0$, существует число $\varepsilon > 0$, для которого $\exists n_0 \forall n > n_0 (|k_n| \geqslant \varepsilon)^n$.

Для любого числа $\varepsilon_1 > 0$, увеличив, если нужно, n_0, можно считать, что $\forall n, m > n_0 (|k_n - k_m| < \varepsilon_1)$.

Тогда $|k_n^{-1} - k_m^{-1}| = \frac{|k_m - k_n|}{|k_n| \cdot |k_m|} \leqslant \frac{\varepsilon_1}{\varepsilon^2}$. Число ε фиксировано, а число ε_1 можно сделать сколь угодно малым. Поэтому последнюю дробь можно сделать сколь угодно малой при всех n, $m > n_0$. Теорема 5 полностью доказана. ■

3. Определение вещественного числа как сечения. В этом пункте мы разберем второй подход к определению вещественного числа. Как и первый подход, он имеет в качестве отправной точки множество рациональных чисел $\boldsymbol{Q}$. Однако теперь множество $\boldsymbol{Q}$ рассматривается не как метрическое пространство и не как поле, а как упорядоченное множество. Системы $\langle \boldsymbol{Q}, \rho \rangle$ и $\langle \boldsymbol{Q}, \leqslant \rangle$ имеют общий носитель, но по существу являются разными объектами. Наша цель состоит не в том, чтобы фактически определить вещественные числа с помощью упорядоченного множества $\langle \boldsymbol{Q}, \leqslant \rangle$ (это делается в курсе «Числовые системы»), а в том, чтобы разобрать сам способ такого определения и выяснить, зависит ли он от каких-то специфических свойств упорядоченного множества $\langle \boldsymbol{Q}, \leqslant \rangle$. С этой целью приведем общее определение класса систем, простейшим представителем которых является конкретная система $\langle \boldsymbol{Q}, +, -, 0, \cdot, \leqslant \rangle$. Пусть K является либо полем $\boldsymbol{Q}$, либо полем $\boldsymbol{R}$.

Определение 1а. **Линеалом называется система вида $\langle \mathfrak{x}, +, -, \cdot, 0, \leqslant \rangle$, где система $\langle \mathfrak{x}, +, -, \cdot, 0 \rangle$ — векторное пространство над полем K, а $\leqslant$ — отношение порядка в множестве $\mathfrak{x}$, для которой выполняются аксиомы:**

1) $\langle \mathfrak{x}, \leqslant \rangle$ — решетка, т. е. для любых двух элементов ξ и η из $\mathfrak{x}$ существуют точные верхняя и нижняя грани множества $\{\xi, \eta\}$;

2) $\xi \leqslant \eta \Rightarrow \xi + \zeta \leqslant \eta + \zeta$, где $\xi, \eta, \zeta \in \mathfrak{x}$;

3) $0 \leqslant \lambda \wedge \xi \leqslant \eta \Rightarrow \lambda\xi \leqslant \lambda\eta$, где $\xi, \eta \in \mathfrak{x}$ и $\lambda \in \boldsymbol{R}$.

Напомним, что точная верхняя грань $\sup\{x_\alpha\}$ множества $\{x_\alpha\}$ определяется как элемент x из частично упорядоченного множества X, для которого выполняются свойства:

1) $x_\alpha \leqslant x$, $\forall \alpha$;
2) $\forall y \in X((\forall \alpha (x_\alpha \leqslant y) \Rightarrow x \leqslant y)$.

Аналогично определяется точная нижняя грань $\inf\{x_\alpha\}$ множества $\{x_\alpha\}$. Здесь $\{x_\alpha\}$ — произвольное семейство элементов какого угодно частично упорядоченного множества X.

Вместо $\sup\{x_\alpha\}$ часто пишут $\sup x_\alpha$ или $\bigvee_\alpha x_\alpha$. Легко проверить: если существует точная верхняя грань, то она единственна. Вместо $\sup\{x, y\}$ обычно пишут $x \vee y$. Аналогично, не может существовать двух $\inf\{x_\alpha\}$. Вместо последнего обозначения часто пишут $\inf_\alpha x_\alpha$ или $\bigwedge_\alpha x_\alpha$. Вместо $\inf\{x, y\}$ пишут $x \wedge y$.

Приведите примеры решеток и примеры отношений порядка, не являющихся решетками. Те и другие упорядоченные множества часто встречаются в элементарной математике.

Определение 1 сохраняет свой смысл, если в нем упорядоченное поле K заменено любым другим упорядоченным полем или упорядоченным кольцом K. (В последнем случае $\mathfrak{x}$ является модулем над K.) Например, можно взять $K = \boldsymbol{Q}$; этот случай будет в дальнейшем основным.

Примерами линеалов являются системы $\langle \boldsymbol{Q}, +, -, 0, \leqslant \rangle$ над полем $\boldsymbol{Q}$ и $\langle \boldsymbol{R}, +, -, 0, \leqslant \rangle$ над полем $\boldsymbol{Q}$, и вторая из этих систем над полем $\boldsymbol{R}$, где $+$, $-$, $\leqslant$ суть обычные операции и отношение порядка между числами. Другими примерами линеалов являются система $\langle \boldsymbol{R}^2, +, -, 0, \leqslant \rangle$ над полем $\boldsymbol{Q}$ и эта же система над полем $\boldsymbol{R}$, где отношение порядка определяется формулой $(\langle x_1, y_1 \rangle \leqslant \langle x_2, y_2 \rangle) \rightleftarrows (x_1 \leqslant x_2 \wedge y_1 \leqslant y_2 \rangle$, а также система $\langle C(X, \boldsymbol{R}), +, -, 0, \leqslant \rangle$ с обычными покоординатными операциями и отношением порядка $(f \leqslant g) \rightleftarrows \forall x \in X (f(x) \leqslant g(x))$.

Условимся в дальнейшем рассматривать множество $\boldsymbol{Q}$ как линеал именно над полем $\boldsymbol{Q}$.

Порядок в множестве X называется условно полным, если любое непустое ограниченное сверху подмножество множества X имеет в X точную верхнюю грань.

Отсюда легко вытекает, что любое непустое ограниченное снизу подмножество множества X имеет в X точную нижнюю грань. Например, порядок в $\boldsymbol{R}$ условно полный, а порядок в $\boldsymbol{Q}$ не является таким.

Определение 1б. **K-пространством называется любой условно полный линеал.**

Определение 1в. **Линеал называется архимедовым, если выполняется условие: $\forall y > 0$ ($\{n \cdot y | n \in N\}$ не ограничено сверху).**

Любое K-пространство архимедово. Действительно, если множество $\{n\cdot y \mid n\in N\}$ ограничено сверху и $y>0$, то обозначим $z\rightleftarrows \rightleftarrows \sup\{n\cdot y \mid n\in N\}$. Тогда $\forall n\,(n\cdot y\leqslant z)$; $(n-1)\cdot y\leqslant z-y$. Поскольку коэффициент $(n-1)$ пробегает весь натуральный ряд, то $\{n\cdot y\}\leqslant z-y$, $z\leqslant z-y$, $0\leqslant -y$, $y\leqslant 0$ — противоречие.

Линеал, не являющийся K-пространством, также может быть архимедовым. Например, архимедовым не K-пространством является интересующий нас линеал Q (над полем Q).

Теорема 6. *Для любого архимедова линеала $\mathfrak{X}$ существует такое K-пространство $\hat{\mathfrak{X}}$ и такая функция $i\colon \mathfrak{X}\to\hat{\mathfrak{X}}$, что $R(i)$ плотно в $\tilde{\mathfrak{X}}$ в следующем смысле:*

$$\forall\xi\in\tilde{\mathfrak{X}}\,\exists\{\xi_\alpha\}\subseteq\mathfrak{X}\ (\xi=\sup\{i(\xi_\alpha)\});$$

кроме того, функция i является инъективным гомоморфизмом линеалов (вложением).

Последнее означает, что функция i инъективна и $i(\xi+\eta)=i(\xi)+i(\eta)$, $i(\lambda\cdot\xi)=\lambda\cdot i(\xi)$, $\xi\leqslant\eta\Leftrightarrow i(\xi)\leqslant i(\eta)$ и $i(\bigvee_j \xi_j)=\bigvee_j i(\xi_j)$.

Конечно, $R(i)$ — подлинеал $\tilde{\mathfrak{X}}$, этот подлинеал обладает свойством, о котором говорится в следующем упражнении. При наличии этого свойства подлинеал называется правильным, так что $R(i)$ — правильный подлинеал в $\tilde{\mathfrak{X}}$.

Упражнение 5. *Для любого подмножества $w\subseteq R(i)$ существование* $\sup w$, *вычисленного относительно $R(i)$, эквивалентно существованию* $\sup w$, *вычисленного относительно $\tilde{X}$, и, если один из этих двух* sup *существует, то они оба равны. Приведите примеры двух упорядоченных множеств X и Y ($X\subseteq Y$ и порядок в X — ограничение порядка в Y), для которых такое утверждение не верно.*

Ниже будет доказано, что K-пространство $\tilde{\mathfrak{X}}$, указанное в теореме 6, определяется по архимедову линеалу $\mathfrak{X}$ единственным образом с точностью до изоморфизма линеалов. Это K-пространство называется пополнением линеала $\mathfrak{X}$.

Второе определение множества вещественных чисел состоит в том, что вещественным числом называется любой элемент пополнения $\tilde{Q}$ архимедова линеала Q. Иными словами, множеством вещественных чисел называется $\tilde{Q}$. По теореме 6 множество $\tilde{Q}$ оказывается условно полным линеалом. Таким образом, второе определение вещественного числа, как и его первое определение, является частным случаем соответствующей общей конструкции, которая применяется далеко не только для случая Q.

Доказательство теоремы 6.

▷ Если $\alpha\subseteq\mathfrak{X}$, то обозначим α^s (соответственно α^i) множество всех верхних (соответственно всех нижних) границ множества α. Конечно, множество α^s или α^i может быть пустым. Обозначим α^{si} результат последовательного вычисления $(\alpha^s)^i$.

Подмножество $\alpha\subseteq\mathfrak{X}$ называется сечением, если $\alpha=\alpha^{si}$. Конечно, всегда выполняется $\alpha\subseteq\alpha^{si}$. Поэтому α сечение, если $\alpha^{si}\subseteq\alpha$. Например, все множество $\mathfrak{X}$ и пустое множество $\varnothing$ являются сечениями.

По определению множество $\tilde{\mathfrak{X}}$ состоит из всех сечений, кроме двух сечений $\mathfrak{X}$ и $\varnothing$.

Упражнение 6. *Докажите, что все сечения в линеале $\langle Q, \leqslant\rangle$ имеют вид $\{r\in Q \mid r\leqslant\lambda\}$, где λ — какое-то фиксированное вещественное число. Отсюда видна связь второго определения вещественного числа с обычным определением вещественного числа с помощью сечения, по Дедекинду.*

Для любого множества $\alpha\subseteq\mathfrak{X}$ множество α^{si} сечение, так как $\alpha^{sis}=\alpha^s$. Упорядочим множество $\tilde{\mathfrak{X}}$ отношением порядка $(\alpha\leqslant\beta)\rightleftarrows\rightleftarrows(\alpha\subseteq\beta)$; здесь и далее α, β — произвольные элементы множества $\mathfrak{X}$. Легко проверить, что это действительно отношение порядка.

Покажем, что упорядоченное множество $\tilde{\mathfrak{X}}$ условно полно. Для непустого ограниченного сверху в $\tilde{\mathfrak{X}}$ множества сечений $\{\alpha_\gamma\}\subseteq\tilde{\mathfrak{X}}$ положим $\alpha\rightleftarrows(\bigcup_\gamma \alpha_\gamma)^{si}$. Если сечение $\beta\in\tilde{\mathfrak{X}}$ ограничивает сверху множество $\{\alpha_\gamma\}$, то $\bigcup_\gamma\alpha_\gamma\subseteq\beta$ и $(\bigcup_\gamma\alpha_\gamma)^{si}\subseteq\beta^{si}=\beta$. Поэтому $\alpha\neq\mathfrak{X}$. Кроме того, $\alpha\neq\varnothing$. Итак, $\alpha\in\tilde{\mathfrak{X}}$ и $\forall\gamma(\alpha_\gamma\leqslant\alpha)$.

Проверим, что $\alpha=\bigvee_\gamma\alpha_\gamma$. Если $\forall\gamma(\alpha_\gamma\leqslant\alpha_0)$, то $\bigcup_\gamma\alpha_\gamma\subseteq\alpha_0$, $(\bigcup_\gamma\alpha_\gamma)^{si}\subseteq\alpha_0^{si}=\alpha_0$, т. е. $\alpha\leqslant\alpha_0$.

Положим $i(\cdot)\colon\xi\mapsto\alpha_\xi$, где $\alpha_\xi\rightleftarrows\{\eta\in\mathfrak{X}\mid\eta\leqslant\xi\}$. (Любое множество вида α_ξ называется компонентой.) Легко видеть, что функция i вида $i\colon\mathfrak{X}\to\tilde{\mathfrak{X}}$ инъективна.

Упражнение 7. *Докажите, что в случае линеала $\langle Q, \leqslant\rangle$ такая функция i является гомоморфизмом линеала Q в линеал его сечений $\tilde{Q}$.*

Теперь нужно в общем случае проверить, что функция i является гомоморфизмом линеалов. Например, проверим, что $i(\bigvee\xi_\gamma)=\bigvee i(\xi_\gamma)$. Если $\xi=\bigvee_\gamma\xi_\gamma$, то нужно проверить, что $i(\xi)=\alpha_\xi=\bigvee_\gamma i(\xi_\gamma)\rightleftarrows\bigvee_\gamma\alpha_{\xi_\gamma}$. Как мы уже видели, $\alpha\rightleftarrows(\bigcup_\gamma\alpha_{\xi_\gamma})^{si}=\bigvee_\gamma\alpha_{\xi_\gamma}$. Достаточно установить, что $\alpha_\xi=\alpha$. Если $\eta\in\alpha_\xi$ то $\eta\leqslant\xi$, но $\xi\in\alpha$ и $\eta\in\alpha$.

Здесь $\lambda\cdot\alpha$ — умножение на скаляр λ, а $\lambda\alpha$ — другое умножение: все элементы α умножаются на λ.

Для α, $\beta\in\mathfrak{X}$ положим $\alpha+\beta\rightleftarrows(\alpha+\beta)^{si}$, где в левой части складываются элементы α, $\beta\in\tilde{\mathfrak{X}}$, а в правой части складываются подмножества α, $\beta\subseteq\mathfrak{X}$. Если $\lambda\in Q$ или $\lambda\in R$ и $\alpha\in\tilde{\mathfrak{X}}$, то в случае $\lambda>0$ положим $\lambda\cdot\alpha\rightleftarrows\{\lambda\cdot\xi\mid\xi\in\alpha\}$; если же $\lambda<0$, то положим $\lambda\cdot\alpha=\lambda\alpha^s$. Наконец, если $\lambda=0$, то положим $\lambda\cdot\alpha\rightleftarrows 0\rightleftarrows\{\eta\mid\eta\leqslant 0\}$.

Проверим, что множество $\lambda\cdot\alpha$ является сечением. Например, рассмотрим случай, когда $\lambda<0$. Если $\xi\in\alpha$, $\eta\in\alpha^s$, то $\lambda\cdot\xi\geqslant\lambda\cdot\eta$; значит, по определению получаем $\lambda\cdot\xi\in(\lambda\alpha^s)^s$, т. е. $\lambda\alpha\subseteq(\lambda\alpha^s)^s$.

Наоборот, если $\eta\in(\lambda\alpha^s)^s$, то $\eta\geqslant\lambda\cdot\xi$, $\forall\xi\in\alpha^s$ и , значит,

$\frac{1}{\lambda}\cdot\eta\leqslant\xi$, $\frac{1}{\lambda}\cdot\eta\in\alpha^{si}=\alpha$, $\eta\in\lambda\alpha$, $(\lambda\alpha^s)^s\subseteq\lambda\alpha$. Итак, $\lambda\alpha=(\lambda\alpha^s)^s$. Проверим теперь, что $(\lambda\alpha^s)^{si}\subseteq\lambda\alpha^s$. Если $\xi\in(\lambda\alpha^s)^{si}$, то $\xi\leqslant\lambda\cdot\eta$, $\forall\eta\in\alpha$ и $\eta\leqslant\frac{1}{\lambda}\cdot\xi$, $\frac{1}{\lambda}\cdot\xi\in\alpha^s$, $\xi=\lambda\eta$, где $\eta\in\alpha^s$, и по определению $\lambda\cdot\alpha$ получаем: $\xi\in\lambda\cdot\alpha$.

Следовательно, операция $\cdot$ вида $\cdot : K\times\tilde{\mathfrak{X}}\to\tilde{\mathfrak{X}}$, где K совпадает с $\boldsymbol{Q}$ или $\boldsymbol{R}$.

Итак, в $\tilde{\mathfrak{X}}$ определены операции $+$ и $\cdot$ нужного вида, имеется элемент 0 и отношение порядка $\leqslant$.

Чтобы убедиться в том, что линеал $\tilde{\mathfrak{X}}$ относительно этих операций и отношения порядка является K-пространством, осталось проверить аксиомы из определения 1а и то обстоятельство, что $\langle\tilde{\mathfrak{X}}, +, \cdot, 0, \leqslant\rangle$ — векторное пространство. Аксиомы 1—3 очевидны. Из аксиом векторного пространства также очевидны все аксиомы, кроме ассоциативности и дистрибутивности операции сложения векторов относительно операции умножения на скаляр.

Их проверку оставим в качестве упражнения.

Итак, линеал $\tilde{\mathfrak{X}}$ есть K-пространство.

Проверим плотность множества $R(i)=\{\alpha_\xi\,|\,\xi\in\mathfrak{X}\}$ в $\tilde{\mathfrak{X}}$. Если $\alpha\in\tilde{\mathfrak{X}}$, то $\alpha=\bigcup\{\alpha_\xi\,|\,\xi\in\alpha\}$, так как α обладает свойством: $\eta\leqslant\xi\in\alpha\Rightarrow\eta\in\alpha$. Поэтому $\alpha=\alpha^{si}=(\bigcup\alpha_\xi)^{si}=\bigvee\{\alpha_\xi\,|\,\xi\in\alpha\}$. ■

Следствие. Пусть $\mathfrak{X}$ и $\mathfrak{Y}$ линеалы и $\mathfrak{Y}$ есть K-пространство и существует вложение i линеала $\mathfrak{X}$ в линеал $\mathfrak{Y}$, образ которого плотен в $\mathfrak{Y}$. Тогда $\mathfrak{Y}$ изоморфно $\tilde{\mathfrak{X}}$.

Слово «изоморфно» означает существование биекции $f:\mathfrak{Y}\leftrightarrow\tilde{\mathfrak{X}}$, для которой

$$f(y_1+y_2)=f(y_1)+f(y_2),$$
$$f(\lambda\cdot y)=\lambda\cdot f(y),$$
$$y_1\leqslant y_2\Leftrightarrow f(y_1)\leqslant f(y_2),$$
$$f(\bigvee_\gamma y_\gamma)=\bigvee_\gamma f(y_\gamma).$$

Для простоты записи будем считать, что $\mathfrak{X}\subseteq\mathfrak{Y}$ и $i(\xi)\equiv\xi$. Этого всегда можно добиться, отождествляя элемент ξ из $\mathfrak{X}$ и элемент $i(\xi)$ из $\mathfrak{Y}$.

▷ Положим $f(\cdot):y\mapsto\alpha_y\rightleftarrows\{\xi\in\mathfrak{X}\,|\,\xi\leqslant y\}$. Очевидно, α_y — сечение в $\mathfrak{X}$. Если $y_1\neq y_2$ и $\alpha_{y_1}=\alpha_{y_2}$, то в силу плотности $\mathfrak{X}$ получаем $y_1=y_2$ — противоречие. Если α — сечение в $\mathfrak{X}$, то положим $y\rightleftarrows\sup_\gamma\alpha$. Тогда $\alpha_y=\alpha$. Действительно, если $\xi\in\alpha_y$, то $\alpha\leqslant y$, $\eta\in\alpha^s\Rightarrow\xi\leqslant y$, $\xi\in\alpha^{si}=\alpha$. Если $\xi\in\alpha$, то $\xi\leqslant y$, $\xi\in\alpha_\gamma$ и $y\mapsto\alpha$. Итак, функция f — биекция множеств $\mathfrak{Y}$ и $\tilde{\mathfrak{X}}$. Легко проверить, что она изоморфизм. ■

В любом линеале $\mathfrak{X}$ определяется функция

$$|\xi|\rightleftarrows\max\{\xi,\ -\xi\}.$$

Она называется абсолютной величиной в $\mathfrak{X}$. В случае $\boldsymbol{R}$ или $\boldsymbol{Q}$ эта функция является (обычной) абсолютной величиной.

Итак, при втором подходе к понятию вещественного числа $\boldsymbol{R}\rightleftarrows\tilde{\boldsymbol{Q}}$ получается как условно полное упорядоченное множество

и одновременно векторное пространство над полем $\boldsymbol{Q}$. Однако пока в множестве $\boldsymbol{R}$ не определена операция умножения. Поэтому умножение на числа (скаляры) из $\boldsymbol{Q}$, которое определено в множестве $\boldsymbol{R}$, продолжим до умножения на числа (скаляры) из $\boldsymbol{R}$. А именно, если $\alpha\in\boldsymbol{R}=\tilde{\boldsymbol{Q}}$, то для любого сечения $\alpha_1\in\boldsymbol{R}$ положим $\alpha\cdot\alpha_1\rightleftarrows\{r\cdot\alpha_1\,|\,r\in\alpha\}$. Получилась операция вида $\boldsymbol{R}^2\to\boldsymbol{R}$. Легко проверить, что для нее выполняются обычные свойства умножения. Поэтому $\boldsymbol{R}$ является кольцом и полем.

4. Определение вещественного числа как последовательности знаков. Третье определение вещественного числа основано на интуитивном представлении о вещественном числе как последовательности целых чисел или, более точно,— последовательности цифр (знаков). Вероятно, это наиболее элементарное, школьное представление о вещественном числе.

Множество вещественных чисел, определенное в соответствии с п. 1 или п. 2, обозначим $\boldsymbol{R}$. Наряду с этим множеством рассмотрим декартово произведение $\boldsymbol{Z}\times(F_q)^N$, где q — фиксированное натуральное число, большее единицы, и $F_q\rightleftarrows\{0, 1, \ldots, q-1\}$.

Рассмотрим множество всех последовательностей вида $x_0, x_1\ x_2\ x_3\ \ldots$, где $x_0\in\boldsymbol{Z}$ и $x_1\in F_q$, $x_2\in F_q$, ... Любой член x_i такой последовательности назовем знаком. Обозначим R' подмножество множества $\boldsymbol{Z}\times(F_q)^N$, состоящее из всех таких последовательностей $x_0, x_1, x_2, x_3, \ldots$, у которых знаки x_i не равны тождественно $(q-1)$, начиная с некоторого места. Например, при $q=10$ множество R' не содержит последовательностей, у которых, начиная с некоторого места, идут одни девятки. Элементы множества R' будем называть q-дробями, или, короче, дробями.

Определим каноническую (фиксированную) биекцию φ между множествами R' и $\boldsymbol{R}$. Пусть дробь $x=x_0,\ x_1x_2x_3\ \ldots$ принадлежит множеству R'.

Положим $r_n\rightleftarrows r_n(x)\rightleftarrows x_0+\ldots+\frac{x_n}{q^n}$, $\varphi(x)\rightleftarrows\lim\limits_{n\to\infty}r_n$. Очевидно, что $\{r_n\}$ — фундаментальная последовательность в $\boldsymbol{Q}$, поэтому существует $\lim r_n$. Итак, $\varphi:R'\to\boldsymbol{R}$.

Чтобы увидеть связь между дробями и измерением отрезков, которое рассматривалось в главе III, определим функцию φ по-другому, не используя рациональные числа r_n. Для этого разделим отрезок $\Delta_0\rightleftarrows[x_0,\ x_0+1]$, содержащийся в $\boldsymbol{R}$, на q равных отрезков, не имеющих общих внутренних точек, которые в соответствии с обычным отношением порядка в $\boldsymbol{R}$ занумеруем слева направо соответственно цифрами 0, 1, 2, ..., $q-1$. Обозначим Δ_1 тот из этих отрезков, который имеет номер x_1.

Аналогичным образом разделим отрезок Δ_1 на q равных отрезков и обозначим Δ_2 тот из полученных отрезков, который имеет номер x_2. Продолжая этот процесс, получим систему вложенных отрезков

$$\Delta_0\supseteq\Delta_1\supseteq\Delta_2\supseteq\ldots,$$

в которой длина n-го отрезка Δ_n равна $\frac{1}{q^n}$. Поэтому длина отрезка Δ_n стремится к нулю с ростом индекса n. По теореме Кантора, которая, как известно, выполняется в множестве $\boldsymbol{R}$, существует единственное вещественное число $\lambda \in \boldsymbol{R}$, принадлежащее всем отрезкам Δ_n. Поскольку все x_i не могут, начиная с некоторого номера, равняться $q-1$, то $\lambda \neq x_0+1$, т. е. $\lambda \in [x_0, x_0+1[$, и точно так же вещественное число λ не равно правому концу никакого отрезка Δ_n. Поэтому при каждом делении можно образовывать отрезки без их правых концов.

Обозначим $\varphi(x)$ полученное число x. Оставим читателю проверку того, что эти два определения функции φ эквивалентны. Биективность функции $\varphi : x \mapsto \varphi(x)$ легко проверяется при первом ее определении. Проверим то же самое для ее второго определения.

Предложение 6. ***Функция $\varphi: R' \to \boldsymbol{R}$ является биекцией.***

▷ Пусть $x = x_0, x_1, x_2 \ldots$ и $y = y_0, y_1, y_2 \ldots$ — две различные дроби из множества R'. Если $x_0 \neq y_0$, то $\varphi(x) \in [x_0, x_0+1[$ и $\varphi(y) \in [y_0, y_0+1[$, поэтому $\varphi(x) \neq \varphi(y)$. Если $x_0 = y_0$ и $x_1 \neq y_1$, то $\varphi(x)$ лежит в x_i-м по счету отрезке (при делении отрезка $[x_0, x_0+1[$ на q равных частей) и $\varphi(x)$ не совпадает с его правым концом, а $\varphi(y)$ аналогичным образом лежит в другом, y_1-м отрезке и не совпадает с его правым концом. Поэтому $\varphi(x)$ и $\varphi(y)$ лежат в непересекающихся отрезках и, следовательно, $\varphi(x) \neq \varphi(y)$. Рассуждая таким образом дальше, получим, что из условия $x \neq y$ вытекает $\varphi(x) \neq \varphi(y)$. Итак, $\varphi(\cdot)$ — инъективное отображение.

Пусть дано вещественное число $\lambda \in \boldsymbol{R}$ и $x_0 \rightleftarrows [\lambda]$ — целая часть λ. Если $\lambda = x_0$, то $\lambda = \varphi(x_0, 000 \ldots)$, где дробь $x_0, 000 \in R'$. Пусть λ — не целое число. Если после деления отрезка $\Delta_0 = [x_0; x_0+1]$ на q равных отрезков λ окажется концом одного из этих отрезков, то $\lambda = \varphi(x_0, x_1 00 \ldots)$, где x_1 — номер того отрезка, левым концом которого является λ. Если λ не совпадает с концом ни одного из этих отрезков, то λ принадлежит ровно одному из этих отрезков, номер которого обозначим x_1. Продолжая этот процесс, мы либо остановимся на каком-то шаге, т. е. получим, что $\lambda = \varphi(x_0, x_1 x_2 \ldots x_n 00\ldots)$, либо построим некоторую дробь вида $x = x_0, x_1 x_2 x_3$ из множества $\boldsymbol{Z} \times (F_q)^{\boldsymbol{N}}$. Легко проверить, что такая дробь x лежит в R' и $\varphi(x) = \lambda$. Тем самым доказана сюръективность отображения φ. ■

Так как R' — подмножество декартова произведения $\boldsymbol{Z} \times (F_q)^{\boldsymbol{N}}$, то естественно ожидать, что операции и отношения, определяемые в декартовом произведении, индуцируют аналогичные по свойствам операции и отношения в R'. В свою очередь операции и отношения в декартовом произведении естественно определять на основе уже имеющихся операций и отношений в сомножителях декартова произведения.

Итак, третье определение вещественного числа состоит в том, что вещественным числом называется любая дробь, т. е. любой элемент множества R'. Множество R' определяется без использования вещественных чисел в смысле их первого и второго определений, поэтому третье определение вещественных чисел не зависит от двух предыдущих определений.

Пока не ясно, как при таком подходе к понятию вещественного числа определить обычные операции и отношения в множестве R'. Конечно, можно воспользоваться биекцией $\varphi : R' \to \boldsymbol{R}$ и перенести операции и отношения, определенные в п. 1 и 2, с множества $\boldsymbol{R}$ на множество R', т. е., например, положить:

$$x + y \rightleftarrows (\varphi^{-1}(\varphi(x) + \varphi(y))).$$

Вместо этого мы хотим определить операции и отношения в множестве R', используя именно структуру декартова произведения в R'. В этом и состоит принципиальное отличие третьего подхода к определению вещественного числа от двух первых подходов. Наряду с пополнением метрических пространств (первый подход), пополнением линеалов (второй подход), наш третий подход — перенос структуры с сомножителей на декартово произведение — также имеет важное общематематическое значение. Перейдем к конкретным построениям.

Определение 2. **Пусть $x = x_0, x_1 x_2 \ldots$ и $y = y_0, y_1 y_2 \ldots$ — две различные дроби из R' и $n \rightleftarrows \min\{k \in \boldsymbol{Z}_{\geqslant 0} \mid x_k \neq y_k\}$. Положим $(x < y) \rightleftarrows (x_n < y_n)$. Иными словами, отношение порядка в R' определяется следующим образом:**

если $x_0 < y_0$, то $x_0, x_1 x_2 \ldots < y_0, y_1 y_2 \ldots$;

если $x_0 = y_0$ и $x_1 < y_1$, то $x_0, x_1 x_2 \ldots < y_0, y_1 y_2 \ldots$ и т. д.

Итак, отношение порядка в декартовом произведении индуцируется отношениями порядка в сомножителях этого произведения. Такой порядок называется лексико-графическим. С аналогичным порядком встречаются при расположении слов в словаре в алфавитном порядке. Например, при $q = 10$ выполняется $0{,}19842\ldots < 0{,}19845\ldots$.

Обозначим $0' \rightleftarrows 0, 00 \ldots$. Дробь $x \in R'$ назовем положительной (отрицательной), если $0' < x$ (соответственно $x < 0'$).

Упражнение 8. *Докажите, что:*

а) для любых дробей x и y из R' выполняется одно и только одно из трех соотношений: $x = y$ (т. е. $x_n = y_n$ при всех $n = 0, 1, 2 \ldots$), или $x < y$, или $y < x$;

б) если $x < y$ и $y < z$, то $x < z$;

в) если $1_n \rightleftarrows 0, 0 \ldots 0100 \ldots$ (знак 1 на n-м месте) и при всех $n \in \boldsymbol{N}$ выполняется $0' \leqslant x < 1_n$, то $x = 0'$;

г) если $x = x_0, x_1 x_2 \ldots \in R'$ и $x_{(n)} \rightleftarrows x_0, x_1 x_2 \ldots x_n 00\ldots \in R'$, то x является точной верхней гранью множества $\{x_{(n)} \mid n \in N\}$;

д) вышеуказанная биекция φ строго монотонна, т. е. $(x < y) \Leftrightarrow (\varphi(x) < \varphi(y))$.

Теорема 7. ***Любое непустое ограниченное сверху множество $X \subseteq R'$ имеет в R' точную верхнюю грань.***

▷ Пусть $c = c_0, c_1 c_2 c_3 c_4 \ldots$ — некоторая фиксированная верхняя грань множества X, т. е. для всех $x \in X$ выполняется $x \leqslant c$. Здесь $x \leqslant c$, как всегда, означает $x < c \vee x = c$. Рассмотрим множество

$X_0 \rightleftarrows \{y \in \mathbf{Z} \mid \exists x \in X (x_0 = y)\}$. Тогда $X_0 \subseteq \mathbf{Z}$ и X_0 непусто, так как X непусто, и X_0 ограничено сверху, так как $x_0 \leqslant c_0$ для всех $x \in X$. Следовательно, в X_0 найдется наибольший элемент. Обозначим его b_0. Конечно, $b_0 \in \mathbf{Z}$.

Рассмотрим множество $X_1 \rightleftarrows \{y \in F_q \mid \exists x \in X (x_0 = b_0 \wedge x_1 = y)\}$. Тогда $X_1 \subseteq F_q = \{0, 1, 2, \ldots q-1\}$ и X_1 не пусто, так как найдутся элементы из X, у которых на нулевом месте стоит b_0. Следовательно, в X_1 найдется наибольший относительно обычного порядка в F_q элемент. Обозначим его b_1; конечно $b_1 \in F_q$.

Если знак b_n уже построен, то положим по индукции $b_{n+1} \rightleftarrows \max X_{n+1}$, где $X_{n+1} \rightleftarrows \{y \in F_q \mid \exists x \in X (x_0 = b_0 \wedge \ldots \wedge x_n = b_n \wedge x_{n+1} = y)\}$.

Если не все b_n, начиная с некоторого номера, равны $q-1$, то положим $b \rightleftarrows b_0, b_1 b_2 \ldots \in R'$.

Если это не так, т. е. при некотором $n \in \mathbf{N}$ выполняется $b_{n+1} = b_{n+2} = b_{n+3} = \ldots = q-1$ и $b_n \neq q-1$,

то положим $b \rightleftarrows b_0, b_1 b_2 \ldots b_{n-1} b_n' 00\ldots$, где $b_n' = b_n + 1$.

Докажем, что $b = \sup X$. Рассмотрим сначала случай, когда последовательность $\{b_n\}$ не имеет «хвоста», состоящего сплошь из знаков $(q-1)$. Пусть $x \in X$. Покажем, что $x \leqslant b$. Если $x \neq b$, то положим $n = \min\{k \mid x_k \neq b_k\}$, т. е. $x_0 = b_0$, $x_1 = b_1$, ..., $x_{n-1} = b_{n-1}$, $x_n \neq b_n$.

Тогда $x_n \in X_n$, а по построению $b_n = \max X_n$, т. е. $x_n \leqslant b_n$ и, значит, $x \leqslant b$. Следовательно, b — верхняя грань X. Рассмотрим любую дробь $x \in R'$, меньшую b, и докажем, что x уже не является верхней гранью X.

Пусть $x_0 = b_0$, $x_1 = b_1$, ..., $x_{n-1} = b_{n-1}$, $x_n < b_n$. Так как множество X_n не пусто, то найдется такая дробь $y \in X$, что $y_0 = b_0$, $y_1 = b_1$, ..., $y_{n-1} = b_{n-1}$, $y_n = b_n > x_n$. Значит, $y > x$, т. е. x не является верхней гранью X.

Случай, когда все знаки b_n, начиная с некоторого номера, равны $q-1$, рассматривается аналогично. ■

Конечной дробью назовем такую дробь $x = x_0, x_1 x_2 x_3 \ldots$, $x \in R'$, у которой все знаки x_n, начиная с некоторого места, равны нулю. Множество всех бесконечных (не конечных) дробей обозначим R''. Арифметические операции в R' определяются сначала для конечных дробей, т. е. для элементов из $R' \setminus R''$, а уже затем для бесконечных дробей. При этом используется равенство $x = \sup x_{(n)}$, приведенное в упражнении 8. По определению положим:

$$x + y \rightleftarrows \sup_n (x_{(n)} + y_{(n)}).$$

При этом сумму конечных дробей $x_{(n)}$ и $y_{(n)}$ определим обычным, школьным правилом:

на местах с номером, большим n, у этой суммы стоят нули;

на n-м месте стоит знак $(a_n + b_n) \pmod q$;

на $(n-1)$-м месте стоит знак $(a_{n-1} + b_{n-1}) \pmod q$, если из «предыдущего разряда не переносится единица», т.е. если $a_n + b_n \in \{0, 1, \ldots, q-1\}$; в противном случае, если $a_n + b_n \in \{q, q+1, \ldots, 2q-2\}$, то на $(n-1)$-м месте стоит знак $(a_{n-1} + b_{n-1} + 1) \pmod q$, аналогично на k-м месте стоит или знак $(a_k + b_k) \pmod q$, или знак $(a_k + b_k + 1) \pmod q$; последнее имеет место тогда и только тогда, когда сумма, которая находится на предыдущем, $(k+1)$-м месте, т. е. $a_{k+1} + b_{k+1}$ или $a_{k+1} + b_{k+1} + 1$, превышает $q-1$; здесь $k = 1, \ldots, n$;

наконец, на 0-м месте стоит знак $a_0 + b_0$ или знак $a_0 + b_0 + 1$ в соответствии с результатом сложения цифр на первом месте.

Упражнение 9. *Докажите, что:*

а) $\sup_n (x_{(n)} + y_{(n)})$ *существует;*

б) операция сложения коммутативна в множестве $R' \setminus R''$ *и, следовательно, в множестве* R';

в) множество R' *относительно этой операции сложения образует группу и множество* $R' \setminus R''$ *— ее подгруппу;*

г) ограничение биекции $\varphi: R' \leftrightarrow \mathbf{R}$ *на множество* $Z' \rightleftarrows \{a \in R' \mid 0 = a_1 = a_2 = \ldots\}$ *является изоморфизмом группы* $\langle Z', +, 0' \rangle$ *и группы* $\langle \mathbf{Z}, +, 0 \rangle$;

д) множество всех конечных дробей счетно, а множество всех бесконечных дробей континуально.

Отметим, что в наших обозначениях, например, при $q = 10$ противоположным элементом для дроби 1, 100... является дробь (-2), 900..., а противоположным элементом для дроби 0,2300... является дробь (-1), 7700... . Таким образом, в равенстве

$$-(1{,}100\ldots) = (-2),\ 900\ldots$$

минус в левой части относится ко всей дроби, а минус в правой части относится только к целой части дроби.

Аналогично определяется и умножение дробей:

если $x \geqslant 0$, $y \geqslant 0$, то $x \cdot y \rightleftarrows \sup_n (x_{(n)} \cdot y_{(n)})$,

если $x < 0$, $y < 0$, то $x \cdot y \rightleftarrows \sup_n ((-x_{(n)}) \cdot (-y_{(n)}))$,

если $x < 0$, $y \geqslant 0$, то $x \cdot y \rightleftarrows -\sup_n ((-x_{(n)}) \cdot y_{(n)})$,

если $x \geqslant 0$, $y < 0$, то $x \cdot y \rightleftarrows -\sup_n (x_{(n)} \cdot (-y_{(n)}))$.

При этом для конечных положительных дробей умножение, как и сложение, определяется по обычному правилу:

$$x \cdot y = (x_0, x_1 \ldots x_n\ 00\ldots) \cdot (y_0, y_1 \ldots y_k\ 00\ldots) \rightleftarrows$$
$$\rightleftarrows q^{-k} \cdot (y_k \cdot x + q y_{k-1} \cdot x + \ldots + q^k y_0 \cdot x).$$

Здесь, как всегда, $m \cdot x = \underbrace{x + \ldots + x}_{m \text{ раз}}$ для любых $m \in \mathbf{N}$ и $x \in R'$.

Для нахождения $q^{-k} \cdot (c_0, c_1 c_2 \ldots c_e 0)$ следует записать целое число c_0 в q-ичной системе счисления и перенести запятую на k знаков влево.

Подчеркнем, что строгое доказательство обычных свойств операций сложения и умножения в R', т. е. проверка того, что система $\langle R', +, \cdot, 0', 1', < \rangle$ является линейно упорядоченным полем,

довольно сложно. Например, почему умножение коммутативно и любой ненулевой элемент обратим? Мы не будем останавливаться на этих деталях.

Перейдем к описанию топологии в множестве R'. Здесь возможны два пути. Первый из них заключается в использовании уже определенного порядка в R': интервалом с концами в точках a и b называется множество всех дробей $x \in R'$, лежащих между a и b. Множество $X \subseteq R'$ называется открытым, если каждая его точка входит в X вместе с некоторым содержащим ее интервалом или, эквивалентным образом, если X является объединением какого-то семейства интервалов. Другими словами, топология в R' определяется базой

$$\{\{x \in R' \mid a < x < b\} \mid a,\ b \in R' \text{ и } a < b\}.$$

Эта топология называется порядковой топологией в R'.

Второй путь основан на том, что R' является подмножеством декартова произведения $\boldsymbol{Z} \times (F_q)^{\boldsymbol{N}}$. В других главах уже не раз использовалось то, что во всяком декартовом произведении топологических пространств каноническим образом определяется топология, а именно топология произведения. Напомним, что базой топологии произведения в декартовом произведении $\prod_{\alpha \in A} X_\alpha$, где X_α какие-то топологические пространства, называется семейство всех множеств вида $\prod_{\alpha \in A} \mathcal{O}_\alpha$, где все $\mathcal{O}_\alpha$ открыты в X_α и за исключением конечного их числа все $\mathcal{O}_\alpha$ совпадают с X_α. Если изменить это определение, считая, что все $\mathcal{O}_\alpha$, кроме какого-то одного, совпадают с X_α, то получится определение предбазы топологии произведения.

В случае декартова произведения $\boldsymbol{Z} \times (F_q)^{\boldsymbol{N}}$ все сомножители по определению имеют дискретную топологию, т. е. в $\boldsymbol{Z}$ и F_q по определению любые подмножества открыты. Поэтому в $\boldsymbol{Z} \times (F_q)^{\boldsymbol{N}}$ возникает определенная топология произведения. Она индуцирует в R' топологию, которую мы также назовем топологией произведения в R'. Оказывается, что порядковая топология и топология произведения в R' почти совпадают.

Рассмотрим конкретный пример при $q = 10$. Элемент x из R' лежит в интервале $]0,\ 19842..., 0,\ 19845...[$, если $x = 0,\ 1984\, x_5 x_6 x_7 \ldots$, где x_5 равно знакам 3 или 4, а цифры x_6, x_7, ... произвольны. Эта часть исходного интервала совпадает с произведением

$$\{0\} \times \{1\} \times \{9\} \times \{8\} \times \{4\} \times \{3,\ 4\} \times F_q \times F_q \times \ldots$$

и, следовательно, принадлежит базе топологии произведения в R'. Аналогичным образом легко исчерпать весь исходный интервал объединением множеств, открытых в топологии произведения. Поэтому база порядковой топологии содержится в топологии произведения, т. е. порядковая топология в R' содержится в топологии произведения в R'.

Обратное включение для этих топологий неверно. Например,

при $q = 10$ рассмотрим следующий элемент предбазы топологии произведения в R':

$$\mathcal{O} \rightleftarrows Z \times \{3\} \times F_q \times F_q \times \ldots .$$

Ясно, что x из R' лежит в $\mathcal{O}$ тогда и только тогда, когда на первом месте в x находится знак 3. Поэтому множество $\mathcal{O}$ совпадает с объединением счетного числа полуинтервалов $\bigcup_{n=-\infty}^{+\infty} [n,\ 300...,\ n,\ 400...[$. Следовательно, $\mathcal{O}$ не является открытым множеством в порядковой топологии.

Однако, если из множества R' и соответственно из множества $\mathcal{O}$ исключить все дроби вида $n,\ 300...$, то в оставшемся от R' множестве оставшееся от $\mathcal{O}$ множество будет открыто в топологии, индуцированной порядковой топологией в R'. Этот конкретный пример нетрудно обобщить (сделайте это в качестве упражнения) и в результате получить следующее

Предложение 7. ***В множестве R'' всех бесконечных дробей совпадают топологии, индуцированные порядковой топологией и топологией произведения в R'.***

Определение вещественного числа как q-дроби имеет две принципиальные особенности: 1) оно зависит от произвольного выбора натурального числа $q > 1$; при этом не видно и на самом деле нет оснований предпочесть одно значение q другому; 2) начальные отрезки $x_{(n)} \rightleftarrows x_0,\ x_1 x_2 \ldots x_n 00\ldots$ произвольной дроби $x = x_0,\ x_1 x_2 x_3 \ldots$, конечно, задают ее приближения рациональными числами r_n в том смысле, что $r_n \to \varphi(x)$ при $n \to \infty$, где $n = 1,\ 2,\ \ldots$. Однако эти приближения отнюдь не являются наилучшими в смысле близости дробей x_n к x.

Сформулируем последнее обстоятельство подробнее.

Определение 3. **Рациональное число $\frac{k}{n}$, где $n \in N$, называется наилучшим приближением числа x, если для всех натуральных чисел $m \leqslant n$ и для всех целых чисел s из неравенства $\frac{s}{m} \neq \frac{k}{n}$ следует неравенство**

$$\left| x - \frac{k}{n} \right| < \left| x - \frac{s}{m} \right| .$$

Оказывается, что возможно такое определение вещественных чисел, также основанное на произведении пространств, что, во-первых, оно не зависит ни от каких параметров (вроде q) и, во-вторых, каждое число x в смысле этого определения задается (определяется) как такая последовательность $x_0,\ x_1 x_2 x_3 \ldots\ x_n \ldots$, у которой начальные отрезки $x_{[n]} \rightleftarrows x_0,\ x_1 x_2 \ldots x_n 00\ldots$ являются наилучшими приближениями числа x.

Такой подход к определению вещественного числа рассматривается в курсе «Числовые системы».

§ 4. ОСНОВНЫЕ ПОДХОДЫ К ОПРЕДЕЛЕНИЮ КОМПЛЕКСНЫХ ЧИСЕЛ

Множество комплексных чисел можно определить просто как $\boldsymbol{R}^2$. Тогда абсолютное значение в $\boldsymbol{C} \rightleftarrows \boldsymbol{R}^2$ определяется по формуле $|z| \rightleftarrows \sqrt{(P_1(z))^2 + (P_2(z))^2}$ и сложение — по естественной формуле $z+u \rightleftarrows \langle P_1(z)+P_1(u),\ P_2(z)+P_2(u)\rangle$. Получается метрическая группа.

Более того, $\boldsymbol{R}^2$, как и любое $\boldsymbol{R}^n$, имеет структуру векторного пространства над $\boldsymbol{R}$: умножение на скаляр $\lambda \in \boldsymbol{R}$ определяется по формуле

$$\lambda \cdot \xi \rightleftarrows \langle \lambda \cdot x_1, \ldots, \lambda \cdot x_n \rangle, \text{ где } \xi = \langle x_1, \ldots, x_n \rangle.$$

Вопрос состоит в том, как определить в $\boldsymbol{R}^2$ операцию умножения. Если положить $z \cdot u \rightleftarrows \langle P_1(z) \cdot P_1(u),\ P_2(z) \cdot P_2(u)\rangle$, то $\langle \boldsymbol{R}^2, +, \cdot \rangle$ образует кольцо, но не поле (почему ?). Следовательно, это не то умножение, которое нам «хотелось бы иметь» в $\boldsymbol{C}$. Правильное умножение в $\boldsymbol{C}$ определяется по следующей нетривиальной формуле:

$$z \cdot z_1 \rightleftarrows \langle x \cdot u - y \cdot v,\ x \cdot v + y \cdot u \rangle \rightleftarrows x \cdot z_1 + y \cdot (i \cdot z_1), \qquad (*)$$

где $z = \langle x, y \rangle$, $z_1 = \langle u, v \rangle$ и $i \cdot z_1 \rightleftarrows J \cdot \begin{pmatrix} u \\ v \end{pmatrix}$.

Здесь матрица J по определению равна

$$J \rightleftarrows \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Подчеркнем, что запись $i \cdot z$ по определению обозначает результат умножения матрицы J на вектор $z = \begin{pmatrix} x \\ y \end{pmatrix}$.

Упражнение 10. *Проверьте, что таким образом получается метрическое поле* $\boldsymbol{C} \rightleftarrows \langle \boldsymbol{C};\ +;\ \cdot;\ 0,\ 1;\ |\cdot| \rangle$.

Формула $(*)$ дает ответ на следующий общий вопрос: можно ли при некоторых значениях n превратить множество $\boldsymbol{R}^n$ в векторное пространство над $\boldsymbol{C}$. Например, при $n=2$ получилось, что $\boldsymbol{C} \rightleftarrows \boldsymbol{R}^2$ — одномерное векторное пространство над $\boldsymbol{C}$. Оказывается, что любое четномерное над полем $\boldsymbol{R}$ векторное пространство $\boldsymbol{R}^{2n}$ можно наделить структурой векторного пространства над полем $\boldsymbol{C}$. Для этого определим в $\boldsymbol{R}^{2n}$ умножение на комплексное число $z = \langle x, y \rangle$ по следующей формуле:

$$z \cdot \xi \rightleftarrows x \cdot \xi + y \cdot (i \cdot \xi),$$

в которой $\xi \in \boldsymbol{R}^{2n}$ и $i \cdot \xi \rightleftarrows J_{2n} \cdot \xi$, где

$J_{2n} \rightleftarrows \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}$ — матрица размера $2n \times 2n$, а E — единичная $(n \times n)$ — матрица и 0 — нулевая $(n \times n)$ — матрица.

При этом над полем $\boldsymbol{C}$ векторное пространство $\boldsymbol{R}^{2n}$ имеет размерность n. Такое наделение множества $\boldsymbol{R}^{2n}$ структурой комплексного векторного пространства играет существенную роль, например в квантовой механике.

Упражнение 11. *Проверьте, что указанная структура в множестве* $\boldsymbol{R}^{2n}$ *действительно является комплексным векторным пространством размерности* n *над* $\boldsymbol{C}$.

Когда $n=1$, умножение векторов из $\boldsymbol{R}^{2n}$ на скаляр из $\boldsymbol{C}$ превращается в определенное выше умножение комплексных чисел.

Второе определение $\boldsymbol{C}$, обычно используемое в алгебре, существенно опирается на структуру поля в $\boldsymbol{R}$. Оно состоит в том, что множество комплексных чисел определяется как простое алгебраическое расширение поля $\boldsymbol{R}$ с помощью любого из квадратных корней вида $\sqrt{(-1)}$. Эти два корня обозначаются соответственно i и $-i$.

В главе IV (с. 230) определялось понятие простого алгебраического расширения. Однако оно использовало поле $\boldsymbol{C}$ и поэтому не годится сейчас, когда мы хотим определить само это поле. Приведем новое определение простого алгебраического расширения произвольного поля K, которое не использует поле $\boldsymbol{C}$. Оно эквивалентно определению из главы IV (докажите).

Многочлен называется унитарным, если его старший коэффициент равен 1.

Пусть f — унитарный, неприводимый над полем K многочлен степени >0. Простым алгебраическим расширением поля K с помощью многочлена f называется фактор-кольцо $L \rightleftarrows K[x]/(f)$. Здесь $K[x]$ — кольцо многочленов с коэффициентами из K от одной переменной x и (f) — главный идеал в этом кольце, порожденный многочленом f, т. е. $(f) \rightleftarrows \{g \cdot f \mid g \in K[x]\}$. Операции в L определяются с помощью исходных операций в кольце $K[x]$.

Упражнение 12. *а) Докажите, что такое фактор-кольцо* L *является полем.*

б) В силу неприводимости многочлен f *не имеет корней в поле* K. *Допустим, что найдется такое поле* E, *что* K *— подполе поля* E *и в* E *уравнение* $f(x)=0$ *имеет какой-то корень* x_0. *Тогда для любого многочлена* $g \in K[x]$ *в поле* E *можно вычислить* $g(x_0)$ *и значение* $g(x_0)$ *принадлежит* E. *Докажите, что поле* L *изоморфно подполю поля* E *вида*

$$L_1 \rightleftarrows \{g(x_0) \mid g \in K[x]\}.$$

Следовательно, поля L и L_1 можно не различать.

Поле L_1 можно использовать вместо поля L, его определение проще и нагляднее, однако во многих случаях неясно, как найти необходимое для его определения поле E и даже существует ли оно. Если поле $\boldsymbol{C}$ уже определено и $K \subseteq \boldsymbol{C}$, то в качестве E всегда можно взять поле $\boldsymbol{C}$ (почему?). Конечно $K \subseteq L_1 \subseteq E$.

Итак, второе определение поля $\boldsymbol{C}$ таково: полем комплексных чисел называется простое алгебраическое расширение поля $\boldsymbol{R}$ с помощью многочлена $f(x) \rightleftarrows x^2+1$.

Оба приведенных определения поля $\boldsymbol{C}$, может быть, не вполне удачны в том отношении, что в них множество $\boldsymbol{C}$ определяется с помощью множества $\boldsymbol{R}$, т. е. комплексные числа появляются как вторичный объект по отношению к вещественным числам. Однако

нет оснований рассматривать комплексные числа как понятие, которое основано на вещественных числах.

Поэтому рассмотрим третье определение комплексных чисел: в множестве $\boldsymbol{Q}^2$ определим структуру метрического поля, при этом умножение определим с помощью формулы $(*)$. Затем по теореме 3 определим множество комплексных чисел как пополнение метрического пространства $\boldsymbol{Q}^2$. При таком подходе метрика в пополнении $(\boldsymbol{Q}^2)^\wedge$ возникает сразу, а алгебраические операции в $(\boldsymbol{Q}^2)^\wedge$ определяются как продолжение алгебраических операций в поле $\boldsymbol{Q}^2$ в соответствии с теоремой 5.

§ 5. РОЛЬ АЛГЕБРАИЧЕСКОЙ ЗАМКНУТОСТИ, ЛОКАЛЬНОЙ КОМПАКТНОСТИ И УПОРЯДОЧЕННОСТИ СРЕДИ СВОЙСТВ КОМПЛЕКСНЫХ И ВЕЩЕСТВЕННЫХ ЧИСЕЛ

1. Рассмотрим теперь комплексные и вещественные числа с новой точки зрения: в какой мере определяют эти поля совокупности их наиболее характерных свойств, принимаемых за аксиомы.

Прежде всего возникает вопрос, каковы эти наиболее характерные свойства. Само собой разумеется, что одним из них является то, что $\boldsymbol{C}$ и $\boldsymbol{R}$ поля.

Для поля $\boldsymbol{C}$ наиболее яркими свойствами, характерными именно для него, являются алгебраическая замкнутость и нулевая характеристика (это чисто алгебраические свойства); континуальная мощность (теоретико-множественное свойство); наличие топологии и непрерывность относительно этой топологии алгебраических операций, т. е. свойство быть топологическим полем (алгебро-топологическое свойство); локальная компактность и недискретность топологии (чисто топологические свойства).

Если какое-то поле K, отличное от $\boldsymbol{C}$, обладает этими свойствами, то вряд ли можно считать, что оно «хуже» или «менее важно», чем поле $\boldsymbol{C}$. Таким образом, возникает вопрос, в какой мере определяют поле $\boldsymbol{C}$ перечисленные выше свойства. Кажется важным, хотя бы в обзорном порядке, познакомиться с ответами на этот вопрос, хотя, к сожалению, мы не сможем привести доказательства соответствующих теорем.

Теорема Штейница. ***Всякое алгебраически замкнутое поле K нулевой характеристики и континуальной мощности алгебраически изоморфно $\boldsymbol{C}$.***

Слова «алгебраически изоморфно» означают наличие биекции вида $f:\boldsymbol{C}\to K$, для которой

$$f(z_1+z_2)=f(z_1)+f(z_2),\ f(0)=0,\ f(z_1\cdot z_2)=f(z_1)\cdot f(z_2),\ f(1)=1.$$

Эта теорема не дает оснований отождествить K и $\boldsymbol{C}$, так как биекция f не является «полным изоморфизмом» в том смысле, что она может не сохранять, например, норму, метрику, абсолютное значение или топологию, которые имеются в $\boldsymbol{C}$. Например, по той

причине, что такие понятия в K не определены; в то же время поле $\boldsymbol{C}$ неразрывно связано с его топологическими и метрическими свойствами.

Теорема Штейница, как и другие подобные ей теоремы, имеет замечательное следствие, которое мы сейчас сформулируем.

Язык теории колец по определению содержит знаки $+$, $-$, $\cdot$, 0, 1 и обычные логические знаки $=$, $|$, $(,)$, $\wedge$, $\vee$, $\neg$, $\Rightarrow$, $\Leftrightarrow$, $\exists$, $\forall$.

Переменной называется любое слово в алфавите $\{|\}$. Формула определяется как слово вида $t_1=t_2$, где t_1, t_2 — записи применения к переменным $\underbrace{|\ \ldots\ |}_{n}$ и постоянным 0, 1 операций $+$, $-$, $\cdot$, а также любое слово вида $\varphi_1\wedge\varphi_2$, $\varphi_1\vee\varphi_2$, $\neg\varphi_1$, $\varphi_1\Rightarrow\varphi_2$, $\varphi_1\Leftrightarrow\varphi_2$, $\forall x\varphi_1$, $\exists x\varphi_1$, где φ_1 и φ_2 — формулы, а x — переменная.

Такие t_1, t_2 называются термами. Например, слова $\|+0$, $\|+|||$, $(\|\cdot(\|+|||))-0$ суть термы.

Если формула φ выполняется (истинна) в поле K, то пишут $K\models\varphi$.

Следствие. Если замкнутая формула φ в языке теории колец истинна в $\boldsymbol{C}$, то для любого алгебраически замкнутого поля K нулевой характеристики и континуальной мощности формула φ истинна в K, т. е. $(\boldsymbol{C}\models\varphi)\Rightarrow(K\models\varphi)$.

▷ Следствие вытекает из утверждения: для любой формулы φ со свободными переменными x_1, ..., x_n и любых комплексных чисел z_1, ..., z_n выполняется $(\boldsymbol{C}\models\varphi(z_1,\ \ldots,\ z_n))\Leftrightarrow(K\models\varphi(f(z_1),\ \ldots,\ f(z_n)))$, где f — биекция из теоремы Штейница. Эта эквивалентность проверяется индукцией по длине слова φ. ■

Область применений теоремы Штейница ограничена в силу слишком жестких ограничений на мощность и характеристику поля K. Дело в том, что существует много важных полей, не имеющих континуальной мощности. Например, поле алгебраических чисел — счетно, а с другой стороны, существуют поля мощности, большей континуума. В равной мере существует много важных полей конечной характеристики. Например, поле $F_p\rightleftarrows\{0,\ \ldots,\ p-1\}$ вычетов по простому основанию p имеет характеристику p; поле $F_p[[x]]$, всех формальных степенных рядов с коэффициентами из F_p, имеет характеристику p. Ко всем этим полям теорема Штейница неприменима.

Упражнение 13. *Проверьте последние утверждения.*

Теорема Робинсона. ***Если $\boldsymbol{C}\models\varphi$, то существует такое натуральное число n, что для любого алгебраически замкнутого поля K характеристики нуль или m, где $m\geqslant n$, выполняется $K\models\varphi$.***

Такого типа теоремы называются теоремами переноса. Это удивительные теоремы. Рассмотрим, например, теорему Робинсона. В поле $\boldsymbol{C}$ имеются такие удобные свойства и структуры, как топология, мера, интеграл Коши, дифференциальные уравнения и т. д. Они ни в коей мере не свойственны произвольному полю K. Даже соответствующие понятия нельзя сформулировать для произволь-

ного поля K. Тем не менее, если с использованием всех этих средств доказано $\boldsymbol{C} \models \varphi$, то отсюда автоматически вытекает, что и для поля K выполняется $K \models \varphi$.

Если посмотреть на $\boldsymbol{C}$ с точки зрения наличия в нем нормы $\|\cdot\|$, то нужно обратить внимание на то, что $\boldsymbol{C}$ — полная нормированная алгебра с делением над полем $\boldsymbol{C}$. Напомним, что алгеброй с делением называется любая алгебра A, в которой всякий отличный от нуля элемент имеет обратный, т. е. алгебра, которая как кольцо является телом. Термин «нормированная» означает, что на алгебре A определена функция $\|\cdot\| : A \to \boldsymbol{R}_{\geqslant 0}$, обладающая свойствами:

$$\|x\| = 0 \Leftrightarrow x = 0, \ \forall x \in A;$$
$$\|\lambda \cdot x\| = |\lambda| \cdot \|x\|, \ \forall \lambda \in \boldsymbol{C}, \ x \in A;$$
$$\|x + y\| \leqslant \|x\| + \|y\|, \ \forall x, \ y \in A;$$
$$\|x \cdot y\| \leqslant \|x\| \cdot \|y\|, \ \forall x, \ y \in A; \ \|e\| = 1,$$

где e — единичный элемент в алгебре A. «Полная» означает, что метрическое пространство A с метрикой $\rho(x, y) \rightleftarrows \|x - y\|$ является полным.

Теорема Гельфанда-Мазура. ***Полная нормированная коммутативная алгебра с делением A над полем C изоморфна C.***

Слово «изоморфна» означает здесь существование биекции вида $f : \boldsymbol{C} \leftrightarrow A$, для которой

$$f(z_1 + z_2) = f(z_1) + f(z_2), \ f(0) = 0, \ f(z_1 \cdot z_2) = f(z_1) \cdot f(z_2),$$
$$f\left(\frac{1}{z}\right) = (f(z))^{-1}, \ f(1) = 1, \ \|z\| = \|f(z)\|.$$

Языком теории нормированных алгебр назовем расширение языка теории колец, полученное добавлением нового сорта переменных и нового знака $\|\cdot\|$ и соответствующим расширением понятий терма и формулы. Старые переменные будем обозначать k, а новые — z. Интуитивно переменная k пробегает алгебру A, а переменная z пробегает поле $\boldsymbol{C}$.

Следствие. Если φ — замкнутая формула в языке теории алгебр и $\boldsymbol{C} \models \varphi$, то для любой алгебры A, удовлетворяющей условиям последней теоремы, выполняется $A \models \varphi$.

▷ Для доказательства индукцией по длине слова φ проверяется утверждение:

$$\boldsymbol{C} \models \varphi(k_1, \ldots, k_n, z_1, \ldots, z_n) \Leftrightarrow$$
$$\Leftrightarrow A \models \varphi(f(k_1), \ldots, f(k_n), z_1, \ldots, z_n),$$

где $f : \boldsymbol{C} \leftrightarrow A$ — биекция из теоремы Гельфанда-Мазура. ■

До сих пор мы имели дело только с полями $\boldsymbol{R}$ и $\boldsymbol{C}$. Существует еще несколько полей, весьма похожих на поля $\boldsymbol{R}$ и $\boldsymbol{C}$. Однако удивительное обстоятельство состоит в обратном: похожих на $\boldsymbol{R}$ и $\boldsymbol{C}$ полей совсем немного: кроме самих $\boldsymbol{R}$ и $\boldsymbol{C}$, ровно два! Прежде чем продолжать нашу линию изложения, придется отвлечься и определить эти два поля.

В п. 1 мы определяли поле $\boldsymbol{R}$ как пополнение по Коши метрического пространства $\langle Q, \rho(\cdot, \cdot)\rangle$, где $\rho(x, y) \rightleftarrows |x - y|$. Можно ли в множестве Q задать другие метрики, не похожие на эту «обычную» метрику $|x - y|$? Можно и следующим образом. Пусть p — фиксированное простое число. Обозначим $\nu_p(x)$ число вхождений p в разложение целого числа x на простые множители. Например, $\nu_p(p^n) = n$; $\nu_p(2) = 0$, если $p \neq 2$; если $p = 3$, то $\nu_p(21) = 1$. Число $x = 0$ не имеет разложения на простые множители. Поэтому определим функцию $\nu_p(\cdot)$ в точке 0 равенством $\nu_p(0) \rightleftarrows \infty$. Ясно, что $\nu_p(\cdot) : \boldsymbol{Z} \to \boldsymbol{N} \cup \{\infty\}$. Продолжим функцию $\nu_p(\cdot)$ на $\boldsymbol{Q}$. Для этого положим $\nu_p\left(\frac{n}{m}\right) \rightleftarrows \nu_p(n) - \nu_p(m)$.

Ясно, что получится функция вида $\nu_p(\cdot) : \boldsymbol{Q} \to \boldsymbol{Z} \cup \{\infty\}$. Функция $\nu_p(\cdot)$ называется p-адическим показателем.

Упражнение 14. *Докажите, что:*

1) $\nu_p(1) = 0$; $\nu_p(0) = \infty$;

2) $\nu_p(x + y) = \min\{\nu_p(x), \nu_p(y)\}$;

3) $\nu_p(x \cdot y) = \nu_p(x) + \nu_p(y)$, $\forall x, \ y \in \boldsymbol{Z}$.

(Любая функция вида $\nu(\cdot) : K \to \Gamma \cup \{\infty\}$, где K — произвольное кольцо и Γ — линейно упорядоченная группа (например, $\Gamma = \boldsymbol{Z}^+$), обладающая свойствами 1—3, называется Γ-нормированием кольца K. Таким образом, функция $\nu_p(\cdot)$ является $\boldsymbol{Z}$-нормированием поля $\boldsymbol{Q}$.)

Положим $|x| \rightleftarrows \left(\frac{1}{p}\right)^{\nu_p(x)} : \boldsymbol{Q} \to \boldsymbol{Q}_{\geqslant 0}$ и $\rho_p(x, y) \rightleftarrows |x - y|_p : \boldsymbol{Q}^2 \to \boldsymbol{Q}_{\geqslant 0}$. Функция $|\cdot|_p$ называется p-адическим абсолютным значением, а функция $\rho_p(\cdot, \cdot)$ — p-адической метрикой.

Упражнение 14. а) *Докажите, что функция $|\cdot|_p$ обладает свойствами:*

1) $|x|_p = 0 \Leftrightarrow x = 0$;

2) $|x + y|_p \leqslant |x|_p + |y|_p$ (и даже $|x + y|_p \leqslant \max\{|x|_p, |y|_p\}$);

3) $|x \cdot y|_p = |x|_p \cdot |y|_p$, $\forall x, \ y \in \boldsymbol{Z}$.

Любая функция вида $|\cdot| : K \to \boldsymbol{R}_{\geqslant 0}$, где K — произвольное кольцо, обладающая свойствами 1—3, называется абсолютным значением в кольце K. Таким образом, функция $|\cdot|_p$ является абсолютным значением в поле $\boldsymbol{Q}$.

б) *Проверьте, что $\langle \boldsymbol{Q}, \rho_p \rangle$ — метрическое пространство.*

По теореме 3 метрическое пространство $\langle \boldsymbol{Q}, \rho_p \rangle$ имеет пополнение. Оно называется множеством p-адических чисел, обозначается $\boldsymbol{Q}_p$ и является полным метрическим пространством. По теореме 5 алгебраические операции в поле $\boldsymbol{Q}$ продолжаются до алгебраических операций в $\boldsymbol{Q}_p$, как это было и в случае $\boldsymbol{R}$. Относительно этой метрики и этих операций $\boldsymbol{Q}_p$ метрическое поле.

Поле рациональных дробей $K(t)$ с коэффициентами в каком-то поле K определяется как множество пар $\langle r, q \rangle$ многочленов из $K[t]$, где $q \neq 0$, факторизованное отношением эквивалентности $(\langle r, q \rangle \sim \langle r_1, q_1 \rangle) \rightleftarrows (r \cdot q_1 = q \cdot r_1)$. Последнее равенство понимается над полем K. Операции в этом множестве определяются, как обычно,

с помощью операций в $(K[t])^2$. По традиции элементы поля $K(t)$ записываются в виде $\frac{r}{q}$ и называются дробями.

Определим в кольце $F_p[t]$ $\mathbf{Z}$-нормирование, а именно: $\nu(x)$ по определению равно наибольшему натуральному числу n, для которого t^n можно вынести из x. Например, $\nu(t^n)=n$, $\nu(1)=0$, $\nu(t^2+2t^3+3t^4+t^6)=2$. Ясно, что функция $\nu(\cdot)$ вида $\nu(\cdot): F_p[t]\to N$. Положим $\nu(0)=\infty$. Продолжим функцию $\nu(\cdot)$ на множество $F_p(t)$, а именно $\nu\left(\frac{x}{y}\right) \rightleftarrows \nu(x)-\nu(y)$, где $x,\ y\in F_p[t]$. Ясно, что теперь функция $\nu(\cdot)$ вида $\nu(\cdot): F_p(t)\to Z\cup\{\infty\}$.

Упражнение 16. *Докажите, что:*

а) функция $\nu(\cdot)$ *действительно является* $\mathbf{Z}$-нормированием в поле $F_p(t)$;

б) функция $|x| \rightleftarrows \left(\frac{1}{2}\right)^{\nu(x)}: F_p(t)\to Q_{\geqslant 0}$ *является абсолютным значением в поле* $F_p(t)$;

в) функция $\rho(x,\ y) \rightleftarrows |x-y| : (F_p(t))^2 \to Q_{\geqslant 0}$ *является метрикой в множестве* $F_p(t)$.

По теореме 3 метрическое пространство $\langle F_p(t),\ \rho\rangle$ имеет пополнение. Оно называется полем формальных степенных рядов над полем F_p, обозначается $F_p[[t]]$ и является полным метрическим пространством. По теореме 5 алгебраические операции в поле $F_p(t)$ продолжаются до операций в $F_p[[t]]$. Относительно этой метрики и этих операций $F_p[[t]]$ есть метрическое поле. Любой элемент x из $F_p[[x]]$ можно представить в виде ряда $x=a_{-m}\cdot t^{-m}+a_{-m+1}\cdot t^{-m+1}+\ldots+a_0x^0+a_1t^1+a_2t^2+\ldots+a_it^i+\ldots$, т. е. в виде $x=\sum\limits_{i=n}^{\infty} a_it^i$, где $n\in Z$ и все a_i из F_p (или, что то же самое, все a_i — целые числа от 0 до $p-1$). Конечно, такой ряд больше похож на ряд Лорана, но по традиции его называют степенным рядом. Продолжение функции $\nu(\cdot)$ на $F_p[[t]]$, определяемое как номер первого ненулевого члена в этом ряду, является $\mathbf{Z}$-нормированием поля $F_p[[t]]$. Замечательно, что и любой элемент x из Q_p также можно представить в виде аналогичного ряда $x=a_{-m}p^{-m}+\ldots+a_0p^0+a_1p^1+\ldots+a_ip^i\ldots$, т. е. в виде $x=\sum\limits_{i=n}^{\infty} a_ip^i$, где $n\in Z$ и $0\leqslant a_i\leqslant p-1$. Аналогичное продолжение функции $\nu_p(i)$ на Q_p, определяемое как номер первого ненулевого члена в этом ряду, является нормированием поля Q_p.

Упражнение 16*. *Докажите, что топологии, определяемые метриками в множествах* $\mathbf{R}$, $\mathbf{C}$, $\mathbf{Q}_p$, $F_p[[t]]$, *локально компактны и не дискретны (отличны от дискретной топологии).*

Поля $\mathbf{R}$, $\mathbf{C}$, $\mathbf{Q}_p$, $F_p[[x]]$ называются классическими локально-компактными полями. Любое из этих полей обозначим P. Нас интересуют поля K, которые содержат поле P и являются конечномерными пространствами над P; иными словами, поля K, которые

являются конечномерными расширениями поля P. Учитывая важность таких полей, опишем их явным образом.

Системой вида P^n, где $n\in N$ и P — классическое локально компактное поле, называется набор $\langle P^n,\ +,\ -,\ 0,\ 1,\ \cdot\rangle$, в котором P^n — декартово произведение носителя поля P самого на себя n раз.

Операции сложения и вычитания в системе P^n определяются покоординатно:

$x+y \rightleftarrows (x_1+y_1,\ \ldots,\ x_n+y_n)$ и $(-x) \rightleftarrows (-x_1,\ \ldots,\ -x_n)$, где $x,\ y\in P^n$ и $x=(x_1,\ \ldots,\ x_n)$, $y=(y_1,\ \ldots,\ y_n)$. Аналогично $0 \rightleftarrows (0,\ \ldots,\ 0)$, где в правой части 0 — нуль поля P, и $1 \rightleftarrows (1,\ 0,\ 0,\ \ldots,\ 0)$, где в правой части 1 — единица поля P.

Очевидно, $\langle P^n,\ +,\ -,\ 0,\ \cdot\rangle$ — векторное пространство размерности n над полем P. И наоборот, любое векторное пространство K над полем P размерности n (с фиксированным в нем базисом) отождествляется с P^n. Нетривиальный момент состоит в том, как определить умножение в системе P^n.

При определении умножения $x\cdot y$ в системе P^n (как и вообще в случае алгебры) руководящей идеей является билинейность функции $f(x,\ y) \rightleftarrows x\cdot y$. Выберем в P^n как векторном пространстве над P какой-то базис $\{i_1,\ \ldots,\ i_n\}$, тогда для любых x и y из P^n выполняется $x=x_1\cdot i_1+\ldots+x_n\cdot i_n$ и $y=y_1\cdot i_1+\ldots+y_n\cdot i_n$.

Принимая билинейность операции умножения за аксиому, получим $x\cdot y=\sum\limits_{l,\ s=1}^{n} x_l\cdot x_s(i_l\cdot i_s)$. Таким образом, вопрос о том, как умножать произвольные векторы x и y из P^n, сводится к вопросу, как умножать n векторов базиса на себя. Умножений вида $i_l\cdot i_s$, где $1\leqslant l,\ s\leqslant n$, ровно n^2; чему равно $i_l\cdot i_s$, не вытекает из аксиомы билинейности умножения. Наиболее прямой способ указать, чему равно $i_l\cdot i_s$ — это задать таблицу следующего вида:

	i_1	i_2	. . .	i_s	. . .	i_n
i_1						
i_2						
.						
.						
.						
i_l						
.						
.						
.						
i_n						

в которой на пересечении l-ой строчки и s-го столбца находится какой-то элемент из P^n. Мы хотим, чтобы относительно этого умножения (как и в случае любой алгебры) существовала единица e, т. е. такой элемент, что $e\cdot x=x\cdot e=x$. Поэтому обозначим $i_1=e$ и по определению положим $i_1\cdot i_s=e\cdot i_s=i_s\cdot e=i_s$. Тогда предыдущую таблицу можно уменьшить до таблицы вида:

	i_2	. . .	i_s	. . .	i_n
i_2					
.					
.					
i_l					
.					
.					
i_n					

Конечно, умножение, которое таким образом определено в множестве P^n, зависит от выбора этой таблицы. Поскольку выражение $i_l \cdot i_s$ заменяется на элемент из P^n вида $c_{ls\,1} \cdot i_1 + \ldots + c_{lsn} \cdot i_n$, то удобно записать $i_l \cdot i_s = \sum_{k=1}^{n} c_{lsk} \cdot i_k$, где $c_{lsk} \in P$, и тогда

$$(x_l \cdot x_s)(i_l \cdot i_s) \rightleftarrows \sum_{k=1}^{n} (x_l \cdot x_s) \cdot c_{lsk} \cdot i_k, \text{ где } (x_l \cdot x_s) \cdot c_{lsk} \in P.$$

Элементы c_{lsk}, очевидно, однозначно определяют таблицу 1, и, следовательно, операцию умножения в P^n. Они называются структурными константами системы P^n.

Легко проверить, что при любом выборе структурных констант система P^n (относительно данных операций сложения и умножения) является кольцом и, следовательно, алгеброй над полем P.

$$\forall x \in P^n (x \neq 0 \Rightarrow \exists y \in P^n (x \cdot y = y \cdot x = e)).$$

Иногда эта алгебра оказывается полем, т. е. кроме предыдущего свойства, в ней выполняется $\forall x, y \in P^n (x \cdot y = y \cdot x)$.

Последнее соотношение, конечно, эквивалентно тому, что $i_l \cdot i_s = i_s \cdot i_l$ для всех l и s.

Итак, телом вида P^n назовем любую систему вида $\langle P^n, +, -, 0, 1, \cdot \rangle$, которая является телом.

Полем вида P^n назовем любую систему вида $\langle P^n, +, -, 0, 1, \cdot \rangle$, которая является полем. Повторим, что везде P — одно из четырех классических локально-компактных полей.

Например, $P = \boldsymbol{R}$, $n = 2$ и $c_{221} = -1$, $c_{222} = 0$, $e = 1$. Тогда получается система вида $\boldsymbol{R}^2$, которая, как легко видеть, совпадает с полем $\boldsymbol{C}$ (проверьте) и, следовательно, является полем вида $\boldsymbol{R}^2$. Другой пример. Пусть $P = \boldsymbol{R}$, $n = 4$, $e = 1$ и умножение определяется следующей таблицей:

	i	j	k
i	-1	k	$-j$
j	$-k$	-1	i
k	j	$-i$	-1

Получается система вида $\boldsymbol{R}^4$, в которой умножение не коммутативно. Будет ли она телом?

Если $x = x_1 + x_2 \cdot i + x_3 \cdot j + x_4 \cdot k$, то положим $\bar{x} \rightleftarrows x_1 - x_2 \cdot i - x_3 \cdot j - x_4 \cdot k$.

Очевидно, $x \cdot x = x_1^2 + x_2^2 + x_3^2 + x_4^2 \in \boldsymbol{R}_{\geqslant 0}$.

Положим $|x| \rightleftarrows \sqrt{x \cdot \bar{x}}$ и $x^{-1} \rightleftarrows \frac{\bar{x}}{|x|^2}$.

Если $x \neq 0$, то $|x| \neq 0$ и элемент x определен. Легко проверить, что $x \cdot x^{-1} = x^{-1} \cdot x = 1$.

Итак, эта система является телом вида $\boldsymbol{R}^4$. Его обозначают H и называют телом кватернионов. Легко проверить, что $|\cdot|$ — абсолютное значение в теле H.

Теорема Понтрягина-Ковальского. ***а) Любое локально-компактное недискретное поле K изоморфно либо полю R, либо полю C, либо одному из полей вида Q_p^m, $(F_p[[x]])^n$ для некоторых значений чисел m, p и n.***

б) Если, кроме того, поле K связно, то оно изоморфно R или C.

в) Если K — локально-компактное, недискретное тело, то оно изоморфно или одному из перечисленных в пункте а) полей, или телу кватернионов H.

Слово «изоморфно» означает существование биекции вида $f: K_0 \leftrightarrow K$, где K_0 — тело одного из пяти перечисленных выше видов, для которой:

$$f(k_1 + k_2) = f(k_1) + f(k_2),$$
$$f(0) = 0,$$
$$f(k_1 \cdot k_2) = f(k_2) \cdot f(k_2),$$
$$f(1) = 1$$

и f — гомеоморфизм.

Языком теории топологических колец назовем расширение языка теории колец, которое получается добавлением нового сорта переменных, обозначаемых $\mathcal{O}$, знака $\in$ и соответствующим расширением понятий терма и формулы. Интуитивно переменная k пробегает множество K, и переменная $\mathcal{O}$ пробегает топологию $\mathcal{L}(K)$.

Следствие. Если φ — замкнутая формула в этом языке и $\boldsymbol{R} \models \varphi$ и $\boldsymbol{C} \models \varphi$, то для любого локально-компактного, недискретного и связного поля K выполняется $K \models \varphi$.

▷ Следствие доказывается точно так же, как и предыдущие следствия, с использованием изоморфизма f из теоремы Понтрягина-Ковальского. ■

Для полноты картины напомним теорему Фробениуса, в частном случае известную из курса «Числовые системы».

Теорема Фробениуса.

Пусть K — максимальное упорядоченное поле, L — тело, являющееся конечномерным векторным пространством над полем K, причем K содержится в центре тела L. Тогда либо $L = K$, либо L — поле вида K^2, либо L — тело вида K^4.

Во втором и третьем случаях умножение в K^2 определяется

так же, как соответственно в поле $\boldsymbol{R}^2$ и теле $\boldsymbol{R}^4$. Поэтому L — аналог вещественных или комплексных чисел, или кватернионов.

(Напомним: если K, L — кольца и $K \subseteq L$, то в L фиксируется структура модуля над K, в которой сложение и вычитание определяются как одноименные операции в L и $k \cdot \xi$, где $k \in K$ и $\xi \in L$, определяется как умножение в L двух элементов k и ξ из L.)

Максимальным называется такое упорядоченное поле K, которое не имеет собственного алгебраического расширения, являющегося упорядоченным полем. (Напомним: понятие упорядоченного поля всегда включает условие о линейности порядка.) Например, $\boldsymbol{Q}(\sqrt{2})$ — не максимальное, а $\boldsymbol{R}$ — максимальное поле. Действительно, всякое алгебраическое расширение K поля $\boldsymbol{R}$ содержится в $\boldsymbol{C}$, так как все алгебраические замыкания $\boldsymbol{R}$ изоморфны между собой и изоморфны $\boldsymbol{C}$. Поскольку $K \neq R$, то K содержит какое-то число $z \in (\boldsymbol{C} \setminus \boldsymbol{R})$. Тогда K содержит число i и, следовательно, совпадает с $\boldsymbol{C}$. Но $\boldsymbol{C}$ нельзя упорядочить, так как $i^2 = -1$.

2. Для $\boldsymbol{R}$ наиболее яркими свойствами, характерными именно для этого поля, являются линейная упорядоченность и условная полнота. В частности, абелева группа $\boldsymbol{R}^+$ — линейно упорядоченная и условно полная группа. (Напомним, что отношение порядка $\leqslant$ в множестве X называется линейным, если выполняется условие $\forall x, y \in X (x \leqslant y \vee y \leqslant x)$. Отношение порядка $\leqslant$ в множестве X называется условно полным, если всякое непустое ограниченное сверху или снизу подмножество множества X имеет точную верхнюю или соответственно точную нижнюю грань.

Теорема 9. ***а) Всякая линейно упорядоченная и условно полная абелева группа G изоморфна одной из групп $\{0\}$, $\boldsymbol{Z}^+$, $\boldsymbol{R}^+$. Слово «изоморфно» означает существование биекции вида $f\colon G_0 \leftrightarrow G$, для которой $f(x_1 + x_2) = f(x_1) + f(x_2)$, $f(0) = 0$, $x \leqslant y \Leftrightarrow f(x) \leqslant f(y)$, где G_0 — носитель одной из трех упомянутых групп.***

б) Всякое линейно упорядоченное и условно полное поле K изоморфно полю $\boldsymbol{R}$. Слово «изоморфно» означает существование биекции вида $f(\cdot)\colon K \leftrightarrow \boldsymbol{R}$, сохраняющей операции и отношение порядка в полях K и $\boldsymbol{R}$.

Поскольку нас интересует поле $\boldsymbol{R}$, ограничимся доказательством п. б); п. а) доказывается аналогично и по существу сводится к упражнению 7 из главы I.

▷ Сначала докажем две леммы.

Лемма 6. ***Характеристика любого упорядоченного поля K равна 0, и K содержит подполе, изоморфное упорядоченному полю $\boldsymbol{Q}$.***

▷ В поле выполняется свойство $\forall k \neq 0 \ (k^2 > 0)$. Действительно, $k \geqslant 0$ или $k \leqslant 0$. В первом случае $k^2 \geqslant 0$. Во втором случае $0 \leqslant -k$, так что $(-k)^2 \geqslant 0$ и $(k)^2 = (-1) \cdot (-1) k \cdot k = (-k)^2 \geqslant 0$, поскольку в любом поле $(-1)^2 = 1$. В частности, $1 = 1^2 > 0$. Поэтому $n \cdot 1 \rightleftarrows \underbrace{1 + \ldots + 1}_{n \text{ раз}} > 0$, так что K — поле характеристики 0.

Функция f_0, по определению равная $f_0(n) \rightleftarrows n \cdot 1$ — биекция между

множеством $\boldsymbol{Z}$ и частью $R(f_0)$ поля K. Положим $f\left(\frac{n}{m}\right) = (n \cdot 1)(m \cdot 1)^{-1}$. Элемент $f\left(\frac{n}{m}\right)$ еще обозначают $\frac{n}{m} \cdot 1$. Эта функция f является продолжением функции f_0. Она — биекция между $\boldsymbol{Q}$ и частью $R(f)$ поля K. Функция f является гомоморфизмом и, следовательно, алгебраическим изоморфизмом полей $\boldsymbol{Q}$ и $R(f)$. Кроме того, функция f_0 сохраняет порядок. Отсюда ясно, что функция f сохраняет порядок. Итак, поле $R(f)$ изоморфно $\boldsymbol{Q}$ и $R(f) \subseteq K$.

Конечно, множество K не обязательно содержит множество $\boldsymbol{Q}$, т. е. неверно, что $\boldsymbol{Q} \subseteq K$, но в K имеется часть $R(f)$, которая биективна с $\boldsymbol{Q}$ посредством функции f. Как обычно, изоморфные объекты не различаются, и поэтому в дальнейшем мы будем обозначать $\boldsymbol{Q} \rightleftarrows R(f)$ или, что то же самое, предполагать $\boldsymbol{Q} \subseteq K$. ■

Лемма 7. а) ***Линейно упорядоченная условно полная абелева группа (например, $\boldsymbol{R}^+$) архимедова.***

б) ***Архимедовость поля K влечет порядковую плотность множества $\boldsymbol{Q}$ в упорядоченном множестве K.***

Слова «порядковая плотность» означают:

$$\forall k_1, k_2 \in K \exists r \in \boldsymbol{Q} (k_1 < k_2 \Rightarrow k_1 < r < k_2).$$

▷ а) Пусть $x > 0$ и допустим, что множество $\{nx \mid n \in \boldsymbol{N}\}$ ограничено каким-то элементом. Тогда существует $y_0 \rightleftarrows \sup\{n \cdot x\}$. Поэтому $y_0 \geqslant (n+1) \cdot x$ и $\forall n ((y_0 - x) \geqslant n \cdot x)$, что противоречит определению y_0.

б) Пусть $k_1 < k_2$. Тогда $k = k_2 - k_1 > 0$ и существует n, для которого $n \cdot k > 1$; поэтому $0 < \frac{1}{n} < k$. ■

Доказательство пункта б) теоремы 9.

По лемме 7 любой элемент $k \in K$ является точной верхней гранью множества всех рациональных чисел $\leqslant k$. Обозначим это множество X_k, $X_k \subseteq \boldsymbol{Q} \subseteq \boldsymbol{R}$. Оно ограничено сверху в $\boldsymbol{R}$; обозначим λ_k его точную верхнюю грань в множестве $\boldsymbol{R}$. Положим $f(k) \rightleftarrows \lambda_k$, $f\colon K \to \boldsymbol{R}$.

Функция f — аддитивный гомоморфизм. Его ядро тривиально. Поэтому функция f инъективна. Положительные элементы поля K переходят в положительные элементы поля $\boldsymbol{R}$. Поэтому функция f сохраняет порядок.

В силу условной полноты поля K функция f сюръективна.

Для k, $l > 0$ выполняется $f(k \cdot l) = f(k) \cdot f(l)$. Поэтому для любых k, $l \in K$ выполняется такое же соотношение. Итак, f — изоморфизм. ■

Упражнение 18. *В кольце $\boldsymbol{Q}[x]$ определим порядок, при котором элемент f из $\boldsymbol{Q}[x]$ положителен, если его старший коэффициент положителен. Проверьте, что таким образом $\boldsymbol{Q}[x]$ становится упорядоченным кольцом. Порядок в нем продолжается на его поле частных $\boldsymbol{Q}(x)$ по формуле $\left(\frac{f}{q} \geqslant 0\right) \rightleftarrows (f \cdot q \geqslant 0)$. Поле $\boldsymbol{Q}(x)$ — неархимедово упорядоченное поле (так как в нем элемент x бесконечно большой). Докажите.*

Языком теории упорядоченных колец назовем расширение языка теории колец, которое получается добавлением нового знака $\leqslant$ и соответствующим расширением понятий терма и формулы.

Следствие. Если φ — замкнутая формула в этом языке и $\boldsymbol{R} \models \varphi$, то для любого линейно упорядоченного и условно полного поля K выполняется $K \models \varphi$.

В этом пункте не использовалось то важное обстоятельство, что $\boldsymbol{R}$ метрическое пространство и в нем определено даже абсолютное значение. Однако отношение порядка уже определяет абсолютное значение в $\boldsymbol{R}$ по формуле $|x| \rightleftharpoons \max\{x, -x\}$. Поэтому следствие остается верным и в том случае, если добавить к языку еще один новый знак $|\cdot|$.

Отметим, что в любом линеале L можно определить функцию $|x| \rightleftharpoons \max\{x, -x\}$, которую иногда называют абстрактной нормой в L. Ее свойства похожи на свойства абсолютного значения $|\cdot|$ в линеале $\boldsymbol{R}$.

§ 6. СВЯЗЬ ПОЛЕЙ ВЕЩЕСТВЕННЫХ И КОМПЛЕКСНЫХ ЧИСЕЛ. ПРОДОЛЖЕНИЕ ЛИНЕЙНОГО ПОРЯДКА С ПОЛЯ НА ЕГО АЛГЕБРАИЧЕСКОЕ РАСШИРЕНИЕ И МЕТРИЧЕСКОЕ ПОПОЛНЕНИЕ

1. Мы привыкли, что поля $\boldsymbol{R}$ и $\boldsymbol{C}$ тесно связаны между собой, например, соотношением $\boldsymbol{C} = \boldsymbol{R}(i)$, и свойства этих полей своеобразно контрастируют между собой: $\boldsymbol{R}$ упорядоченное поле и даже максимальное упорядоченное поле, но оно алгебраически не замкнуто, а $\boldsymbol{C}$ не упорядочиваемое, но зато алгебраически замкнутое поле. Является ли такая связь проявлением общей закономерности? На этот вопрос положительно отвечает следующая

Теорема Эйлера-Лагранжа. ***Пусть K — упорядоченное поле. Следующие свойства эквивалентны:***

а) поле K — максимально;

б) поле $K(i)$ — алгебраически замкнуто, где i один из квадратных корней из элемента (-1) поля K;

в) всякий положительный элемент поля K является квадратом и каждый многочлен нечетной степени над полем K имеет корень в K.

2. Для любого упорядоченного поля выполняется свойство

$$\forall k \neq 0 \ (k^2 > 0).$$

Действительно, если $k > 0$, то, умножая на k, получим $k^2 \geqslant 0$; если $k < 0$, то, прибавляя $-k$, получим $0 < -k$ и $0 \leqslant k^2$, но равенство $k^2 = 0$ влечет $k = 0$; отсюда получаем проверяемое свойство.

Когда рассматривают какое-то поле K, то обычно стремятся определить на его носителе такое отношение линейного порядка, при котором оно будет упорядоченным полем. Причина этого понятна: упорядоченные поля легче изучать. Особенно естественной

кажется такая попытка в том случае, если поле K является расширением (в алгебраическом смысле) или пополнением (в метрическом смысле) какого-то упорядоченного поля K_0, где $K_0 \subset K$. Тогда порядок в K_0 хотелось бы продолжить до порядка в K. Однако это не всегда возможно.

Например, порядок в поле $\boldsymbol{Q}$ продолжается до линейного порядка в поле $\boldsymbol{R}$, но порядок в $\boldsymbol{R}$ нельзя продолжить до линейного порядка в $\boldsymbol{C}$. Действительно, если бы в $\boldsymbol{C}$ существовал какой угодно (пусть даже не продолжающий порядок в $\boldsymbol{R}$) линейный порядок, согласованный со структурой поля в $\boldsymbol{C}$, т. е. относительно него $\boldsymbol{C}$ являлось упорядоченным полем, то мы получили бы противоречие: $i^2 > 0$, $-1 > 0$, $0 \geqslant 1$, $0 > 1$, $1 = 1^2 > 0$, $1 > 1$. При этом не помогает то обстоятельство, что $\boldsymbol{C}$ — простое алгебраическое расширение $\boldsymbol{R}$ вида $R(\sqrt{-1})$, хотя линейный порядок продолжается с $\boldsymbol{Q}$ на его простое алгебраическое расширение $\boldsymbol{Q}(\sqrt{2})$.

Аналогичным образом дело обстоит в случае пополнения даже такого простого поля, как $\boldsymbol{Q}$. Существуют такие пополнения поля $\boldsymbol{Q}$ (относительно двух разных метрик), что на одно из них продолжается порядок в $\boldsymbol{Q}$, а на другое не продолжается. А именно, линейный порядок в поле $\boldsymbol{Q}$ продолжается на его пополнение относительно обычной метрики $\rho(x, y) \rightleftharpoons |x - y|$ и обычного абсолютного значения $|\cdot|$ в $\boldsymbol{Q}$; этим пополнением является $\boldsymbol{R}$. Линейный порядок в кольце $\boldsymbol{Z}$ продолжается на пополнение $\boldsymbol{Z}$ относительно той же метрики и того же абсолютного значения, что и в предыдущем примере. В этом случае пополнение совпадает с исходным $\boldsymbol{Z}$. Однако линейный порядок в $\boldsymbol{Q}$ не продолжается на его пополнение относительно p-адической метрики и p-адического абсолютного значения, т. е. не продолжается на $\boldsymbol{Q}_p$, и аналогичным образом линейный порядок в $\boldsymbol{Z}$ не продолжается на его пополнение относительно p-адического абсолютного значения $|\cdot|_p$, т. е. на $\boldsymbol{Z}_p$, например, при $p = 5$. В этом случае пополнение $\boldsymbol{Z}_p$ не совпадает с исходным $\boldsymbol{Z}$.

Поскольку $\boldsymbol{Z}_5 \subseteq \boldsymbol{Q}_5$, то для поля $\boldsymbol{Q}_5$ это утверждение вытекает из более сильного утверждения: кольцо $\boldsymbol{Z}_5$ не допускает линейного упорядочения, согласованного с операциями в нем. Последнее вытекает из следующей теоремы.

Теорема 10. ***В кольце $\boldsymbol{Z}_5$*** существует элемент ε, ***для которого*** $\varepsilon^2 = -1$.

Отсюда точно так же, как и в случае $i^2 = -1$, получаем невозможность линейного упорядочения, согласованного с операциями в кольце $\boldsymbol{Z}_5$.

Итак, простые алгебраические расширения $\boldsymbol{Q}(\sqrt{2})$ и $\boldsymbol{R}(\sqrt{-2})$ соответственно допускают и не допускают продолжение порядка с исходного поля. Аналогично метрические пополнения $\boldsymbol{R}$ и $\boldsymbol{Q}_5$ соответственно допускают и не допускают продолжения порядка с исходного поля $\boldsymbol{Q}$.

Оставшаяся часть этого пункта посвящена доказательству теоремы 10, при первом чтении ее можно пропустить.

Прежде чем доказывать теорему 10, отметим, что даже такие

привычные поля, как F_p, могут обладать неожиданными свойствами. Например, многочлены x^p и x над полем F_p, как функции на этом поле, тождественно равны, так как по теореме Ферма $\forall x \in F_p (x^p \equiv x (\mathrm{mod}\, p))$, хотя как многочлены они различны. Таким образом, многочлены, совпадающие как функции, могут иметь разные степени и разные коэффициенты при соответствующих степенях. Такое невозможно в поле характеристики нуль.

Поскольку $(-1) \equiv (p-1) (\mathrm{mod}\, p)$, то для одних p в поле F_p существует корень $\sqrt{-1}$, а для других p в F_p такого корня не существует. Например, в F_5 выполняется $\sqrt{-1} = 2$, так как $-1 \equiv 2^2 (\mathrm{mod}\ 5)$, но в F_7 уравнение $x^2 = -1$ не имеет решения, т. е. не существует $\sqrt{-1}$. Как по простому числу p узнать, существует ли в поле F_p корень $\sqrt{-1}$? Пусть $m \in \boldsymbol{N}$; положим:

$$\left(\frac{-1}{m}\right) \rightleftarrows \begin{cases} +1, \text{ если уравнение } x^2 \equiv -1 (\mathrm{mod}\, m) \\ \quad \text{имеет решение в кольце } F_m, \\ -1, \text{ в противном случае} \end{cases}$$

или общее:

$$\left(\frac{a}{m}\right) \rightleftarrows \begin{cases} +1, \text{ если уравнение } x^2 \equiv a (\mathrm{mod}\, m) \\ \quad \text{имеет решение в кольце } F_m, \\ -1, \text{ в противном случае,} \end{cases}$$

где $a \in \boldsymbol{Z}$.

Обозначение $\left(\frac{a}{m}\right)$ называется *символом Якоби*. Если m — простое число, то оно называется *символом Лежандра*.

В курсе теории чисел доказывается, что $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$, где $p \neq 2$ и p простое число.

Следовательно, в поле F_p существует корень $\sqrt{-1}$ в том и только в том случае, если $\frac{p-1}{2}$ четное число.

Отсюда мы снова получаем, что в F_7 (и в F_3) корень $\sqrt{-1}$ не существует, а в F_5 существует.

Если существует решение уравнения

$$x^2 \equiv a (\mathrm{mod}\, m),$$

то a называется квадратичным вычетом по модулю $m \in \boldsymbol{N}$.

Лемма 8. ***Пусть $p \neq 2$ и $x_0 \not\equiv 0 (\mathrm{mod}\, p)$.***

Число $\alpha = \sum_{i=0}^{\infty} x_i \cdot p^i$ (где $0 \leqslant x_i \leqslant p-1$) является квадратом какого-то числа в $\boldsymbol{Z}_p$ в том и только в том случае, если число x_0 — квадратичный вычет по модулю p.

Отсюда сразу вытекает доказательство теоремы 10, так как в поле F_5 элемент -1 является квадратом и, следовательно, по лемме 8 в кольце Z_5 элемент $-1 = (-1) \cdot p^0 + 0 \cdot p^1 + 0 \cdot p^2 + \ldots$ также является квадратом. (Отметим, что элементы (-1) в поле F_5 и кольце Z_5 совсем разные.)

▷ Осталось проверить лемму 8. Пусть коэффициент x_0 является квадратичным вычетом по модулю p, т. е. существует число x, для которого $x_0 \equiv x^2 (\mathrm{mod}\, p)$. Положим $f(\gamma) \rightleftarrows (\gamma)^2 - \alpha$, где $\gamma \in \boldsymbol{Z}_p$. Найдем $\beta \in \boldsymbol{Z}_p$, для которого $f(\beta) = 0$. Для этого построим фундаментальную последовательность $\{\beta_n\}$ в $\boldsymbol{Z}_p$, для которой $\beta_n \equiv x (\mathrm{mod}\, p)$ и $f(\beta_n) \equiv 0 (\mathrm{mod}\, p^{n+1})$ для всех $n \geqslant 0$. При $n = 0$ положим $\beta_0 = x$. При этом $\gamma = f(\beta_0) = x^2 - \alpha$ является последовательностью с нулевым начальным членом.

Поскольку $\nu_p(\gamma) \geqslant 1$ и $\gamma = p^{\nu_p(\gamma)} \cdot \gamma_0$, то $\frac{\gamma}{p} = p^{\nu_p(\gamma)-1} \cdot \gamma_0 \in \boldsymbol{Z}_p$, т. е. p делит $f(\beta_0)$.

Допустим, уже построены элементы $\beta_0, \ldots, \beta_{n-1}$.

Разложим функцию $f(\gamma) = f(\beta_{n-1}) + 2 \cdot \beta_{n-1} \cdot (\gamma - \beta_{n-1}) + 2 \cdot (\gamma - \beta_{n-1})^2$ по степеням $(\gamma - \beta_{n-1})$, где $f(\beta_{n-1}) = p^n \cdot \gamma_0$.

Здесь $2 \cdot \beta_{n-1}$ не делится на p, так как иначе x делилось бы на p и x_0 делилось бы на p^2, т. е. $x_0 \equiv 0 (\mathrm{mod}\, p)$. Поэтому $2 \cdot \beta_{n-1}$ обратим в кольце $\boldsymbol{Z}_p$. Следовательно,

$$f(\beta_{n-1} + y \cdot p^n) = p^n \cdot \gamma_0 + 2 \cdot \beta_{n-1} \cdot y \cdot p^n + 2 \cdot y^2 \cdot p^{2n} = p^n \cdot (\gamma_0 + 2\beta_{n-1} y) + 2 \cdot p^{2n} \cdot y^2$$

и можно выбрать такое $y \in \boldsymbol{Z}_p$, что

$$\gamma_0 + 2\beta_{n-1} \cdot y \equiv 0 (\mathrm{mod}\ p).$$

Тогда $f(\beta_{n-1} + y \cdot p^n) \equiv 2 \cdot p^{2n} \cdot y^2 \equiv 0 (\mathrm{mod}\, p^{n+1})$.

Положим $\beta_n \rightleftarrows \beta_{n-1} + y \cdot p^n$.

Первое из условий, которые мы хотим получить, выполняется:

$$\beta_n - x = (\beta_{n-1} - x) + y \cdot p^n \equiv 0 (\mathrm{mod}\, p),$$

а второе условие очевидно.

Последовательность $\{\beta_n\}$ фундаментальна, так как

$$\nu_p(\beta_n - \beta_{n-1}) \geqslant n \text{ и } \nu_p(\alpha + \beta) \geqslant \min\{\nu_p(\alpha), \nu_p(\beta)\}.$$

В силу полноты $\boldsymbol{Z}_p$ предел этой последовательности существует в $\boldsymbol{Z}_p$. Обозначим его β. По непрерывности многочлена получаем $f(\beta) = \lim f(\beta_n)$ и $\nu_p(f(\beta_n)) \to \infty$, поэтому $\nu_p(f(\beta)) = \infty$ и $f(\beta) = 0$.

Теперь докажем лемму в обратную сторону: пусть число α является квадратом, т. е. $\alpha = \beta^2$. Тогда, в частности, $x_0 \equiv y_0^2 (\mathrm{mod}\, p)$. ■

ПРИЛОЖЕНИЕ 1
(к главе I)

1. Группы, изоморфные прямой и окружности. Цель этого приложения состоит в доказательстве теоремы 3, которая утверждает, что любая связная топологическая группа G, нейтральный элемент e которой имеет окрестность, гомеоморфную какому-нибудь открытому интервалу на прямой $\boldsymbol{R}$, изоморфна топологической группе $\boldsymbol{R}^+$ или топологической группе T. Поскольку обычно все математические объекты рассматриваются с точностью до изоморфизма, то эта теорема говорит о том, что существуют ровно две связные топологические группы, единичные элементы которых имеют окрестности, гомеоморфные интервалу на прямой $\boldsymbol{R}$. Более того, эта теорема конкретно указывает эти две группы.

В основном тексте главы мы встречались с такими конкретными топологическими группами, как $\boldsymbol{R}^+$, $\boldsymbol{R}^*$, T, S, An, O_2, SO_2. Сейчас нам потребуются некоторые свойства произвольной топологической группы G. Условимся в этом дополнении записывать групповую операцию мультипликативно.

Предложение 1. а) ***Если g_0 некоторый фиксированный элемент топологической группы G, то отображения $g\mapsto g\cdot g_0$, $g\mapsto g_0 g$ и $g\mapsto g^{-1}$ являются гомеоморфизмами G на себя.***

б) ***Пусть g_0 — некоторый фиксированный элемент топологической группы G, а F — замкнутое, $\mathcal{O}$ — открытое и X произвольное подмножество G. Тогда множества $F\cdot g_0 \rightleftarrows \{g\cdot g_0 \mid g\in F\}$, $g_0\cdot F \rightleftarrows \{g_0\cdot g \mid g\in F\}$, $F^{-1}\rightleftarrows\{g^{-1}\mid g\in F\}$ замкнуты в G, а множества $\mathcal{O}\cdot X$, $X\cdot\mathcal{O}$ и $\mathcal{O}^{-1}$ открыты в G.***

в) ***Гомоморфизм ψ из топологической группы G_1 в топологическую группу G_2 непрерывен в том и только в том случае, если он непрерывен в одной точке его области определения — единице e_1 группы G_1.***

▷ а) Пусть $\varphi(g)\rightleftarrows g\cdot g_0$ для всех $g\in G$. Из равенства $\varphi(g_1)=\varphi(g_2)$ следует, что $g_1=\varphi(g_1)g_0^{-1}=\varphi(g_0)\cdot g_0^{-1}=g_2$, т. е. φ инъективное отображение G в себя. Кроме того, $\varphi(gg_0^{-1})=(gg_0^{-1})g_0=g$, что доказывает сюръективность отображения φ.

Докажем непрерывность отображения φ в произвольной точке $g\in G$. Для любой окрестности $\mathcal{O}$ точки $\varphi(g)=g\cdot g_0$ в силу непрерывности операции умножения в G найдутся окрестность $\mathcal{O}_1$ точки g и окрестность $\mathcal{O}_2$ точки g_0, для которых $\mathcal{O}_1\cdot\mathcal{O}_2\subseteq\mathcal{O}$. Поэтому, в частности, $\varphi(\mathcal{O}_1)=\mathcal{O}_1 g_0\subset\mathcal{O}_1\mathcal{O}_2\subseteq\mathcal{O}$. Итак, для любой окрестности $\mathcal{O}$ точки $\varphi(g)$ найдена окрестность $\mathcal{O}_1$ точки g, образ которой при отображении φ содержится в $\mathcal{O}$; это доказывает непрерывность отображения φ.

Докажем непрерывность обратного к φ отображения φ^{-1}. Так как $\varphi^{-1}(g)=g\cdot g_0^{-1}$ для всех $g\in G$, то φ^{-1} отличается от φ заменой g_0 на g_0^{-1}. В силу произвольности g_0 можно считать, что g_0 и есть g_0^{-1}. По доказанному выше получаем, что φ^{-1} непрерывное отображение.

Гомеоморфность отображений $g\mapsto g\cdot g$ и $g\mapsto g^{-1}$ доказывается аналогично.

б) Рассмотрим прообраз замкнутого множества F при отображении $g\to g\cdot g_0^{-1}$, которое, как доказано в пункте а), непрерывно. Нетрудно видеть, что этот прообраз является множеством $F\cdot g_0=\{g\cdot g_0 \mid g\in F\}$. Но прообраз замкнутого множества при непрерывном отображении замкнут. Поэтому замкнуто и множество $F\cdot g_0$.

Аналогично доказывается замкнутость множеств $g_0\cdot F$ и F^{-1} и открытость множеств $\mathcal{O}\cdot g_0$, $g_0\cdot\mathcal{O}$, $\mathcal{O}^{-1}$, где $\mathcal{O}$ открытое подмножество в G. Тогда множества $\mathcal{O}\cdot X\rightleftarrows\{g\cdot g_0 \mid g\in\mathcal{O},\ g_0\in X\}=\bigcup\limits_{g_0\in X}\mathcal{O}\cdot g_0$ и $X\cdot\mathcal{O}=\bigcup\limits_{g_0\in X} g_0\cdot\mathcal{O}$ открыты как объединения открытых множеств.

в) Если гомоморфизм $\psi: G_1\to G_2$ непрерывен в каждой точке $g\in G_1$, то он, в частности, непрерывен и в единице $e_1\in G_1$. Пусть наоборот, гомоморфизм $\psi: G_1\to G_2$ непрерывен в единице e_1 группы G_1. Докажем, что функция ψ непрерывна в каждой точке $g_1\in G_1$. Для любой окрестности $\mathcal{O}$ точки $g_2\rightleftarrows\psi(g_1)\in G_2$ рассмотрим множество $g_2^{-1}\cdot\mathcal{O}$, которое по пункту б) является окрестностью единицы e_2 в группе G_2. В силу непрерывности отображения ψ получаем, что найдется такая окрестность $\mathcal{O}_1$ единицы e_1 в группе G_1, что $\psi(\mathcal{O}_1)\subseteq g_2^{-1}\cdot\mathcal{O}$. Рассмотрим множество $g_1\mathcal{O}_1$, которое по пункту б) является окрестностью точки g_1 в группе G_1. Так как отображение ψ гомоморфно, то $\psi(g_1\mathcal{O}_1)\subseteq\psi(g_1)\psi(\mathcal{O}_1)$ и, следовательно,

$$\psi(g_1\mathcal{O}_1)\subseteq\psi(g_1)\psi(\mathcal{O}_1)=g_2\psi(\mathcal{O}_1)\subseteq g_2(g_2^{-1}\mathcal{O})=\mathcal{O}.$$

Итак, для любой окрестности $\mathcal{O}$ точки $\psi(g_1)$ найдется такая окрестность $g_1\mathcal{O}_1$ точки g_1, что $\psi(g_1\mathcal{O}_1)\subseteq\mathcal{O}$. Это доказывает непрерывность функции ψ (см. рис. 94). ■

Упражнение 1. *Докажите, что:*

а) *если множество $\mathcal{O}$ открыто в топологической группе G, то в*



Рис. 94

ней открыты множества $\mathcal{O}^2 \rightleftarrows \mathcal{O}\cdot\mathcal{O} = \{g\cdot g_1 \mid g, g_1 \in \mathcal{O}\}$, $\mathcal{O}^3$, ..., $\mathcal{O}^n$ *и т. д. Опишите вид таких множеств в случае, когда* $G = \boldsymbol{R}^*$ *и* $\mathcal{O} = \left]-\frac{1}{2}; 2\right[$;

б) *если* g_1, ..., g_k — *некоторые элементы в топологической группе* G *и* $g \rightleftarrows g_1^{n_1} g_2^{n_2} \ldots g_k^{n_k}$, *где* $n_i \in \boldsymbol{Z}$, *то для любой окрестности* $\mathcal{O}$ *элемента* g *найдутся такие окрестности* $\mathcal{O}_1$, ..., $\mathcal{O}_k$ *элементов* g_1, ..., g_k, *что* $\mathcal{O}_1^{n_1}\mathcal{O}_2^{n_2}\ldots\mathcal{O}_k^{n_k} \subseteq \mathcal{O}$;

в) *для того чтобы любое одноточечное подмножество топологической группы* G *было замкнуто в ней, необходимо и достаточно, чтобы для любых элементов* g_1 *и* g_2 *из* G *нашлись такие открытые множества* $\mathcal{O}_1$ *и* $\mathcal{O}_2$, *что* $g_1 \in \mathcal{O}_1$, $g_2 \in \mathcal{O}_2$ *и* $\mathcal{O}_1 \cap \mathcal{O}_2 = \varnothing$; $g_1 \neq g_2$.

Указание. Для окрестностей V_1 и V_2 элементов g_1 и g_2, таких, что $g_2 \notin V_1$ и $g_1 \notin V_2$, рассмотрите окрестность $\mathcal{O}$ единицы e, для которой $\mathcal{O}\mathcal{O}^{-1} \subseteq (g_1^{-1}V_1) \cap (g_2^{-1}V_2)$; положите $\mathcal{O}_1 \rightleftarrows g_1\mathcal{O}$ и $\mathcal{O}_2 \rightleftarrows g_2\mathcal{O}$.

Напомним, что топологическое пространство называется *отделимым (хаусдорфовым)*, если для любых двух точек x и y, $x \neq y$ найдутся окрестности $\mathcal{O}_x$ и $\mathcal{O}_y$ этих точек $x \in \mathcal{O}_x$ и $y \in \mathcal{O}_y$, которые не пересекаются между собой. Для топологической группы по упражнению 1, в это условие эквивалентно замкнутости любого одноточечного подмножества. Всюду ниже рассматриваются только отделимые топологические группы и специальное упоминание об отделимости опускается.

Мы разобьем доказательство теоремы 3 на три шага.

Определение 1. **Отображение $f: X \to Y$ топологического пространства X в топологическое пространство Y называется открытым, если образ любого открытого в X множества при отображении f открыт в Y.**

Лемма 1. ***Пусть $f: \boldsymbol{R}^+ \to G$ открытый непрерывный гомоморфизм топологической группы $\boldsymbol{R}^+$ на топологическую группу G, $G \neq \{e\}$. Тогда G изоморфна $\boldsymbol{R}^+$ или T.***

▷ Рассмотрим $\operatorname{Ker} f$ — ядро гомоморфизма f. Множество $\operatorname{Ker} f \rightleftarrows f^{-1}(\{e\})$ замкнуто в $\boldsymbol{R}^+$, так как f непрерывно и множество $\{e\}$ замкнуто в G. Кроме того, $\operatorname{Ker} f$, как и ядро любого гомоморфизма, является подгруппой в $\boldsymbol{R}^+$. Но замкнутые подгруппы в $\boldsymbol{R}^+$ уже описаны нами в упражнении 8 главы I, § 5, поэтому $\operatorname{Ker} f = \boldsymbol{R}^+$, или $\operatorname{Ker} f = \{0\}$, или $\operatorname{Ker} f = \{n\cdot a \mid n \in \boldsymbol{Z}\}$ при некотором $a > 0$. Случай $\operatorname{Ker} f = \boldsymbol{R}^+$ невозможен, так как иначе получаем $f(\boldsymbol{R}^+) = \{e\}$, а это противоречит условиям $f(\boldsymbol{R}^+) = G$ и $G \neq \{e\}$.

Пусть $\operatorname{Ker} f = \{0\}$. Тогда гомоморфизм f является биекцией $\boldsymbol{R}^+$ на G. Легко видеть, что обратная биекция f^{-1} является гомоморфизмом группы G на группу $\boldsymbol{R}^+$. Поэтому для доказательства изоморфности топологических групп $\boldsymbol{R}^+$ и G достаточно проверить, что f^{-1} непрерывно. Пусть U — открытое в $\boldsymbol{R}^+$ множество. Так как f является биекцией, то прообраз U при отображении f^{-1} совпадает с образом U при отображении f. Но $f(U)$ открыто в G по условию этой леммы. Осталось рассмотреть третий случай.

Пусть при некотором $a > 0$ выполняется $\operatorname{Ker} f = \{n\cdot a \mid n \in \boldsymbol{Z}\}$. Рассмотрим диаграмму из непрерывных гомоморфизмов, показанную на рисунке 95, на ней $l_a(x) \rightleftarrows a\cdot x$ (т. е. l_a — непрерывный гомоморфизм $\boldsymbol{R}^+$ на $\boldsymbol{R}^+$) и $\varphi: \boldsymbol{R}^+ \to \boldsymbol{R}^+/\boldsymbol{Z}^+ \rightleftarrows T$ — канонический непрерывный гомоморфизм $\boldsymbol{R}^+$ на T (см. гл. I, § 2). Построим изоморфизм ψ топологических групп T и G.

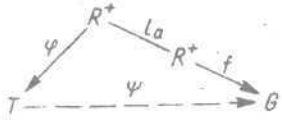


Рис. 95

Для любого $[t] \in T$ положим $\psi([t]) \rightleftarrows f\circ l_a(t)$. Проверим, что элемент $\psi([t])$ не зависит от выбора представителя в классе эквивалентности $[t]$. Если $t_1 \in [t]$, то $t_1 - t \in \boldsymbol{Z}$ и $l_a(t_1) = l_a(t) + n\cdot a$, где $n \in \boldsymbol{Z}$, и $(f\circ l_a)(t_1) = (f\circ l_a)(t) + f(n\cdot a) = (f\circ l_a)(t)$, так как f гомоморфизм и его ядро содержит число $n\cdot a$.

Докажем, что ψ непрерывно. Пусть $\mathcal{O}$ открыто в G. Тогда $\psi^{-1}(\mathcal{O}) = \varphi((f\circ l_a)^{-1}(\mathcal{O}))$, но $f\circ l_a$ непрерывно как композиция непрерывных функций, а φ открытое отображение; ψ — гомоморфизм:

$$\psi([t_1]\cdot[t_2]) = \psi([t_1\cdot t_2]) = (f\circ l_a)(t_1\cdot t_2) = (f\circ l_a)(t_1)\cdot(f\circ l_a)(t_2) = \psi([t_1])\cdot\psi([t_2]),$$

так как $f\circ l_a$, как композиция гомоморфизмов, является гомоморфизмом.

Отображение ψ инъективно: если $[t_1] \neq [t_2]$, то $t_1 - t_2 \notin \boldsymbol{Z}$ и $(l_a(t_1) - l_a(t_2)) \notin \operatorname{Ker} f$, т. е.

$$f(l_a(t_1) - l_a(t_2)) \neq 0,\ f(l_a(t_1)) - f(l_a(t_2)) \neq 0,\ (f\circ l_a)(t_1) \neq (f\circ l_a)(t_2).$$

Оно сюръективно, так как $f\circ l_a$ сюръективно и достаточно использовать определение ψ. Итак, ψ биекция и изоморфизм.

Обратное отображение ψ^{-1} непрерывно, так как ψ открыто. Проверим последнее: если $\mathcal{O}$ открыто в T, то $\psi(\mathcal{O}) = (f\circ l_a)(\varphi^{-1}(\mathcal{O}))$ и $\varphi^{-1}(\mathcal{O})$ открыто в $\boldsymbol{R}^+$, а $f\circ l_a$ открытое отображение (почему?). ■

Определение 2. **Две топологические группы G_1 и G_2 называются локально изоморфными, если их единичные элементы имеют такие окрестности $\mathcal{O}_1$ и $\mathcal{O}_2$ и найдется такая биекция f из $\mathcal{O}_1$ на $\mathcal{O}_2$, что:**

а) функция f гомеоморфизм,

б) для всех $x, y \in \mathcal{O}$, если $x\cdot y \in \mathcal{O}$, то произведение $f(x)f(y)$ принадлежит $\mathcal{O}_2$ и $f(x)\cdot f(y) = f(x\cdot y)$. Сама биекция $f: \mathcal{O}_1 \leftrightarrow \mathcal{O}_2$ называется локальным изоморфизмом топологических групп G_1 и G_2.

Замечание. Если в определении 2 опустить условие биективности f и соответственно условие непрерывности f^{-1}, то получится определение локального гомоморфизма группы G_1 в группу G_2.

Лемма 2. ***Связная топологическая группа G, локально изоморфная топологической группе $\boldsymbol{R}^+$, изоморфна группам $\boldsymbol{R}^+$ или T.***

▷ Пусть f — локальный изоморфизм групп $\boldsymbol{R}^+$ и G. Покажем, что можно построить открытый, непрерывный гомоморфизм $\boldsymbol{R}^+$ на G. Тогда по лемме I получим, что G изоморфна $\boldsymbol{R}^+$ или T.

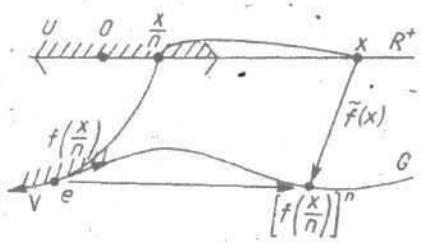


Рис. 96

Пусть U — окрестность нуля в $\boldsymbol{R}^+$ и V — окрестность e в G, между которыми функция f устанавливает биективное соответствие. Перейдем, может быть, к меньшей, чем U, но связной окрестности нуля в $\boldsymbol{R}^+$ (т. е. к интервалу) и образу этой окрестности относительно отображения f. Эти окрестности также обозначим U и V. Сужение f на U также обозначим f. Тогда биекция $f\colon U \leftrightarrow V$ является локальным изоморфизмом. Для каждого вещественного числа x по архимедовости группы $\boldsymbol{R}^+$ найдется такое натуральное число n, что $\frac{x}{n} \in U$. Положим $\tilde{f}(x) \rightleftarrows \left(f\left(\frac{x}{n}\right)\right)^n$ (рис. 96).

Докажем, что элемент $\tilde{f}(x)$ не зависит от выбора натурального числа n. Пусть $\frac{x}{m}$ также принадлежит U при некотором $m \in \boldsymbol{N}$. Тогда $\frac{x}{n \cdot m} \in U$ и в силу локальной гомоморфности f получаем:

$$\left(f\left(\frac{x}{n}\right)\right)^n = \left(f\left(m\left(\frac{x}{mn}\right)\right)\right)^n = \left(f\left(\frac{x}{mn}\right)\right)^{mn} = \left(f\left(n\frac{x}{mn}\right)\right)^m = \left(f\left(\frac{x}{m}\right)\right)^m.$$

Докажем, что $\tilde{f}$ — гомоморфизм $\boldsymbol{R}^+$ в G. Для этого при любых $x, y \in \boldsymbol{R}^+$ выберем такое $n \in \boldsymbol{N}$, что $\frac{x}{n}, \frac{y}{n}, \frac{x+y}{n} \in U$. Тогда $f\left(\frac{x+y}{n}\right) = f\left(\frac{x}{n}\right) \cdot f\left(\frac{y}{n}\right)$, но $f\left(\frac{y}{n}\right) \cdot f\left(\frac{x}{n}\right) = f\left(\frac{y+x}{n}\right)$, а $\frac{x+y}{n} = \frac{y+x}{n}$, т. е. в группе G элементы $f\left(\frac{x}{n}\right)$ и $f\left(\frac{y}{n}\right)$ коммутируют между собой. Отметим, что группа G заранее не предполагается коммутативной. Поэтому

$$\tilde{f}(x+y) = \left(f\left(\frac{x+y}{n}\right)\right)^n = \left(f\left(\frac{x}{n}\right) \cdot f\left(\frac{y}{n}\right)\right)^n =$$
$$= \underbrace{\left(f\left(\frac{x}{n}\right) \cdot f\left(\frac{y}{n}\right)\right) \ldots \left(f\left(\frac{x}{n}\right) \cdot f\left(\frac{y}{n}\right)\right)}_{n \text{ раз}} =$$
$$= \left(f\left(\frac{x}{n}\right)\right)^n \cdot \left(f\left(\frac{y}{n}\right)\right)^n = \tilde{f}(x) \cdot \tilde{f}(y).$$

Ясно, что гомоморфизм $\tilde{f}$ в окрестности U совпадает с локальным изоморфизмом f и, в частности, $\tilde{f}$ непрерывно в нуле топологической группы $\boldsymbol{R}^+$. Следовательно, функция $\tilde{f}$ непрерывна в каждой точке топологической группы $\boldsymbol{R}^+$. Аналогично из открытости локального

изоморфизма f и совпадения f и $\tilde{f}$ в окрестности U следует открытость отображения $\tilde{f}$ (проверьте). Итак, $\tilde{f}$ — непрерывный и открытый гомоморфизм из $\boldsymbol{R}^+$ в G. Остается доказать сюръективность отображения $\tilde{f}$.

Рассмотрим открытие в G множества $V^1 \rightleftarrows \{g_1 g_2 \mid g_1,\ g_2 \in V\}$, $V^2 \rightleftarrows V^1 \cdot V$, $V^2 \rightleftarrows V^2 \cdot V$, ... и открытое множество V^∞, по определению равное объединению $\bigcup_{n=1}^{\infty} V^n$. Покажем, что V^∞ не только открыто в G. Если $G \setminus V^\infty$ не пусто, то мы получаем разбиение G множества V^∞. Рассмотрим окрестность $g_0 \cdot V^{-1} \rightleftarrows \{g_0 g^{-1} \mid g \in V\}$ точки g_0. Так как g_0 — предельная точка V^∞, то в этой окрестности найдется хотя бы один элемент из V^∞, отличный от g_0. Тогда при некотором n этот элемент принадлежит V^n. Поэтому найдутся такие $g_1, g_2, \ldots, g_n$ из V, что $g_1 g_2 \ldots g_n = g_0 g^{-1}$, где $g \in V$.

Поэтому $g_0 = g_1 g_2 \ldots g_n g \in V^{n+1} \subseteq V^\infty$, т. е. V^∞ замкнуто в топологии группы G. Поскольку V^∞ замкнуто в G, то дополнение $G \setminus V^\infty$ открыто в G. Если $G \setminus V^\infty$ не пусто, то мы получаем разбиение G на два непустых непересекающихся и открытых множества V^∞ и $G \setminus V^\infty$. Это противоречит связности G. Следовательно, $V^\infty = G$. Пусть теперь $g \in G$ и $g_1, \ldots, g_n$ — такие элементы из V, что $g = g_1 g_2, \ldots, g_n$. Так как $g_1, \ldots, g_n$ из V, то найдутся числа $x_1, \ldots, x_n$ из U, для которых $g_1 = f(x_1), \ldots, g_n = f(x_n)$. Поэтому $\tilde{f}(x_1 + \ldots + x_n) = \tilde{f}(x_1) \cdot \ldots \cdot \tilde{f}(x_n) = f(x_1) \cdot \ldots \cdot f(x_n) = g_1 \ldots g_n = g$. Следовательно, отображение $\tilde{f}$ сюръективно. ■

Лемма 2 устанавливает связь между локальными свойствами связной топологической группы и ее глобальными свойствами. Неожиданным следствием леммы 2 является, например, коммутативность группы G. Другое следствие — наличие у каждого элемента $g \in G$ окрестности, замыкание которой компактно в G. Оба эти следствия трудно получить другим способом.

Следующая лемма 3 удивительна в том отношении, что, основываясь на чисто топологическом свойстве группы G (локальная гомеоморфность с R), она позволяет получить сведения об ее алгебраической (групповой) структуре. Интересно также, что в ее доказательстве приходится существенно использовать один из основных результатов главы III, а именно, теорему 5 о моделях системы положительных скалярных величин. Поскольку здесь мы доказываем теорему 3, являющуюся основой многих построений в главе I, то оказывается, что тема элементарных функций и тема измерений величин имеют общие основания.

Лемма 3. ***Пусть G — топологическая группа, нейтральный элемент e которой имеет окрестность, гомеоморфную некоторому интервалу в $\boldsymbol{R}$. Тогда топологическая группа G локально изоморфна $\boldsymbol{R}^+$.***

Краткая формулировка: если топологическая группа G локально гомеоморфна $\boldsymbol{R}$, то она локально изоморфна $\boldsymbol{R}^+$.

▷ Пусть U окрестность нейтрального элемента e и $]a, b[$ — интервал в $\boldsymbol{R}$, о котором говорится в лемме 3, а $h\colon U \to]a, b[$ —

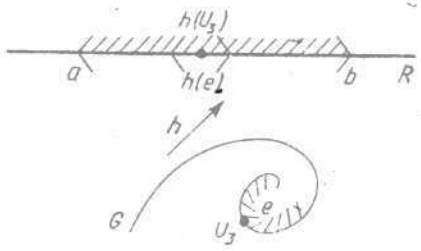


Рис. 97

соответствующий гомеоморфизм окрестности U и интервала $]a, b[$. Для начала «улучшим» окрестность U. Так как произведение и переход к обратному элементу являются непрерывными операциями в топологической группе G, то найдется такая окрестность $\mathcal{O}$ единичного элемента e, что $\mathcal{O}\cdot\mathcal{O}\subseteq U\cap U^{-1}$, $\mathcal{O}^{-1}\cdot\mathcal{O}\subseteq U$, $\mathcal{O}\cdot\mathcal{O}^{-1}\subseteq U$. Положим $U\rightleftarrows\mathcal{O}\cup\mathcal{O}^{-1}$ и получим $U_1\subseteq U$. Так как $h: U\to]a, b[$ — гомеоморфизм, то $h(U_1)$ — открытое подмножество $]a, b[$. Поэтому $h(U_1)$ состоит из не более чем счетного числа попарно непересекающихся интервалов. Пусть $]a^1, b^1[$ — тот из этих интервалов, который содержит число $h(e)$. Тогда $U_2\rightleftarrows h^{-1}(]a^1, b^1[)$ — окрестность e. Докажем, что $h(U_2^{-1})$, как и $h(U_2)=]a^1, b^1[$, является открытым интервалом, содержащим число $h(e)$. Каждому числу $t\in]a^1, b^1[$ сопоставим число $h(h^{-1}(t))^{-1}$, которое принадлежит $]a, b[$, так как $h^{-1}(t)\in U_2$. Такое соответствие задает непрерывную функцию, являющуюся композицией биекций и потому отображающую биективно $]a^1, b^1[$ на $h(U_2^{-1})$. Образ непрерывной биекции на интервале — интервал (докажите). Поэтому $h(U_2^{-1})$ — интервал, который, очевидно, содержит $h(e)$. Наконец, положим $U_3\rightleftarrows U_2\cup U_2^{-1}$ и получим окрестность e, для которой $U_3=U_3^{-1}$, $U_3\cdot U_3\subseteq U_1\cdot U_1\subseteq U$ и $h(U_3)$ — интервал (рис. 97).

Пусть $g_1, g_2\in U$. По определению положим $g_1\leqslant g_2$, если $h(g_1)\leqslant h(g_2)$, т. е. перенесем порядок с интервала $]a, b[$ на окрестность U. Определим $V\rightleftarrows\{g\in U\mid g\geqslant e\}$ и $I\rightleftarrows\{g\in U_3\mid g\geqslant e\}$.

Докажем, что пятерка объектов $\langle V, I, \cdot, e, \leqslant\rangle$, где $\cdot$ есть умножение в группе G, является системой положительных скалярных величин (по поводу последнего понятия см. главу III, определение 1). Тогда по теореме 1 из главы III получим, что существует строго возрастающая функция $f: I\to\boldsymbol{R}_{\geqslant 0}$, для которой $f(e)=0$, $f(x\cdot y)=f(x)+f(y)$ при всех $x, y\in I$ и f биективно отображает I на $[0, \lambda]$ или $[0, \lambda[$, где $0\leqslant\lambda\leqslant\infty$. Эта функция f и будет искомым локальным изоморфизмом.

Итак, проверим условия из определения 1 в главе III.

1) Операция $\cdot$ ассоциативна и e — ее нейтральный элемент по определению групповой операции в G.

2) Отношение $\leqslant$ является линейным порядком на V, так как это отношение является переносом обычного линейного порядка в интервале $]a, b[$. По определению V элемент e — наименьший относительно этого порядка элемент в V. Пусть $x\in I$, $y\in V$ и $y\leqslant x$, т. е. $y\in U$ и $h(e)\leqslant h(y)\leqslant h(x)$. Так как $x\in U_3$ и $h(U_3)$ — интервал, то $h(y)\in h(U_3)$, т. е. $y\in U_3$.

Поэтому $y\in I$.

3) Докажем, что для любых $x, y, z\in U_3$ из условия $x<y$ следуют условия $x\cdot z<y\cdot z$ и $z\cdot x<z\cdot y$. Рассмотрим функцию $\varphi: U_3\to\boldsymbol{R}$, определяемую равенством $\varphi(z)\rightleftarrows h(y\cdot z)-h(x\cdot z)$.

Так как групповые операции в G и $\boldsymbol{R}$ непрерывны и функция h непрерывна, то и функция φ непрерывна. Поэтому непрерывной функцией на интервале $h(U_3)$ является композиция $\varphi\circ h^{-1}\mid h(U_3)$. Эта функция нигде не обращается в нуль, так как в противном случае получим: $(h(yz)=h(xz))$, что эквивалентно $yz=xz$ и $y=x$.

Так как $(\varphi\circ h^{-1})(h(e))=\varphi(e)=h(y)-h(x)>0$, то получим, что функция $\varphi\circ h^{-1}$ положительна на всем интервале $h(U_3)$. Поэтому для любого z из U_3 выполняется $(h(yz)-h(xz)>0)$, т. е. $(xz<yz)$.

Аналогично доказывается второе неравенство $zx<zy$.

4) Пусть $x, y\in I$ и $x<y$. Так как окрестность U_3 симметрична, то $x^{-1}\in U_3$ и по свойству 3 $e=x^{-1}\cdot x\leqslant x^{-1}\cdot y$. Так как $e\leqslant x$, то аналогично получаем: $x^{-1}=x^{-1}\cdot e\leqslant x^{-1}\cdot x=e$. Поэтому $e\leqslant x^{-1}\cdot y\leqslant y$. Обозначим $z\rightleftarrows x^{-1}\cdot y\in V$. По последнему свойству 2) получим $z\in I$. Поэтому $y=x\cdot z$ для z из I.

5) Рассмотрим любой элемент x из U_3, отличный от e. Если $x>e$, то $I_{\geqslant e}$ не пусто. Если $x<e$, то по свойству 3) получим $e\leqslant x^{-1}$ и $x^{-1}\in U_3$, т. е. $I_{>e}$ не пусто.

Допустим, что m — наименьший элемент в $I_{>e}$. Пусть t любое число из интервала $]h(e), h(m)[$. Тогда $t\in h(U_3)$, так как $h(U_3)$ — интервал, содержащий $h(e)$ и $h(m)$, т. е. $h^{-1}(t)\in U_3$ и $e<h^{-1}(t)$ по определению отношения $<$. Поэтому $h^{-1}(t)$ — элемент в $I_{>e}$, который меньше m. Получили противоречие.

6) Пусть $x_1\leqslant x_2\leqslant\ldots$ возрастающая последовательность элементов из I. Тогда $h(x_1)\leqslant h(x_2)\leqslant\ldots$ — возрастающая последовательность чисел из интервала $h(U_3)$. Если последовательность $\{x_n\}$ ограничена сверху элементом $x\in I$, то последовательность $\{h(x_n)\}$ ограничена сверху числом $h(x)$ из $h(U_3)$. Так как $h(U_3)$ — интервал, то точная верхняя грань последовательности $\{h(x_n)\}$ принадлежит $h(U_3)$. Легко видеть, что значение функции h^{-1} на этой точной верхней грани совпадает с $\sup\{x_n\}$.

Итак, доказано, что пятерка $\langle V, I, \cdot, e, \leqslant\rangle$ является системой положительных скалярных величин и, следовательно, к ней действительно применима теорема 1 из главы III. Пусть $f: I\to\boldsymbol{R}_{\geqslant 0}$ — строго возрастающая функция из этой теоремы.

Для элементов x из $U_3\setminus I$ определим $f_I(x)$ равенством $f(x)\rightleftarrows -f(x^{-1})$, где x^{-1} уже обязано принадлежать I (почему?). Получилась функция $f: U_3\to\boldsymbol{R}$. Она строго возрастает на множестве U_3. Действительно, если $e\leqslant x<y$, то по упомянутой теореме 1 $f(x)<f(y)$; если $x<e\leqslant y$, то $f(x)<0\leqslant f(y)$; если $x<y<e$, то $e<x^{-1}<y^{-1}$ и $0<f(x^{-1})<f(y^{-1})$, т. е. $f(x)<f(y)$. Так как окрестность U_3 симметрична, то и множество $f(U_3)$ симметрично относительно нуля. По теореме 1 из главы III множество $f(I)$ является

числовым промежутком в $\boldsymbol{R}$, содержащим нуль. Поэтому $f(U_3)$ является числовым промежутком, содержащим нуль строго внутри себя.

Докажем, что $f(x\cdot y)=f(x)+f(y)$ для всех таких точек x, $y\in U_3$, что $x\cdot y\in U_3$. Если x, $y\in I$, то $e\leqslant x\cdot y$ и, следовательно, $x\cdot y\in I$. Для таких x и y равенство $f(x\cdot y)=f(x)+f(y)$ обеспечивается теоремой 1. Если x, $y\in U_3\backslash I$, то $x<e$, $y<e$ и, следовательно, $x\cdot y<e$. Тогда

$$f(x\cdot y)=-f((x\cdot y)^{-1})=-f(y^{-1}\cdot x^{-1})=-(f(y^{-1})+f(x^{-1}))=f(x)+f(y).$$

Если $x<e\leqslant y$ и $x\cdot y<e$, то $f(x)=f(xy\cdot y^{-1})=f(xy)+f(y^{-1})$ и $f(x\cdot y)=f(x)-f(y^{-1})=f(x)+f(y)$.

Аналогично разбирается случай, когда $x<e\leqslant y$ и $xy\geqslant e$.

Итак, доказано, что f — локальный гомоморфизм.

Так как f — строго возрастающая функция на U_3, то $f\circ h^{-1}\restriction h(U_3)$ — строго возрастающая функция на интервале $h(U_3)$.

При этом образ $(f\circ h^{-1})(h(U_3))$ совпадает с образом $f(U_3)$, который, как отмечалось выше, является промежутком, симметричным относительно нуля. Следовательно, функция $f\circ h^{-1}$ непрерывна на интервале $h(U_3)$, а функция $f=(f\circ h^{-1})\circ h$ непрерывна на U_3.

Аналогично функция $h\circ f^{-1}$ на промежутке $f(U_3)$ непрерывна как обратная к непрерывной, строго возрастающей функции $f\circ h^{-1}$. Следовательно, и функция $f^{-1}=h^{-1}\circ(h\circ f^{-1})$ непрерывна. ■

Следствие. Связная топологическая группа G, локально гомеоморфная $\boldsymbol{R}$, коммутативна.

Замечание. Без условия связности следствие неверно. Например, пусть $G=O_2$. Тогда G не коммутативна, но у ее единичного элемента $E\rightleftarrows\begin{pmatrix}1&0\\0&1\end{pmatrix}$ имеется окрестность, гомеоморфная интервалу в $\boldsymbol{R}$, так как $O_2=SO_2\cup(O_2\backslash SO_2)$, а множество SO_2 связно и гомеоморфно окружности.

▷ Доказательство теоремы 3. Пусть G — связная топологическая группа, нейтральный элемент которой имеет окрестность, гомеоморфную какому-нибудь интервалу на числовой прямой. Тогда по лемме 3 группа G локально изоморфна $\boldsymbol{R}^+$ и по лемме 2 она изоморфна $\boldsymbol{R}^+$ или T.

В частности, если G компактна, то G изоморфна T. В противном случае G изоморфна $\boldsymbol{R}^+$. ■

2. Длина дуги. Определение функций косинус и синус числового аргумента на основе понятия длины дуги.

1. Обозначим $|a, b|$ множество вещественных чисел, которое является отрезком, или (конечным или бесконечным) интервалом, или (конечным или бесконечным) полуинтервалом. Иными словами, по определению $|a, b|$ — любое связное подмножество множества $\boldsymbol{R}$.

Напомним, что кривой (точнее, гладкой кривой) называется любое отображение α вида $\alpha: |a, b|\to\boldsymbol{R}^2$, которое обладает свойством: функции $x(t)=P_1(\alpha(t))$ и $y(t)=P_2(\alpha(t))$ имеют непрерывные производные в каждой точке их области определения. Итак,

$\alpha(t)=\begin{pmatrix}x(t)\\y(t)\end{pmatrix}\in\boldsymbol{R}^2$. Напомним, что $\alpha'(t)\rightleftarrows\begin{pmatrix}x'(t)\\y'(t)\end{pmatrix}\in\boldsymbol{R}^2$ и $|\xi|^2=(\xi,\xi)=x^2+y^2$, где $\xi=\begin{pmatrix}x\\y\end{pmatrix}$. Длиной кривой $\alpha: |a, b|\to\boldsymbol{R}^2$ называется число или символ $+\infty$, которые определяются как

$$l_\alpha\rightleftarrows\int_a^b\sqrt{(x'(t))^2+(y'(t))^2}\,dt=\int_a^b|\alpha'(t)|dt. \qquad (*)$$

Если область определения кривой α совпадает с интервалом или полуинтервалом с концами a и b, где $-\infty\leqslant a$, $b\leqslant+\infty$, то интеграл в $(*)$ может оказаться несобственным; в этом случае формулу $(*)$ нужно понимать так:

$$l_\alpha\rightleftarrows\lim_{\substack{\varepsilon\to0\\\varepsilon>0}}\int_{a+\varepsilon}^{b-\varepsilon}|\alpha'(t)|dt,$$

если a и b конечные величины, и

$$l_\alpha\rightleftarrows\int_{-\infty}^{+\infty}|\alpha'(t)|dt,$$

если $a=-\infty$ и $b=+\infty$ (аналогично интерпретируются случаи $a=-\infty$, $b<+\infty$ и $-\infty<a$, $b=+\infty$).

Дугой в плоскости $\boldsymbol{R}^2$ называется любое подмножество множества $\boldsymbol{R}^2$, являющееся образом некоторой кривой. Любую кривую α, образ которой совпадает с дугой M, назовем заданием дуги M. Ясно, что для одной той же дуги можно найти много различных заданий. Например, отрезок в плоскости $\boldsymbol{R}^2$ с концами в точках $\langle 0, 0\rangle$ и $\langle 1, 1\rangle$ является дугой, и любая из следующих кривых служит ее заданием:

$$\begin{aligned}
&\alpha_1(t)\rightleftarrows\langle t, t\rangle, && 0\leqslant t\leqslant1;\\
&\alpha_2(t)\rightleftarrows\langle 3t+1, 3t+1\rangle, && -\tfrac{1}{3}\leqslant t\leqslant0;\\
&\alpha_3(t)\rightleftarrows\langle t^3, t^3\rangle, && 0\leqslant t\leqslant1;\\
&\alpha_4(t)\rightleftarrows\langle t^2, t^2\rangle, && -1\leqslant t\leqslant1;\\
&\alpha_5(t)\rightleftarrows\langle t^2, t^2\rangle, && -\tfrac{1}{2}\leqslant t\leqslant1.
\end{aligned}$$

Легко подсчитать, что $l_{\alpha_1}=l_{\alpha_2}=l_{\alpha_3}=\sqrt{2}$, а $l_{\alpha_4}=2\sqrt{2}$, $l_{\alpha_5}=\frac{5\sqrt{2}}{4}$.

В качестве еще одного примера рассмотрим различные задания единичной окружности $S=\{\langle x, y\rangle\in\boldsymbol{R}^2\,|\,x^2+y^2=1\}$:

$$\begin{aligned}
&\alpha_1(t)\rightleftarrows\langle\cos t, \sin t\rangle, && 0\leqslant t\leqslant2\pi;\\
&\alpha_2(t)\rightleftarrows\langle\cos(t^2), \sin(t^2)\rangle, && \sqrt{\pi}\leqslant t\leqslant\sqrt{3\pi};\\
&\alpha_3(t)\rightleftarrows\langle\cos t, \sin t\rangle, && 0\leqslant t\leqslant4\pi.
\end{aligned}$$

Легко подсчитать, что $l_{\alpha_1}=l_{\alpha_2}=2\pi$, а $l_{\alpha_3}=4\pi$.

Итак, различные кривые, задающие одну и ту же дугу, могут иметь различную длину. Выясним, при каких условиях два различных задания $\gamma(t)$, $a\leqslant t\leqslant b$ и $\alpha(\tau)$, $c\leqslant\tau\leqslant d$ одной и той же дуги M имеют одинаковую длину.

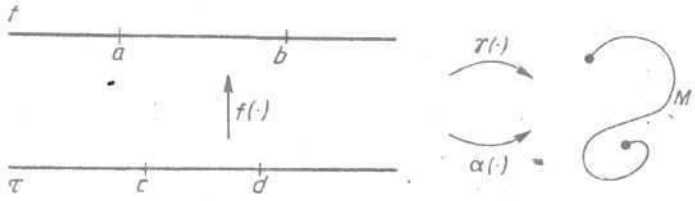


Рис. 98

Предложение 1. ***Пусть*** $\gamma : |a, b| \to \boldsymbol{R}^2$ ***и*** $\alpha : |c, d| \to \boldsymbol{R}^2$ ***две кривые, задающие одну и ту же дугу*** M, ***то есть*** $\gamma(|a, b|) = \alpha(|c, d|) = M$. ***Если найдется такая функция*** $f : |c, d| \leftrightarrow |a, b|$, ***что*** $\alpha(\tau) = \gamma(f(\tau))$ ***и производная*** $f'(\tau)$ ***непрерывна и отлична от нуля при всех*** $\tau \in |c, d|$, ***то*** $l_\gamma = l_\alpha$ ***(см. рис. 98).***

▷ Так как функция $f'(\tau)$ непрерывна и отлична от нуля на множестве $|c, d|$, то $f'(\tau)$ либо всюду на этом множестве положительна, либо всюду отрицательна. Рассмотрим, например, случаи, когда $f'(\tau) < 0$ при всех $\tau \in |c, d|$. Тогда функция f монотонно убывает и $f(c) = b$ и $f(d) = a$. Рассмотрим замену переменной $t = f(\tau)$. По теореме о замене переменной в определенном интервале получаем для кривой $\gamma(t) \rightleftarrows \langle x(t), y(t) \rangle$:

$$l_\gamma = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}\, dt = \int_d^c \sqrt{[x'(f(\tau))]^2 + [y'(f(\tau))]^2}\, f'(\tau) d\tau =$$

$$= \int_c^d \sqrt{[(x \circ f)'(\tau)]^2 + [(y \circ f)'(\tau)]^2}\, d\tau = l_\alpha,$$

где последнее равенство справедливо по определению l_α, так как по условию

$$\alpha(\tau) = \gamma(f(\tau)) = \langle x(f(\tau)),\ y(f(\tau)) \rangle = \langle (x \circ f)(\tau),\ (y \circ f)(\tau) \rangle. \quad \blacksquare$$

Определение 1.

а) Дуга M называется простой, если найдется такое ее задание $\alpha : |a, b| \to \boldsymbol{R}^2$, что функция α инъективна и $\alpha'(t) \neq \langle 0, 0 \rangle$ для всех t из $|a, b|$.

б) Параметризацией простой дуги M называется любое ее задание $\alpha : |a, b| \to \boldsymbol{R}^2$, удовлетворяющее условиям п. а), т. е. функция α инъективна и $\alpha'(t) \neq \langle 0, 0 \rangle$ для всех t из $|a, b|$.

в) Длиной простой дуги M называется длина любой ее параметризации.

Это определение простой дуги существенно шире обычного; мы принимаем его ради того, чтобы окружность S оказалась простой дугой.

Предложение 2. ***Пусть*** $\alpha : [a, b] \to M$ ***произвольная параметризация простой дуги*** M. ***Функция*** α ***является гомеоморфизмом отрезка*** $[a, b]$ ***и множества*** M.

▷ По условию функция α инъективна. Она сюръективна, так как

по определению дуги выполняется $M = \alpha([a, t])$. Поэтому функция $\alpha : [a, b] \to M$ — биекция. В множествах $[a, t]$ и M, конечно, рассматриваются топологии, индуцированные соответственно из $\boldsymbol{R}$ и $\boldsymbol{R}^2$. По условию α непрерывна и, следовательно, непрерывна в топологиях множеств $[a, b]$ и M (почему?). Осталось проверить, что функция $\alpha^{-1} : M \to [a, b]$ непрерывна, или, что то же самое, функция $\alpha : |a, b] \to M$ открыта, или, что то же самое, функция α замкнута. Последнее по определению означает: для любого замкнутого в $[a, b]$ множества F его образ $\alpha(F)$ замкнут в M. Но последнее утверждение является общим фактом: если $f : X \to Y$ — непрерывная биекция компактного топологического пространства X в отделимое топологическое пространство Y, то функция f замкнута. Действительно, если F — замкнутое подмножество X, то F компактно (почему?); образ компактного множества F относительно непрерывной функции f компактен (почему?). Поэтому множество $f(F)$ компактно. Любое компактное подмножество множества Y замкнуто (почему?). (Где в этом рассуждении используются компактность X, отделимость Y и непрерывность f?) ■

Проверим, что длина простой дуги M не зависит от выбора параметризации этой дуги.

Предложение 3. ***Пусть*** $\gamma : |a, b| \to \boldsymbol{R}^2$ ***и*** $\alpha : |c, d| \to \boldsymbol{R}^2$ ***две параметризации одной и той же простой дуги*** M. ***Тогда*** $l_\gamma = l_\alpha$.

▷ По предложению 1 для доказательства достаточно найти такую функцию $f : |a, b| \leftrightarrow |c, d|$, что функция $f'(\tau)$ непрерывна и отлична от нуля при всех τ и $\alpha(\tau) = \gamma(f(\tau))$ при всех $\tau \in |c, d|$. По предложению 2 функция α (и аналогично функция γ) является гомеоморфизмом любого отрезка $[\tau_0, \tau_1]$, содержащегося в ее области определения, и подмножества $\alpha([\tau_0, \tau_1])$ множества M. Рассмотрим функцию f, определенную равенством $f(\tau) \rightleftarrows \gamma^{-1}(\alpha(\tau))$, где $\tau \in |c, d|$. Очевидно, f — биекция. Проверим, что функция f непрерывна. Для этого достаточно проверить непрерывность функции f в каждой точке ее области определения — промежутка $|c, d|$. Пусть точка τ_0 из $|c, d|$; положим $t_0 \rightleftarrows f(\tau_0) = \gamma^{-1}(\alpha(\tau_0))$ и выберем отрезок $[t_1, t_2]$, содержащий точку t_0 и содержащийся в области определения функции γ — промежутке $|a, b|$. При этом отрезок $[t_1, t_2]$ должен содержать точку t_0 внутри, т. е. $t_0 \in]t_1, t_2[$. Последнее невозможно только в том случае, если промежуток $|a, b|$ содержит один из концов и этот конец совпадает с t_0 (пусть, например, $a = t_0$), тогда в качестве отрезка $[t_1, t_2]$ выберем $[a, t_2]$, и точка t_0 опять-таки принадлежит открытому в $|a, b|$ множеству $[a_1, t_2[$. По предложению 2 непрерывная биекция γ, суженная на отрезок $[t_1, t_2]$, является гомеоморфизмом. Поэтому $M_1 \rightleftarrows \gamma(]t_1, t_2[)$ — открытое в M множество, которое содержит точку $\alpha(\tau_0)$. Тогда $\mathcal{O} \rightleftarrows \alpha^{-1}(M_1) = f^{-1}(]t_1, t_2[)$ открытое в $|c, d|$ множество, содержащее точку τ_0. Функция $\gamma^{-1} : M_1 \to]t_1, t_2[$ непрерывна, и поэтому функция $f : \mathcal{O} \to]t_1, t_2[$ непрерывна как композиция непрерывных функций $f = \gamma^{-1} \circ \alpha$. В частности, она непрерывна в точке τ_0.

Итак, функция f является непрерывной биекцией; любая непре-

рывная биекция вида $|c, d| \to |a, b|$ является строго монотонной функцией (почему?). Поэтому дробь $\frac{f(\tau_1)-f(\tau)}{\tau_1-\tau}$ имеет постоянный знак при всех значениях τ, τ_1 из $|c, d|$ и $\tau \neq \tau_1$.

Остается проверить, что функция f непрерывно дифференцируема и функция $f'(\tau) \neq 0$ при всех $\tau \in |c, d|$. Пусть $\tau, \tau_1 \in |c, d|$ и $\tau_1 \neq \tau$ обозначим $t_1 = f(\tau_1)$, $t = f(\tau)$; $t_1 \neq t$. По теореме Лагранжа $\alpha(\tau_1) - \alpha(\tau) = \langle x(\tau_1) - x(\tau), \ y(\tau_1) - y(\tau) \rangle = \langle (\tau_1 - \tau) \cdot x'(\tau_2), \ (\tau_1 - \tau) \times y'(\tau_3) \rangle$, где τ_2 и τ_3 какие-то числа, лежащие между τ_1 и τ. Аналогично, применяя теорему Лагранжа к функции $\gamma(t) = \langle x_1(t), y_1(t) \rangle$, получим:

$$\begin{gathered}\alpha(\tau_1) - \alpha(\tau) = \gamma(f(\tau_1)) - \gamma(f(\tau)) = \\ = \gamma(t_1) - \gamma(t) = \langle x_1(t_1) - x_1(t), \ y_1(t_1) - y_1(t) \rangle = \\ = \langle (t_1 - t) \cdot x_1'(t_2), \ (t_1 - t) \cdot y_1'(t_3) \rangle,\end{gathered}$$

где t_2 и t_3 какие-то числа, лежащие между числами t и t_1. Найдем длину вектора $\alpha(\tau_1) - \alpha(\tau)$ по первой и второй из этих формул и приравняем полученные результаты:

$$(\tau_1 - \tau)^2 ((x'(\tau_2))^2 + (y'(\tau_3))^2) = (t_1 - t)^2 ((x_1'(t_2))^2 + (y_1'(t_3))^2).$$

Поэтому

$$\left(\frac{f(\tau_1)-f(\tau)}{\tau_1-\tau}\right)^2 = \left(\frac{t_1-t}{\tau_1-\tau}\right)^2 = \frac{(x'(\tau_2))^2 + (y'(\tau_3))^2}{(x_1'(t_2))^2 + (y_1'(t_3))^2}.$$

Перейдем к пределу при $\tau_1 \to \tau$. Тогда $\tau_2 \to \tau$, $\tau_3 \to \tau$ и в силу непрерывности функций $x'(\cdot)$ и $y'(\cdot)$ получим:

$$(x'(\tau_2))^2 + (y'(\tau_3))^2 \to |\alpha'(\tau)|^2 \neq 0.$$

В силу непрерывности f получим: $t_1 \to t$, $t_2 \to t$, $t_3 \to t$ и по непрерывности функций $x_1'(\cdot)$, $y_1'(\cdot)$

$$(x_1'(t_2))^2 + (y_1'(t_3))^2 \to |\gamma'(t)|^2 \neq 0.$$

Итак, доказано, что $\lim\limits_{\tau_1 \to \tau} \left(\frac{f(\tau_1)-f(\tau)}{\tau_1-\tau}\right)^2 = \frac{|\alpha'(\tau)|^2}{|\gamma'(f(\tau))|^2} \neq 0$. Как отмечалось выше, дробь $\frac{f(\tau_1)-f(\tau)}{\tau_1-\tau}$ имеет постоянный знак. Поэтому, извлекая корень, получим $f'(\tau) = \varepsilon \cdot \frac{|\alpha'(\tau)|}{|\gamma'(f(\tau))|} \neq 0$ при всех $\tau \in |c, d|$, где ε — постоянная, равная $+1$ или -1.

Функция $f'(\cdot)$ непрерывна, так как она является частным двух непрерывных функций и $|\gamma'(f(\tau))| = \sqrt{[x_1'(f(\tau))]^2 + [y_1'(f(\tau))]^2} \neq 0$. ■

2. Цель этого приложения — определить тригонометрические функции с помощью длины дуги окружности. В частности, определить $\cos x$, где $x \in]0, \pi[$, как первую координату такой точки z на верхней полуокружности $S_+ \rightleftarrows \{\langle x, y \rangle \in S \mid y \geqslant 0\}$ окружности S, что дуга $\overset{\smile}{1, z}$, соединяющая точки 1 и z и лежащая в S_+, имеет

длину x. Для этого нужно доказать, что множество $\overset{\smile}{1, z}$ действительно является простой дугой, т. е. по определению 1 нужно найти ее параметризацию, затем вычислить ее длину и определить число π. При этом нельзя пользоваться тригонометрическими функциями, так как мы их собираемся еще определить. Следовательно, в нашем распоряжении остаются только рациональные функции, частные от деления двух многочленов.

Оказывается, что окружность S можно параметризовать кривой, не связанной с тригонометрическими функциями. А именно, ее можно параметризовать такой кривой $\alpha(t) = \langle x(t), \ y(t) \rangle$, что функции $x(t)$ и $y(t)$ будут рациональными функциями от переменной t, т. е. они будут равны частному двух многочленов от одной переменной t. Такие параметризации называются рациональными.

Чтобы найти рациональную параметризацию окружности S, вспомним (см. § 5 этой главы) определение стереографической проекции p окружности S на прямую $\boldsymbol{R}$ из точки $\langle a, b \rangle$, лежащей на S.

Обычно в качестве точки $\langle a, b \rangle$ выбирается точка $\langle 0, 1 \rangle$, но сейчас удобнее выбрать в этом качестве точку $\langle -1, 0 \rangle$, так как в этом случае $p(\langle 0, 0 \rangle) = 0$. Итак, по определению каждой точке $\langle x, y \rangle \in S$, отличной от точки $\langle -1, 0 \rangle$, сопоставим точку $p(\langle x, y \rangle)$, равную пересечению прямой, проходящей через точки $\langle -1, 0 \rangle$ и $\langle x, y \rangle$, с осью ординат. Отображение $p : \langle x, y \rangle \mapsto p(\langle x, y \rangle)$ является гомеоморфизмом множеств $S \setminus \{\langle -1, 0 \rangle\}$ и $\boldsymbol{R}$ (см. рис. 99).

Обозначим α функцию, обратную к биекции p; она имеет вид $\alpha : \boldsymbol{R} \to S \setminus \{\langle -1, 0 \rangle\}$, оказывается, что функция α как раз определяет рациональную параметризацию окружности.

Предложение 4. ***Определенная выше функция $\alpha : \boldsymbol{R} \to S \setminus \{\langle -1, 0 \rangle\}$ задается формулой***

$$\alpha(t) = \left\langle \frac{1-t^2}{1+t^2}, \ \frac{2t}{1+t^2} \right\rangle \qquad (**)$$

и является параметризацией простой дуги ($S \setminus \{\langle -1, 0 \rangle\}$).

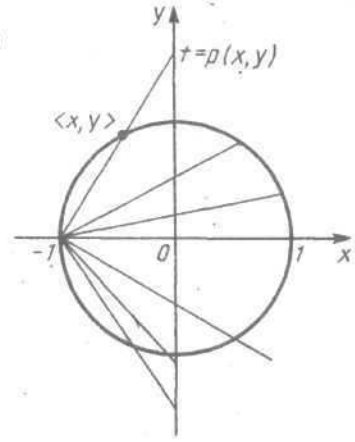


Рис. 99

Сначала проверим, что формула $(**)$ действительно определяет функцию, обратную к стереографической проекции p, т. е. проверим, что для всех $t \in \boldsymbol{R}$ точка $\alpha(t)$, вычисленная по формуле $(**)$, совпадает с точкой $\langle x, y \rangle$, являющейся пересечением окружности S и прямой, проходящей через точки $\langle -1, 0 \rangle$ и $\langle 0, t \rangle$ (рис. 99). Действительно, эта прямая задается уравнением $y = tx + t$. Поэтому для точки $\langle x, y \rangle$ получаем:

$$1=x^2+y^2=x^2+t^2(1+x)^2;\ x^2(1+t^2)+2xt^2+t^2-1=0;$$
$$x=\frac{-t^2\pm\sqrt{t^4-(t^2-1)}}{1+t^2}=\frac{-t^2\pm 1}{1+t^2}=\frac{1-t^2}{1+t^2},$$

так как по условию $x\neq -1$. Тогда $y=t(x+1)=\frac{2t}{1+t^2}$, т. е. $\alpha(t)=\langle x, y\rangle$.

Из доказанного, в частности, следует инъективность функции α, определенной на $\boldsymbol{R}$, и равенство $\alpha(\boldsymbol{R})=(S\setminus\{\langle -1, 0\rangle\})$. Проверим, что функция $\alpha(t)=\langle x(t), y(t)\rangle$ является гладкой кривой и $\forall t\in\boldsymbol{R}(\alpha'(t)\neq\langle 0, 0\rangle)$. Вычислим x' и y':

$$x'(t)=\frac{2(1+t^2)-2t2t}{(1+t^2)^2}=\frac{2(1-t^2)}{(1+t^2)^2};\ y'(t)=\frac{2t(1+t^2)-(t^2-1)2t}{(1+t^2)^2}=\frac{4t}{(1+t^2)^2}$$

и получим, что они — непрерывные функции, одновременно не обращающиеся в нуль ни в какой точке $t\in\boldsymbol{R}$, так как $|\alpha'(t)|\neq 0$. Следовательно, $(S\setminus\{\langle -1, 0\rangle\})$ простая дуга и α ее параметризация. Отметим для дальнейшего, что $|\alpha'(t)|=\frac{2}{1+t^2}$.

Упражнение 2. *Аналогично доказательству предложения 4 проверьте, что любая компонента связности в множестве $S\setminus\{z_1, z_2\}$, где z_1, z_2 точки из S (т. е. любая дуга в наивном смысле этого слова) является простой дугой в смысле определения 1; кроме того, дуга $\overset{\smile}{1, z}$ (определяемая как $\{\langle x, y\rangle\in S_+ \mid 1\geqslant x\geqslant \operatorname{Re} z\}$, если $z\in S_+$, и точно так же с заменой S_+ на $S\setminus S_+$, если $z\in S\setminus S_+$) параметризуется кривой $\alpha\restriction[0, t]$, где $\alpha(t)=z$.*

Определим число π как длину верхней полуокружности S_+.

Из упражнения 2 ясно, что эта полуокружность является простой дугой с параметризацией. Поэтому наше определение числа π означает

$$\pi=\int_{-1}^{1}|\alpha'(t)|dt=2\int_{-1}^{1}\frac{dt}{1+t^2}.$$

Предложение 5. ***а) Длина простой дуги $(S\setminus\{\langle -1, 0\rangle\})$ равна 2π.***

б) Длины всех дуг вида $(S\setminus\{\langle a, b\rangle\})$, где $\langle a, b\rangle$ — фиксированная точка из S, одинаковы и равны 2π.

▷ а) По предложению 3 длина дуги $S\setminus\{\langle -1, 0\rangle\}$ равна $\int_{-\infty}^{+\infty}\frac{2dt}{1+t^2}$.

Следовательно, требуется доказать, что

$$\int_{-\infty}^{+\infty}\frac{2dt}{1+t^2}=2\cdot\int_{-1}^{+1}\frac{2dt}{1+t^2},$$

что, в силу четности функции $\frac{1}{1+t^2}$, эквивалентно равенству

$$\int_{0}^{+\infty}\frac{2dt}{1+t^2}=2\cdot\int_{0}^{1}\frac{2dt}{1+t^2}.$$

Так как $\int_0^{\infty}f(t)dt=\int_0^1 f(t)dt+\int_1^{+\infty}f(t)dt$, то остается доказать равенство

$$\int_{0}^{1}\frac{dt}{1+t^2}=\int_{1}^{+\infty}\frac{dt}{1+t^2}.$$

Выполним замену переменной $t=\frac{1}{\tau}$ в несобственном интеграле $\int_1^{+\infty}\frac{dt}{1+t^2}$:

$$\int_{1}^{+\infty}\frac{dt}{1+t^2}=\int_{1}^{0+}\frac{-\frac{1}{\tau^2}d\tau}{1+\left(\frac{1}{\tau}\right)^2}=-\int_{1}^{0+}\frac{d\tau}{1+\tau^2}=\int_{0+}^{1}\frac{d\tau}{1+\tau^2},$$

где $\tau\in[0, 1]$.

Так как значение интеграла от непрерывной функции $\frac{1}{1+\tau^2}$ на отрезке $[0, 1]$ равно значению интеграла от этой же функции на полуинтервале $]0, 1]$, то получаем

$$\int_{1}^{+\infty}\frac{dt}{1+t^2}=\int_{0}^{1}\frac{d\tau}{1+\tau^2},$$

где $\tau\in[0, 1]$.

б) Определим следующую параметризацию γ: $\boldsymbol{R}\to\boldsymbol{R}^2$ дуги $S\setminus\{\langle a, b\rangle\}$:

$$\gamma(t)=\langle -ax+by,\ -bx-ay\rangle,\ \text{где } x=\frac{1-t^2}{1+t^2},\ y=\frac{2t}{1+t^2}.$$

Иными словами, параметризация γ является композицией $A\circ\alpha$ параметризации α дуги $S\setminus\{\langle -1, 0\rangle\}$ и поворота $A:\boldsymbol{R}^2\to\boldsymbol{R}^2$, при котором точка $\langle -1, 0\rangle$ перейдет в точку $\langle a, b\rangle$. Тогда, учитывая равенство $a^2+b^2=1$, получим

$$l_\gamma=\int_{-\infty}^{+\infty}\sqrt{((-a\cdot x(t)+b\cdot y(t))')^2+((b\cdot x(t)+a\cdot y(t))')^2}\,dt=$$
$$=\int_{-\infty}^{+\infty}\sqrt{(x'(t))^2+(y'(t))^2}\cdot dt=l_\alpha=2\pi.\ \blacksquare$$

Упражнение 3. *Докажите, что:*
а) длина дуги $S\cap\{\langle x, y\rangle \mid y\leqslant 0\}$ равна π;

б) длина любой полуокружности, т. е. любой компоненты связности множества $(S\setminus\{z,\ -z\})$, где z — фиксированная точка окружности, равна π; тот же результат получается, если к полуокружности добавить любой из ее концов или оба конца;

в) длины всех четвертей окружности одинаковы и равны $\frac{\pi}{2}$ независимо от того, содержат ли они свои концы.

Все упомянутые длины, как и численное значение постоянной π, вычисляются без использования тригонометрических функций. Действительно, во всех случаях возникают интегралы с подынтегральной функцией $\frac{1}{1+t^2}$. Эта функция является суммой членов бесконечно убывающей прогрессии со знаменателем $(-t^2)$, т. е.

$$\frac{1}{1+t^2}=\sum_{n=0}^{\infty}(-t^2)^n \text{ при } |t|<1.$$

Поэтому интегрирование сводится к вычислению значений по существу многочленов. Например: $\pi=\int_{-1}^{1}\frac{dt}{1+t^2}=\lim_{\varepsilon\to 0-}\int_{-1+\varepsilon}^{1-\varepsilon}\sum_{n=0}^{\infty}(-t^2)^n\times$

$$\times dt=\lim_{\varepsilon\to 0+}\left(\sum_{n=0}^{\infty}(-1)^n\cdot\frac{2(1-\varepsilon)^{2n+1}}{2n+1}+\sum_{n=m+1}^{\infty}(-1)^n\cdot\frac{2(1-\varepsilon)^{2n+1}}{2n+1}\right)=2\times$$

$$\times\sum_{n=0}^{\infty}(-1)^n\cdot\frac{1}{2n+1}+2\cdot\lim_{\varepsilon\to 0+}\sum_{n=m+1}^{\infty}(-1)^n\cdot\frac{(1-\varepsilon)^{2n+1}}{2n+1}.$$

По признаку Лейбница сходимости знакопеременных рядов

$$\left|\sum_{n=m+1}^{\infty}(-1)^n\cdot\frac{(1-\varepsilon)^{2n+1}}{2n+1}\right|\leqslant\frac{|1-\varepsilon|^{2m+1}}{2m+1}.$$

Поэтому при любом фиксированном m модуль разности между числом π и суммой $2\left(1-\frac{1}{3}+\frac{1}{5}-\ldots\frac{(-1)^m}{2m+1}\right)$ не превышает $\frac{2}{2m+1}$, т. е. равняется сумме числового ряда $2\cdot\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}$ и сколь угодно точно приближается начальным отрезком этого ряда.

Упражнение 4. *Докажите, что вся единичная окружность является простой дугой и ее длина равна 2π.*

Указание: соедините параметризации двух полуокружностей, например правой полуокружности, включающей точки $\binom{0}{1}$ и $\binom{0}{-1}$, и левой полуокружности, не включающей этих точек; получится параметризация, определенная на полуинтервале.

Теперь перейдем к определению функции $\cos(\cdot)$ и $\sin(\cdot)$.

Определение 3. **Пусть число $x\in]-\pi,\pi[$ и ξ — такое число, что длина простой дуги $\overset{\smile}{0,\alpha}(\xi)$ равна числу x, т. е. $\int_0^{\xi}\frac{2dt}{1+t^2}=x$.**

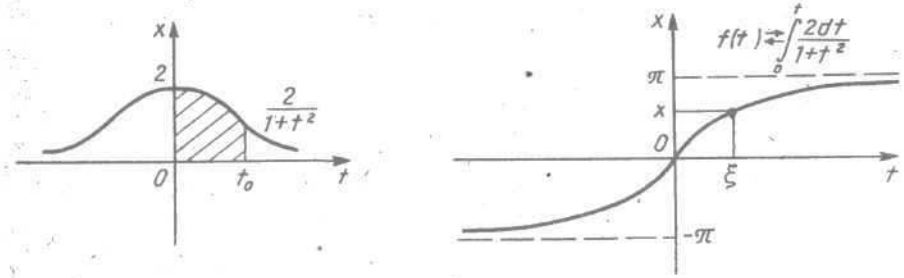


Рис. 100

Тогда положим $\cos x\rightleftarrows x(\xi)=\frac{1-\xi^2}{1+\xi^2}$ и $\sin x\rightleftarrows y(\xi)=\frac{2\xi}{1+\xi^2}$. Доопределим эти функции $\cos x$ и $\sin x$ в точках $x=\pm\pi$ соответственно значениями (-1) и 0 и затем продолжим их на множество всех вещественных чисел R периодически с периодом 2π.

Итак, получено новое (по сравнению с § 5, 6) определение тригонометрических функций.

Существование и единственность числа $\xi\in R$, для которого $\int_0^{\xi}\frac{2dt}{1+t^2}=x$ можно получить и независимо от понятия простой дуги, как следствие монотонности и непрерывности интеграла $\int_0^{\xi}\frac{2dt}{1+t^2}$ функции с переменным верхним пределом (рис. 100).

Упражнение 5. *Проверьте, что таким образом определенные функции $\cos x$ и $\sin x$ удовлетворяют аксиоматическому определению одноименных функций в § 5, 6. Отсюда по теореме 8 о единственности аксиоматически определенных тригонометрических функций получаем, что оба определения функций $\cos x$ и $\sin x$ совпадают.*

ПРИЛОЖЕНИЕ 2

(к главе III)

Доказательство теоремы о моделях системы положительных скалярных величин.

Итак, докажем теорему 1 из главы III. Пусть элементы $x,\ y\in I$ и $y\neq x_0$. Обозначим наибольшее целое число $n>0$, для которого y^n определено и не превосходит x (рис. 101):

$$(x:y)\rightleftarrows\max\{n\,|\,n\in\mathbf{Z}_{\geqslant 0}\text{ и }y^n\leqslant x\}.$$

Лемма 1. ***а) Для любых $x,\ y,\ z\in I,\ z\neq x_0$ выполняется $(x:z)+(y:z)\leqslant((x\cdot y):z)\leqslant(x:z)+(y:z)+1$.***

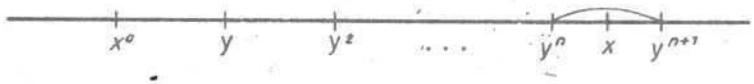


Рис. 101

б) Для любых $x, y, z \in I$, $z \neq y_0$, $y \neq x_0$ выполняется

$$(x : z)+1 \leqslant ((y : z)+1)\,((x : y)+1).$$

▷ а) Пусть $(x : z)=n$, $(y : z)=m$. Тогда $z^n \leqslant x \leqslant z^{n+1}$ и $z^m \leqslant y \leqslant z^{m+1}$. По аксиоме 2 из определения 1 системы положительных скалярных величин (из главы III) в части, касающейся I, получаем: z^n, $z^m \in I$ и $z^n \cdot z^m$ определено. Так как на I умножение $\cdot$ и порядок $\leqslant$ согласованы (аксиома 3), то $z^n \cdot z^m = z^{n+m} \leqslant x \cdot y$, т. е. $n+m \leqslant ((x \cdot y) : z)$.

Если $z^{n+1} \cdot z^{m+1}$ определено, то аналогичным образом $x \cdot y < z^{n+m+2}$, т. е. $((x \cdot y) : z) < n+m+2$.

Если же $z^{(n+m+2)}$ не определено, то по определению функции $(x : y)$ также получаем $((x \cdot y) : z) < n+m+2$.

б) Пусть $(y : z)=n$, $(x : y)=m$. Тогда $z^n \leqslant y < z^{n+1}$ и $y^m \leqslant x < y^{m+1}$.

Если $z^{(n+1)(m+1)}$ определено, то $x < y^{m+1} < \underbrace{z^{n+1} \cdot \ldots \cdot z^{n+1}}_{m+1 \text{ раз}} = z^{(n+1)(m+1)}$, т. е. $(x : z) < (n+1)(m+1)$ и $(x : z)+1 \leqslant (n+1)(m+1)$.

Если $z^{(n+1)(m+1)}$ не определено, то по определению функции $(x : y)$ также получаем $(x : z) < (n+1)(m+1)$ и $(x : z)+1 \leqslant (n+1)(m+1)$. ■

Лемма 2. ***а) Пусть $a \neq x_0$ любая фиксированная величина из I. Найдется такая последовательность $\{z_n\} \in I$, $z_n \neq x_0$, что последовательность $\{(a : z_n)\}$, $n \in N$ строго возрастает.***

б) В множестве I найдется счетное всюду плотное подмножество I_∞, т. е. такое счетное подмножество, что для любых $x, y \in I$, $y < x$ найдется такое $z \in I_\infty$, что $y < z < x$.

▷ а) Произвольно выберем $z_1 \in I$, $z_1 \neq x_0$ и произвольно выберем z_1' из условия $x_0 < z_1' < z_1$. По аксиоме 4 найдется такое z_1'', что $z_1 = z_1' \cdot z_1''$. Из аксиомы 3 следует: $x_0 < z_1'' < z_1$. Положим $z_2 = \min\{z_1', z_1''\}$.

Произвольно выберем z_2' из условия $x_0 < z_2' < z_2$. По аксиоме 4 $z_2 = z_2' \cdot z_2''$. Положим $z_3 = \min\{z_2', z_2''\}$ и т. д.

По построению $z_{n+1}^2 < z_n$ и $(a : z_{n+1}) \geqslant 2 \cdot (a : z_n)$. ■

▷ б) Обозначим $\{z_n\}$ последовательность величин, построенную в п. а). Для каждого $n \in N$ определим множество

$$I_n \rightleftarrows \{z_n^k \mid k \in N,\ z_n^k \in I\}.$$

Положим $I_\infty = \bigcup_{n=1}^{\infty} I_n \subseteq I$.

Докажем, что для любого $z \in I$, $z \neq x_0$ найдется $z_n < z$. Действительно, если $z_n \geqslant z$ при всех n, то $(a : z_n) < (a : z)$ при всех n, что противоречит неограниченности последовательности $\{(a : z_n)\}$.

Тем самым лемма 2,б уже доказана для $y = x_0$. Рассмотрим случай, когда $x_0 < y < x$. Пусть z такое, что $x = y \cdot z$. Тогда $x_0 < z$ и можно выбрать z_n так, что $x_0 < z_n < z$. Рассмотрим элемент $z_n^{((y : z_n)+1)} \in I_\infty$. По определению числа $(y : z_n)$ он больше y. Кроме того, из неравенства $z_n^{(y : z_n)} \leqslant y$, $z_n < z$ и равенства $x = y \cdot z$ следует, что

$$z_n^{((y : z_n)+1)} = z_n^{(y : z_n)} \cdot z_n \leqslant y \cdot z_n < y \cdot z = x. \quad ■$$

Фиксируем некоторый элемент $a \in I$, $a \neq x_0$ («единицу измерения»). Определим для каждого $z \in I$, $z \neq x_0$ функцию $f_z : I \to \boldsymbol{Q}_{>0}$ следующим образом: $f_z(x) \rightleftarrows \dfrac{(x : z)+1}{(a : z)+1}$.

Отметим, что по лемме 1,б для всех x, $z \in I$, $z \neq x_0$ выполняется $f_z(x) \leqslant (x : a)+1$.

Учитывая пример, разобранный в главе III после формулировки теоремы 1, можно интуитивно представить себе $f_z(x)$ как измерение x относительно единицы измерения a с точностью z (с избытком). Сейчас, используя теорему Тихонова о компактности топологического произведения топологических компактных пространств относительно топологии произведения, мы построим функцию f, предельную для семейства функций $\{f_z \mid z \in I\}$, которая и будет искомым изоморфизмом структур $\langle V, I, \cdot, x_0, \leqslant \rangle$ и $\langle \boldsymbol{R}_{>0}, [0, \lambda|+, C, \leqslant \rangle$.

Для каждого $x \in I$ рассмотрим отрезок вещественных чисел $[0, (x : a)+1]$ и рассмотрим декартово произведение этих отрезков по всем $x \in I$, т. е. определим

$$X \rightleftarrows \prod_{x \in I} [0, (x : a)+1].$$

Элементами X являются все такие функции вида $\varphi : I \to \boldsymbol{R}$, что $0 \leqslant \varphi(x) \leqslant (x : a)+1$, $\forall x \in I$.

Так как по определению $f_z(x)$ неотрицательно, то при любом $z \in I$, $z \neq x_0$, получаем $f_z \in X$.

Открытая окрестность $\mathcal{O}(\varphi)$ элемента φ. Пусть $\varphi \in X$. Фиксируем конечное число элементов $x_1, x_2, \ldots, x_n \in I$ и конечное число строго положительных чисел $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$. Для каждого значения индекса i фиксируем в отрезке $[0, (x : a)+1]$ открытую ε_i окрестность $\varphi(x_i)$.

Напомним, что топология произведения $\mathcal{T}_X$ в множестве X определяется таким образом, что состоит из всех тех функций $\psi \in X$, значения которых в точках x_i попадают в фиксированные ε_i — окрестности (рис. 102). В остальных точках значения функции произвольны. Иными словами

$$\mathcal{O}(\varphi) \rightleftarrows \mathcal{O}_{x_1, \ldots, x_n,\ \varepsilon_1, \ldots, \varepsilon_n}(\varphi) \rightleftarrows \{\psi \in X \mid |\varphi(x_i) - \psi(x_i)| < \varepsilon_i,\ \forall i (1 \leqslant i \leqslant n)\}.$$

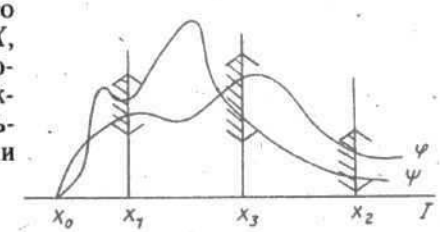


Рис. 102

Упражнение. *Семейство множеств вида* $\mathscr{O}_{x_1,\ldots,x_n,\ldots,\varepsilon_1,\ldots,\varepsilon_n}(\varphi)$, *где* $n\in\boldsymbol{N}$, $x_i\in I$, $\varepsilon_i\in\boldsymbol{R}_{>0}$, $\varphi\in X$ *образуют базу топологии в множестве* X. *Докажите.*

Лемма 3. а) ***Топологическое пространство*** $\langle X, \mathscr{T}_X\rangle$ ***компактно.***

б) ***Пусть*** $X_z\rightleftarrows\{f_{z'}\mid x_0<z'<z\}$ ***и*** $\bar{X}_z$ ***— замыкание множеств*** X_z ***в топологическом пространстве*** X. ***Тогда пересечение*** $\bigcap\{\bar{X}_{\bar{z}}\mid z\in I_{>0}\}$ ***непусто.***

▷ а) Этот пункт по существу составляет содержание упомянутой выше теоремы Тихонова, его доказательство мы опустим.

б) Пусть $z_1, z_2, \ldots, z_n\in I_{>x_0}$. Тогда для $z\rightleftarrows\min\{z_1,\ldots,z_n\}$ выполняется $X_z\subseteq\bigcap_{i=1}^{n}X_{z_i}\subseteq\bigcap_{i=1}^{n}\bar{X}_{z_i}$, т. е. пересечение любого конечного числа множеств вида $\bar{X}_z$ непусто.

Допустим, что $\bigcap\{\bar{X}_z\mid z\in I_{>z_0}\}=\varnothing$. Тогда для всякого $\varphi\in X$ найдется такое $z\in I_{>z_0}$, что $\varphi\notin\bar{X}_z$ и, следовательно, φ не принадлежит $\bar{X}_z$ вместе с некоторой своей окрестностью $\mathscr{O}(\varphi)$. Получили открытое покрытие всего компактного пространства. Выберем его конечное подпокрытие $\mathscr{O}(\varphi_1)$, $\mathscr{O}(\varphi_2)$, ..., $\mathscr{O}(\varphi_n)$.

Тогда пересечение соответствующих множеств $\bar{X}_{z_1},\ldots,\bar{X}_{z_n}$ пусто, так как всякая точка $\varphi\in X$ лежит в некотором множестве $\mathscr{O}(\varphi_i)$ и по построению $\mathscr{O}(\varphi_i)_n$ не содержится в X_{z_i} и, следовательно, не содержится в $\bigcap_{i=1}^{n}\bar{X}_{z_i}$. Получили противоречие. ■

Пусть f — любой элемент из пересечения $\bigcap\{\bar{X}_z\mid z\in I_{>x_0}\}$. Докажем, что функция $f: I\to\boldsymbol{R}$ является искомым изоморфизмом.

Лемма 4. ***Фиксируем*** $x, y\in I$. ***Для всех*** $x, y\in I$ ***выполняется*** $f(x\cdot y)=f(x)+f(y)$.

▷ Добавив по единице к частям неравенства

$$(x:z)+(y:z)\leqslant((x\cdot y):z)\leqslant(x:z)+(y:z)+1$$

(лемма 1,а) и разделив их почленно $(a:z)+1$, получим, что при всех $z\in I_{>z_0}$ выполняется

$$f_z(x)+f_z(y)-\frac{1}{(a:z)+1}\leqslant f_z(x\cdot y)\leqslant f_z(x)+f_z(y),$$

т. е. $0\leqslant f_z(x)+f_z(y)-f_z(x\cdot y)\leqslant\frac{1}{(a:z)+1}$. Выберем такую последовательность z_n, что последовательность $\{(a:z_n)\}$ строго возрастает (лемма 2,а), и рассмотрим открытую окрестность

$$\mathscr{O}(f)\rightleftarrows\mathscr{O}_{x,\,y,\,x\cdot y,\,\frac{1}{n},\,\frac{1}{n},\,\frac{1}{n}}(f)=\left\{\psi\in X\mid|\psi(z)-f(z)|<\frac{1}{n}\right\}$$

для трех значений z равных x, y, $x\cdot y$.

Так как $f\in\bigcap\{\bar{X}_z\mid z\in I_{>x_0}\}$, то в окрестности $\mathscr{O}(f)$ найдется элемент g из $\bar{X}_{z_n}$. Поэтому $\mathscr{O}(f)$ окрестность и элемент g. Поэтому в $\mathscr{O}(f)$ найдется элемент из X_{z_n} вида функции $f_{z_n'}$, где $x_0<z_n'<z_n$. По определению $\mathscr{O}(f)$ имеем

$$f_{z_n'}(x)-\frac{1}{n}<f(x)<f_{z_n'}(x)+\frac{1}{n},\quad f_{z_n'}(y)-\frac{1}{n}<f(y)<f_{z_n'}(y)+\frac{1}{n},$$

$$-f_{z_n'}(x\cdot y)-\frac{1}{n}<-f(x\cdot y)<-f_{z_n'}(x\cdot y)+\frac{1}{n}.$$

Складывая эти неравенства, получим

$$-\frac{3}{n}\leqslant f_{z_n'}(x)+f_{z_n'}(y)-f_{z_n'}(x\cdot y)-\frac{3}{n}<f(x)+f(y)-f(x\cdot y)<f_{z_n'}(x)+$$
$$+f_{z_n'}(y)-f_{z_n'}(x\cdot y)+\frac{3}{n}\leqslant\frac{1}{(a:z_n')+1}+\frac{3}{n}\leqslant\frac{3}{n}+\frac{1}{(a:z_n)+1}.$$

Устремляя n к бесконечности, получим $f(x)+f(y)-f(xy)=0$. ■

Лемма 5. ***Выполняется а)*** $f(x_0)=0$.

б) f ***строго возрастающая функция на*** I.

▷ а) $f(x_0)=f(x_0\cdot x_0)=f(x_0)+f(x_0)$. Поэтому $f(x_0)=0$.

б) Допустим, что при некотором $z>x_0$ выполняется $f(z)=0$. Тогда для любого z' из условия $x_0<z'<z$ по аксиоме 4 получаем: $z=z'\cdot z''$ и $0=f(z)=f(z'\cdot z'')=f(z')+f(z'')\geqslant 0$. Следовательно, $f(z')=0$. Более того: так как для любого $x\in I$ выполняется $x=z^{(x:z)}\times z'$, где $z'<z$, то $f(x)=(x:z)\cdot f(z)+f(z')=0$ и, следовательно, f тождественно равно нулю на I. Но это противоречит равенству $f(a)=1$, которое вытекает из того, что $f_z(a)=1$ при всех $z\in I_{>x_0}$.

Итак, если $z>x_0$, то $f(z)>0$. Если теперь $x>y$, то $x=y\cdot z$, где $z>x_0$ и $f(x)=f(y\cdot x)=f(y)+f(z)>f(y)$. ■

Лемма 6. ***Для любого элемента*** $x\in I$ ***и для любого числа*** t, ***удовлетворяющего условию*** $0\leqslant t\leqslant f(x)$, ***найдется такое*** $y\in I$, ***что*** $f(y)=t$.

▷ По лемме 2,б выберем в I счетное всюду плотное подмножество I_∞, при этом в качестве z_1 выберем элемент a. Докажем, что множество $f(I_\infty)$ всюду плотно на отрезке $[0, f(x)]$. Напомним, что I_∞ включает в себя такие элементы z_n, что $(a:z_{n+1})\geqslant 2\cdot(a:z_n)$. Выполняется неравенство $(a:z_{n+1})\geqslant 2^n$. Поэтому $(z_{n+1}^{2^n})\leqslant a$.

Так как функция f строго возрастает, то $2^n\cdot f(z_{n+1})=f(z_{n+1}^{2^n})\leqslant f(a)=1$, т. е. $f(z_{n+1})\leqslant 2^{-n}$.

Следовательно, множество точек $f(x_0)$, $f(z_{n+1})$, $f(z_{n+1}^2)$, ..., $f(z_{n+1}^k)$, где $k=(x:z_{n+1})+1$, с одним и тем же шагом, меньшим 2^{-n}, расположено в порядке возрастания на отрезке $[0, f(z_{n+1}^k)]\supseteq[0, f(x)]$, поэтому множество $f(I_\infty)$ всюду плотно на $[0, f(x)]$.

Пусть t произвольное число из отрезка $[0, f(x)]$ и $\{f(y_n)\}$ — такая возрастающая последовательность чисел из множества $[0, f(x)]\cap f(I_\infty)$, предел которой равен t. В силу монотонности f последовательность $\{y_n\}$ монотонно возрастает в I_∞ и она ограничена элементом x. По аксиоме 6 существует $y\rightleftarrows\sup y_n$, который принадлежит I. По монотонности f получаем: $f(y)\geqslant\sup f\{y_n\}=t$. Допустим, что $f(y)>t$. Тогда найдется такой $z\in I_\infty$, что $t<f(z)<f(y)$, и по монотонности f получаем $y_n<z<y$, $\forall n$, что противоречит равенству $y=\sup_n y_n$. ■

Леммы 4—6 доказывают теорему 1.

ПРИЛОЖЕНИЕ 3

(к главе IV)

Доказательство некоторых вспомогательных алгебраических утверждений

Мы начнем это приложение со следующего упражнения, продолжая нумерацию из главы IV.

Упражнение 20. *Докажите, если поле L конечномерно над полями P и K и $P \subseteq K$, то поле K* конечномерно над полем P и

$$\dim_P L = \dim_K L \cdot \dim_P K.$$

Лемма 10. ***Для того чтобы поле L являлось нормальным расширением поля P, необходимо и достаточно, чтобы поле L являлось полем разложения некоторого многочлена над P.***

▷ Пусть поле L является нормальным расширением поля P, т. е.:

а) поле L конечномерно над полем P;

б) если L содержит один корень неприводимого над P уравнения, то L содержит и все корни этого уравнения.

Условие конечномерности L над P означает наличие конечного базиса $\{k_1, \ldots, k_s\}$ в поле L над полем P. В частности, $L = P(k_1, \ldots, k_s)$.

Числа $k_1, \ldots, k_s$ являются корнями некоторых неприводимых уравнений над P. Действительно, рассмотрим, например, последовательность степеней числа k_1:

$$1 = k_1^0,\ k_1,\ k_1^2,\ k_1^3,\ \ldots .$$

В силу конечномерности L и P при некотором n число k_1^n равно линейной комбинации $t_0 + t_1 k_1 + \ldots + t_{n-1} k_1^{n-1}$, где $t_i \in P$. Поэтому k_1 является корнем уравнения $g(x) = x^n - t_{n-1}x^{n-1} - \ldots - t_1 x - t_0 = 0$ над P. Раскладывая многочлен $g(x)$ над полем P на неприводимые множители, получим неприводимый многочлен, корнем которого является k.

Для чисел $k_1, \ldots, k_s$ обозначим эти неприводимые многочлены соответственно $g_1, \ldots, g_s$. По условию б все корни всех этих многочленов содержатся в поле L. Поэтому все корни многочлена $g_0(x) \rightleftarrows g_1(x) \cdots g_s(x)$ также содержатся в поле L. Следовательно, поле L содержит поле разложения P_{g_0} многочлена g_0, т. е. $P_{g_0} \subseteq L$.

С другой стороны, числа $k_1, \ldots, k_s$ — некоторые корни многочлена g_0, поэтому $L = P(k_1, \ldots, k_s) \subseteq P_{g_0}$, и, следовательно, $L = P_{g_0}$.

Докажем теперь обратное утверждение: поле разложения P_f любого многочлена f над P является нормальным расширением поля P.

Конечномерность поля P_f над полем P вытекает из упражнения 4, предложения 4,б и упражнения 20.

Пусть k — некоторый элемент поля P_f, являющийся корнем некоторого неприводимого над P многочлена h_1. По предложению 4,в найдется такой многочлен $g(x_1, \ldots, x_n)$ над P, что $k = g(\alpha_1, \ldots, \alpha_n)$, где $\alpha_1, \ldots, \alpha_n$ — все корни многочлена f.

Рассмотрим многочлен

$$h(x) \rightleftarrows \prod_{\sigma \in S_n} (x - g_\sigma(\alpha_1, \ldots, \alpha_n))$$

степени $n!$, где многочлены $g_\sigma(\cdot)$ определяются равенствами

$$g_\sigma(x_1, \ldots, x_n) \rightleftarrows g(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

Докажем, что коэффициенты многочлена $h(x)$ лежат в поле P. Рассмотрим, например, коэффициент при $x^{n!-1}$, равный $-\sum_{\sigma \in S_n} g_\sigma(\alpha_1, \ldots, \alpha_n)$. Многочлен $-\sum_{\sigma \in S_n} g_\sigma(x_1, \ldots, x_n)$, значением которого является этот коэффициент, очевидно, не меняет своего значения под действием на его аргументы любой перестановки σ из S_n. Иными словами, он является алгебраическим инвариантом группы S_n. Такой многочлен называется симметрическим.

В курсе алгебры[1] доказывается, что значение любого симметрического многочлена над (числовым) полем P на корнях $\alpha_1, \ldots, \alpha_n$ любого многочлена f над P принадлежит полю P. Поэтому этот и, аналогичным образом, любой другой коэффициент многочлена h принадлежит P.

Итак, h — многочлен над P степени $n!$ и число $k = g(\alpha_1, \ldots, \alpha_n)$ является корнем h. По условию k является также корнем неприводимого многочлена h_1. В силу неприводимости h_1, многочлен h нацело делится на h_1, т. е. $h(x) = h_1(x) \cdot h_2(x)$.

Поэтому все корни многочлена h_1 являются корнями многочлена h, т. е. лежат среди чисел вида $g_\sigma(\alpha_1, \ldots, \alpha_n)$, $\sigma \in S_n$, которые очевидно лежат в $P_f = P(\alpha_1, \ldots, \alpha_n)$. ■

Следствие. Если поле K является полем разложения какого-то многочлена f над полем P, то число элементов группы Галуа $G(K, P)$ совпадает с размерностью $\dim_P K$ поля K над полем P.

▷ По условию поле K имеет вид $K = P(x_1, \ldots, x_n)$, где $x_1, \ldots, x_n$ корни многочлена f.

По упражнению 13 найдется такое число a, что $K = P(a)$. По конечномерности K над P существует неприводимый над P многочлен g, корнем которого является число a. Все его корни различны. По лемме 10 поле K является нормальным расширением поля и поэтому содержит все корни уравнения $g(x) = 0$.

Обозначим эти корни $x_1 = a, x_2, \ldots, x_m$. По предложению 4,б выполняется $\dim_P K = \deg g = m$.

Установим биекцию между элементами группы $G(K, P)$ и корнями $x_1, \ldots, x_m$. Элементу $W \in G(K, P)$ по определению соответствует корень $W(a)$ уравнения $g(x) = 0$, т. е. отображение из $G(K, P)$ в множество $\{x_1, \ldots, x_m\}$ зададим соответствием $W \mapsto W(a)$. Это соответствие инъективно, так как

$$W(a) = W_1(a) \Rightarrow W_1^{-1}(W(a)) = a$$

[1] См.: *Куликов Л. Я.*, с. 498.

и по предложению 4,б функция $(W_1^{-1}\circ W)$ тождественна на K, т. е. $W^{-1}\circ W=\mathrm{id}$, и поэтому $W_1=W$. Докажем сюръективность этого отображения.

Пусть x_i любой из корней $x_1, \ldots, x_m$. По все тому же предложению 4,б любой элемент k из K однозначно представим в виде $k=t_0+t_1\cdot a+\ldots+t_{m-1}\cdot a^{m-1}$. Определим функцию W_i из поля K в себя следующим образом:

$$W_i(t_0+t_1a+\ldots+t_{m-1}a^{m-1})\rightleftarrows t_0+t_1x_i+\ldots+t_{m-1}\cdot x_i^{m-1}.$$

Так как $W_i(a)=x_i$, то для доказательства сюръективности отображения $W\mapsto W(a)$ достаточно проверить, что функция W_i является элементом группы Галуа $G(K, P)$. Ясно, что $W_i|P=\mathrm{id}$ и что $W_i(k+t)=W_i(k)+W_i(t)$ для всех $k_1, t\in K$. Остается проверить, что $W_i(k\cdot t)=W_i(k)\cdot W_i(t)$ для всех $k, t\in K$.

Пусть h — произвольный многочлен над P и многочлен r является остатком от деления h на g, т. е.

$$h(x)\equiv h_0(x)\cdot g(x)+r(x).$$

Так как степень многочлена r строго меньше m, то по определению W_i получаем, что $W_i(r(a))=r(W_i(a))$. Так как числа $W_i(a)=x_i$ и a являются корнями многочлена g, то получаем:

$$W_i(h(a))=W_i(r(a))=r(W_i(a))=h(W_i(a)).$$

Итак, для любого многочлена h над полем P выполняется

$$W_i(h(a))=h(W_i(a)).$$

Пусть теперь k и t произвольные элементы поля K. Тогда предложению 4,в найдутся такие многочлены h_1 и h_2 над полем P, что $k=h_1(a)$ и $t=h_2(a)$.

Используя доказанное выше равенство $W_i(h(a))=h(W_i(a))$, получаем

$$W_i(k\cdot t)=W_i(h_1(a)\cdot h_2(a))=W_i((h_1h_2)(a))=(h_1\cdot h_2)(W_i(a))=h_1(W_i(a))\times$$
$$\times h_2(W_i(a))=W_i(h_1(a))\,W_i(h_2(a))=W_i(k)\,W_i(t).$$

Теперь нетрудно установить инъективность функции $W_i: K\to K$, так как в силу ее линейности достаточно проверить равенство $\mathrm{Ker}\ W_i=\{0\}$.

Действительно, если k — отличный от нуля элемент K, для которого $W_i(k)=0$, то

$$1=W_i(1)=W_i(k\cdot k^{-1})=W_i(k)\cdot W_i(k^{-1}),$$

что противоречит равенству $W_i(k)=0$.

Итак, построенная функция W_i линейно и инъективно отображает поле K в себя. Рассматривая поле K как векторное пространство над полем P и учитывая, что всякая линейная инъективная функция из конечномерного векторного пространства в себя является сюръективной функцией, получаем, что функция W_i сюръективна и, значит,

она является изоморфизмом поля K, оставляющим на месте любой элемент поля P. Другими словами, доказано, что $W_i\in G(K, P)$. ■

Лемма 11. ***Любое радикальное расширение поля P содержится в некотором нормальном радикальном расширении поля P.***

▷ Доказательство проведем индукцией по длине s радикального ряда $P\subset P_1\subset\ldots\subset P_{s-1}\subset P_s\rightleftarrows K$, в котором по определению все поля P_i являются простыми радикальными расширениями полей P_{i-1}.

В частности, $K=P_{s-1}(\zeta, \sqrt[m]{a-1})$, где $a\in P_{s-1}$ и $\zeta=e^{\frac{2\pi}{m}i}$.

Если $s=1$, то поле $K=P_s$, как это проверялось на с. 255, является полем разложения уравнения $x^m-a=0$ над P. Следовательно, по лемме 10 K само является нормальным расширением поля P.

Допустим, что лемма доказана для всех полей K, имеющих радикальный ряд длины, меньшей s. Тогда поле P_{s-1} уже содержится в некотором нормальном радикальном расширении L поля P. Построим искомое радикальное расширение K поля P.

Для этого рассмотрим уравнение

$$x^m-a=0,\ a\in P_{s-1}.$$

Полем его разложения является поле $K=P_s$. Для числа a рассмотрим любой неприводимый над P многочлен $g(x)$, среди корней которого содержится число a. Обозначим все его корни $a_1=a, a_2, \ldots, a_k$. В силу нормальности L над P все эти корни принадлежат L.

Пусть $\alpha_1, \alpha_2, \ldots, \alpha_k$ — произвольные корни соответственно уравнений $x^m-a_1=0, \ldots, x^m-a_k=0$ над L. Положим поле $\bar{K}$ равным расширению поля L с помощью чисел $\alpha_1, \ldots, \alpha_k$ и числа $\zeta=e^{\frac{2\pi}{m}i}$, т. е.

$$\bar{K}\rightleftarrows L(\zeta, \alpha_1, \ldots, \alpha_k).$$

Ясно, что поле $\bar{K}$ является радикальным расширением поля L. Так как L радикально над P, а $\bar{K}$ радикально над L, то поле $\bar{K}$ радикально над полем P.

Докажем, что $K\subseteq\bar{K}$. Для этого достаточно заметить, что все корни уравнения

$$x^m-a=0,\ a\in P_{s-1}$$

исчерпываются числами $\alpha_1, \zeta\alpha_1, \zeta^2\alpha_1, \ldots, \zeta^{m-1}\alpha_1$, т. е. поле K содержится в поле $P_{s-1}(\zeta, \alpha_1)$, и так как $P_{s-1}\subseteq L$, то

$$K\subseteq L(\zeta, \alpha_1)\subseteq L(\zeta, \alpha_1, \ldots, \alpha_k)=\bar{K}.$$

Наконец, докажем нормальность $\bar{K}$ над P. Так как L нормально над P, то по лемме 10 поле L является полем разложения некоторого многочлена $h(x)$ над P. Обозначим все его корни $x_1, \ldots, x_n$. Тогда $L=P(x_1, \ldots, x_n)$ и $\bar{K}=P(x_1, \ldots, x_n, \zeta, \alpha_1, \ldots, \alpha_k)$. Рассмотрим многочлен

$$f(x)\rightleftarrows h(x)\cdot g(x^m)=h(x)(x^m-a_1)(x^m-a_2)\ldots(x-a_k),$$

где второе равенство получается из соотношения $g(x)=(x-a_1)...$ $...(x-a_k)$.

Коэффициенты многочлена f лежат в поле P, так как коэффициенты многочленов h и g лежат в P. Поле разложения P_f многочлена f над P содержит поле $\bar{K}$, так как числа $x_1, ..., x_n, \alpha_1, ..., \alpha_k, \zeta=\frac{\zeta\cdot\alpha_i}{\alpha_i}$ содержатся в $\bar{K}$.

Наоборот, любой корень многочлена f равен либо одному из чисел $x_1, ..., x_n$, либо одному из чисел $\zeta^j\cdot a_i$, $0\leqslant j\leqslant n$, и поэтому $P_f\subseteq\bar{K}$. Следовательно, поле $\bar{K}$ совпадает с полем разложения P_f. Отсюда по лемме 10 получаем нормальность поля $\bar{K}$ над полем P. ■

Лемма 9 (см. § 2 п. 6). ***Любое число из поля разложения P_f многочлена f над полем P, неподвижное относительно действия всех изоморфизмов из группы Галуа $G(P_f, P)$, принадлежит P.***
▷ Обозначим $K\rightleftarrows P_f$ и $L\rightleftarrows\{x\in K\mid\forall W\in G(K, P)(W(x)\}$. Ясно, что L — поле и промежуточное поле, т. е. $P\subseteq L\subseteq K$. Докажем, что $L=P$.

Сначала докажем, что группы $G(K, L)$ и $G(K, P)$ совпадают. Первая из них является подгруппой второй, так как $P\subseteq L\subseteq K$. Наоборот, если $W\in G(K, P)$ и $x\in L$, то по определению поля L $W(x)=x$, т. е. $W\in G(K, L)$.

По лемме 16 поле K является полем разложения некоторого многочлена над полем L.

Так как поле K по условию также является полем разложения некоторого многочлена над полем P, то можно применить следствие к лемме 10 и получить, что $\dim_P K=\dim_P L$. Поскольку $P\subseteq L\subseteq K$, то по упражнению 20 получаем, что $\dim_P L=1$, т. е. $L=P$. ■

Лемма 12. ***Любая подгруппа и любая фактор-группа разрешимой группы являются разрешимыми группами.***
▷ Пусть

$$G\supset G_1\supset ...\supset G_n\supset\{e\}$$

— цепочка вложенных друг в друга групп, в которой все G_i являются нормальными делителями G_{i-1} и все фактор-группы G_{i-1}/G_i циклични. Для подгруппы $H\subset G$ рассмотрим цепочку групп

$$H\supset H_1\supset ...\supset H_n=\{e\},$$

где $H_i\rightleftarrows G_i\cap H_{i-1}$.

Можно определить группу H_i несколько иначе. Рассмотрим гомоморфизм

$$\varphi_{i-1}: H_{i-1}\to G_{i-1},\ \varphi_{i-1}(g)\rightleftarrows g.$$

Тогда группа H_i — прообраз группы G_i, т. е. $H_i=\varphi_{i-1}^{-1}(G_i)$. По упражнению 5 прообраз нормального делителя при гомоморфизме является нормальным делителем (проверьте). Поэтому из гомоморфности функции φ_{i-1} вытекает, что H_i — нормальный делитель H_{i-1}. Следовательно, определена фактор-группа H_{i-1}/H_i, которая

очевидным образом вкладывается в циклическую группу G_{i-1}/G_i. А именно $[g]_{H_{i-1}/H_i}\mapsto[g]_{G_{i-1}/G_i}$. Итак, группа H_{i-1}/H_i изоморфна подгруппе циклической группы G_{i-1}/G_i. Но всякая подгруппа циклической группы сама циклична. Следовательно, все фактор-группы H_{i-1}/H_i циклични, что доказывает разрешимость подгруппы $H\subset G$.

Если H — нормальный делитель группы G и $\varphi: G\to G|H$ — канонический гомоморфизм G на фактор-группу G/H, то из цепочки групп для G возникает цепочка групп для G/H:

$$G/H=\varphi(G)\supset\varphi(G_1)\supset ...\supset\varphi(G_n)\supset\varphi(\{e\})=\{e'\},$$

где e' — единичный элемент фактор-группы G/H. При этом $\varphi(G_i)$ как гомоморфный образ нормального делителя в G_{i-1} является нормальным делителем в $\varphi(G_{i-1})$, а также фактор-группа $\varphi(G_{i-1})/\varphi(G_i)$ изоморфна фактор-группе G_{i-1}/G_i и поэтому циклична. Следовательно, фактор-группа G/H разрешима. ■

Лемма 13. ***Любая конечная абелева группа разрешима.***

Доказательство проведем индукцией по числу n элементов в группе G. Если $n=1$, то утверждение леммы очевидно. Пусть утверждение леммы доказано для всех абелевых групп с числом элементов, меньших n, а G — абелева группа, которая состоит из n различных элементов. Возьмем любой элемент g группы G, отличный от единичного, и рассмотрим циклическую подгруппу H группы G, порожденную этим элементом. Так как $g\neq e$, то подгруппа H группы G нетривиальна, и поэтому число элементов в фактор-группе G/H строго меньше n (фактор-группу G/H можно рассматривать, так как в абелевой группе любая подгруппа является нормальным делителем). По индуктивному предположению абелева группа $G_1'\rightleftarrows G/H$ разрешима.

Рассмотрим следующую диаграмму:

$$\begin{array}{ccccccccc} G_1'\rightleftarrows G/H & \supset & G_2' & \supset & ... & \supset & G_n' & \supset & \{e'\} \\ \varphi\uparrow & & \varphi\uparrow & & & & \varphi\uparrow & & \varphi\uparrow \\ G_1\rightleftarrows G & \supset & G_2 & \supset & ... & \supset & G_n & \supset & H, \end{array}$$

где $\varphi: G\to G/H$ — канонический гомоморфизм группы G на фактор-группу G/H и $G_1'\supset G_2'\supset ...\supset G_n'\supset\{e'\}$ — разрешимый ряд для группы G_1^{-1} с единичным элементом e' и по определению $G_2\rightleftarrows\varphi^{-1}(G_2')$, ..., $G_n\rightleftarrows\varphi^{-1}(G_n')$. Отметим, что $\varphi^{-1}(\{e'\})=H$ и $\varphi^{-1}(G_1')=G$.

Ряд $G_1\supset G_2\supset ...\supset G_n\supset H$ естественно назвать опусканием ряда $G_1'\supset G_2'\supset ...\supset G_n'\supset\{e'\}$ при помощи гомоморфизма φ. По построению фактор-группы G_i'/G_{i+1}' циклические при всех $i=1, 2, ..., n$.

Так как фактор-группы $\varphi^{-1}(G_k')/\varphi^{-1}(G_{k+1}')$ изоморфны G_k'/G_{k+1}', то получаем, что и фактор-группы G_k/G_{k+1} циклични.

Учитывая, что группа H по построению циклична, окончательно получаем, что ряд

$$G_1\supset G_2\supset ...\supset G_n\supset H\supset\{e\}$$

является разрешимым рядом для группы G. ■

Лемма 14. ***Если для группы G можно указать цепочку групп***

$$G=G_0\supset G_1\supset\ldots\supset G_s=\{e\},$$

в которой все G_i — нормальные делители G_{i-1} и все группы G_{i-1}/G_i разрешимы, то группа G разрешима.

▷ Пусть $\varphi_0: G\to G/G_1$ — канонический гомоморфизм групп $G=G_0$ на фактор-группу G/G_1. Опустим, как в лемме 7, разрешимый ряд для группы G/G_1 с помощью гомоморфизма φ_0. Получим ряд групп:

$$G_0\supset H_{0,1}\supset H_{0,2}\supset\ldots\supset H_{0,n_0}\supset G_1,$$

в котором все фактор-группы $G_0/H_{0,1}$, $H_{0,1}/H_{0,2}$, ..., $H_{0,n_0}/G_1$ циклични.

Далее с помощью канонического гомоморфизма φ_1 группы G_1 на фактор-группу G_1/G_2 опустим разрешимый ряд для группы G_1/G_2. Получим ряд групп:

$$G_1\supset H_{1,1}\supset H_{1,2}\supset\ldots\supset H_{1,n_1}\supset G_2,$$

в котором все фактор-группы $G_1/H_{1,1}$, $H_{1,1}/H_{1,2}$, ..., $H_{1,n_1}/G_2$ циклични.

Произведя аналогичные опускания разрешимых рядов для групп G_{i-1}/G_i с помощью канонических гомоморфизмов φ_{i-1}: $G_{i-1}\to G_{i-1}/G_i$ при всех $i=0, 1, \ldots, s-1$, получаем разрешимый ряд для группы G:

$$\begin{aligned}G=G_0\supset H_{0,1}\supset\ldots\supset H_{0,n_0}\supset G_1=\\ =G_1\supset H_{1,1}\supset\ldots\supset H_{1,n_1}\supset G_2=\ldots=\\ =G_{s-1}\supset H_{s-1,1}\supset\ldots\supset H_{s-1,n_{s-1}}\supset G_s=\{e\}.\end{aligned}$$ ■

Лемма 15. ***Если поле K является нормальным расширением поля P и поле разложения P_f некоторого многочлена f над P содержится в K, то группа Галуа $G(P_f, P)$ изоморфна фактор-группе $G(K, P)/G(K, P_f)$.***

▷ Доказательство заключается в построении такого гомоморфизма $h: G(K, P)\to G(P_f, P)$, что его ядро $\operatorname{Ker} h$ совпадает с $G(K, P_f)$, а образ этого гомоморфизма совпадает с $G(P_f, P)$. Тогда по упражнению 7 получим, что группа $G(P_f, P)=R(h)$ изоморфна фактор-группе $G(K, P)/G(K, P_f)=G(K, P)/\operatorname{Ker} h$.

Пусть $W\in G(K, P)$. По определению положим $h(W)=W|P_f$. Чтобы доказать $h(W)\in G(P_f, P)$, достаточно проверить, что $h(W)$ переводит элементы поля P_f в элементы поля P_f. Фиксируем $k\in P_f$. В силу конечномерности K над P найдется неприводимый над P многочлен $g(x)$, корнем которого является число k, т. е. $g(k)=0$. Ясно, что число $W(k)$ также является корнем уравнения $g(x)=0$. В силу нормальности поля P_f над полем P (лемма 10) все корни уравнения $g(x)=0$ лежат в P_f. Поэтому $W(k)\in P_f$, т. е. $h(W)\in G(P_f, P)$.

Итак, построено отображение $h: G(K, P)\to G(P_f, P)$. Так как $h(W_1\circ W_2)=(W_1\circ W_2)|P_f=W_1|P_f\circ W_2|P_f=h(W_1)\circ h(W_2)$, то h является гомоморфизмом группы $G(K, P)$ в группу $G(P_f, P)$. Так как $(W\in\operatorname{Ker} h)\Leftrightarrow(W|P_f=\mathrm{id}_P)\Leftrightarrow(W\in G(K, P_f))$, то ядро этого гомоморфизма совпадает с группой Галуа $G(K, P_f)$, т. е. $\operatorname{Ker} h=G(K, P_f)$. Осталось проверить, что образ $R(h)$ гомоморфизма h совпадает с $G(P_f, P)$, т. е. проверить сюръективность гомоморфизма h.

По упражнению 13 найдется такой элемент a, что $P_f=P(a)$. Рассмотрим неприводимый над P многочлен g степени l, корнем которого является число a. По предложению 4,б размерность поля $P_f=P(a)$ над P равна l, и по следствию к лемме 10 получаем, что число элементов в группе Галуа $G(P_f, P)$ равно l.

Покажем, что фактор-множество $G(K, P)/G(K, P_f)$ содержит по крайней мере l элементов. Действительно, в силу нормальности поля K все корни $a_1\rightleftarrows a, a_2, \ldots, a_l$ уравнения $g(x)=0$ лежат в K. Для любого a_i из этих l корней найдется по предложению 6 такой изоморфизм $W\in G(K, P)$, что $W(a)=a_i$. Этот изоморфизм W обозначим W_i. Все l изоморфизмов $W_1, \ldots, W_l$ задают попарно различные классы эквивалентности в множестве $G(K, P)/G(K, P_f)$, так как $([W_j]=[W_m])\Leftrightarrow(W_j^{-1}\circ W_m\in G(K, P_f))\Rightarrow(W_j^{-1}\circ W_m(a)=a)\Rightarrow a_j=a_m$; мы использовали, что $P_f=P(a)$. Равенство $a_j=a_m$ приводит к противоречию с тем, что у неприводимого над P уравнения $g(x)=0$ нет кратных корней.

Итак, доказано неравенство $|G(K, P)/G(K, P_f)|\geqslant l$, где $|X|$ — число элементов в множестве X.

По упражнению 7 $|R(h)|=|G(K, P)/G(K, P_f)|\geqslant l$, а так как $R(h)\subset G(P_f, P)$, то $R(h)\leqslant l$, т. е. $R(h)=G(P_f, P)$. ■

Лемма 18. Пусть K поле разложения некоторого многочлена над полем P и H произвольная подгруппа группы Галуа $G(K, P)$, а также $L_H\rightleftarrows\{k\in K\mid\forall W\in H\ \ (W(k)=k)\}$ — промежуточное поле, т. е. $P\subseteq L_H\subseteq G(K, P)$. Тогда $G(K, L)=H$.

▷ Очевидно, $H\subseteq G(K, L_H)$. Обозначим m число элементов в группе H. Если порядок группы $G(K, L_H)$ меньше или равен m, то все доказано. По следствию к лемме 10 порядок группы $G(K, L_H)$ равен $\dim_{L_H}K$, т. е. нужно показать, что $(\dim_{L_H}K)\subseteq m$.

Перенумеруем все элементы группы H, начиная, как обычно, с единичного элемента: $W_1\rightleftarrows\mathrm{id}, W_2, \ldots, W_m$. Поле K, как поле разложения, имеет вид $K=P(x_1, \ldots, x_n)$, где $x_1, \ldots, x_n$ все корни одного и того же уравнения. По упражнению 13 найдется такое число $a\in K_m$, что $K=P(a)$.

Рассмотрим многочлен $h(x)=\prod\limits_{i=1}^{m}(x-W_i(a))$. Множество его корней совпадает с множеством чисел $\{a, W_2(a), \ldots, W_m(a)\}$. Для любого изоморфизма $W\in H$ это множество чисел переходит в себя, т. е. $\{a, \ldots, W_m(a)\}=\{W(a), \ldots, (W\cdot W_m)(a)\}$. Поэтому любой симметрический многочлен от корней $a, \ldots, W_m(a)$ и, в частности, любой коэффициент многочлена $h(x)$ не меняется под действием любого W из H. Следовательно, все коэффициенты многочлена h принадлежат L_H, т. е. h многочлен над L_H.

Если h приводим над L_H, то перейдем к неприводимому над L_H сомножителю h_1, корнем которого является число a.

Итак, $\deg h_1 \leqslant \deg h = m$. По предложению 4б выполняется $\dim_{L_H} K = \deg h_1 \leqslant m$, так как $K = L_H(a)$. ■

ПРИЛОЖЕНИЕ 4

СФЕРИЧЕСКАЯ, ГИПЕРБОЛИЧЕСКАЯ И ЭЛЛИПТИЧЕСКАЯ ПЛОСКОСТИ

§ 1. Точки, прямые и отрезки в сферической, эллиптической и гиперболической плоскостях

1. В данном приложении все ссылки без указания главы относятся к главе II.

Определение 1. **а) Сферической точкой (s-точкой) называется любая точка $\langle x, y, z\rangle$ из R^3, для которой $x^2+y^2+z^2=1$, т. е. любая точка сферы радиуса 1 с центром в точке $\bar{O} \rightleftarrows \langle 0, 0, 0\rangle \in R^3$.**

б) Сферической прямой (s-прямой) называется множество s-точек, лежащих в любой плоскости, проходящей через начало координат $\bar{O}$ (рис. 103).

Определим предикат «s-точка A лежит на s-прямой l» как теоретико-множественную принадлежность, т. е. $A \in l$. Аналогично определяется предикат «s-точки A и B равны между собой»: он означает теоретико-множественное равенство троек чисел A и B.

Легко видеть, что при таких определениях на всякой s-прямой l лежат хотя бы две различные s-точки A и B и найдутся три попарно разные s-точки, не лежащие ни на какой одной s-прямой l. Тем самым при нашем определении «точек», «прямых» и предикатов «лежит», «равно» выполняются аксиомы I_2 и I_3 из аксиоматики Евклида — Гильберта (см. гл II, § 3). Однако аксиома I_1 не выполняется для s-точек и s-прямых.

Действительно, через любые две диаметрально противоположные s-точки можно провести континуальное число различных s-прямых. Отметим, что нарушение аксиомы I_1 происходит только для диаметрально противоположных s-точек, так как если s-точки A и B не диаметрально противоположны, то точки $\bar{O}$, A, B не лежат на одной прямой в R^3 и через них можно провести единственную плоскость, которая в пересечении с S^2 даст ту единственную s-прямую, на которой лежат s-точки A и B.

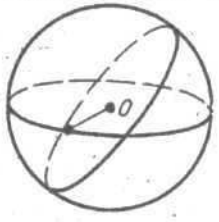


Рис. 103

Сферической плоскостью называется множество S^2. Модель аксиоматически определенной эллиптической плоскости (см. § 3) получается из сферической плоскости в результате попытки

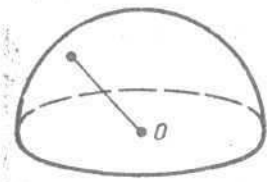


Рис. 104

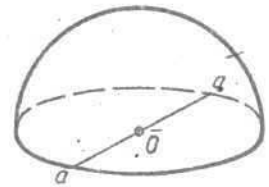


Рис. 105

обеспечить выполнение аксиомы I_1, т. е. попытки приблизить сферическую плоскость к евклидовой. Для этого нужно отождествить диаметрально противоположные точки в S^2.

2. Определение 2. **а) Эллиптической точкой (e-точкой) называется любая неупорядоченная пара диаметрально противоположных точек в S^2.**

б) Эллиптической прямой (e-прямой) называется множество неупорядоченных пар диаметрально противоположных точек, содержащихся в пересечении S^2 с плоскостью, проходящей через начало координат $\bar{O}$.

Следовательно, e-точки и e-прямые — это проективные точки и проективные прямые в уже знакомой нам проективной плоскости P_2 (см. § 2). Как и в случае s-плоскости, предикаты «e-точка A лежит на e-прямой» и «e-точка A равна e-точке B» определяются теоретико-множественно.

Упражнение 1. *Проверьте для множеств e-точек и e-прямых выполнение аксиом первой группы из аксиоматики Евклида-Гильберта.*

Замечание. Как и в случае проективной плоскости P_2, можно не рассматривать точек нижней полусферы:

$$S^2_- \rightleftarrows \{\langle x, y, z\rangle \in S^2 \mid z < 0\}.$$

Напомним, что верхняя полусфера определяется как $S^2_+ \rightleftarrows \{\langle x, y, z\rangle \in S^2 \mid z \geqslant 0\}$.

При таком подходе e-точкой называется любая точка верхней полусферы S^2_+, в которой точки границы заменены парами диаметрально противоположных точек (рис. 104). При этом e-прямой является верхняя полуокружность любого экватора, у которой концы отождествлены (рис. 105). Из такого представления видно, что e-прямые «похожи» на окружности.

Перейдем к аксиомам второй группы. В первую очередь отметим, что в этих s- и e-плоскостях нельзя ввести понятие «между», согласующееся с понятием «конгруэнтно», которое определяется с помощью движений в S^2. Иными словами, нельзя определить предикат «между», для которого выполняются все аксиомы второй и третьей групп аксиоматики Евклида — Гильберта. Пока иллюстрируем это обстоятельство следующим образом. Если для s-точек A, B, C, изображенных на рисунке 106, s-точка B лежит между A и C, то, выполняя движения — повороты на 120° и 240°, по-

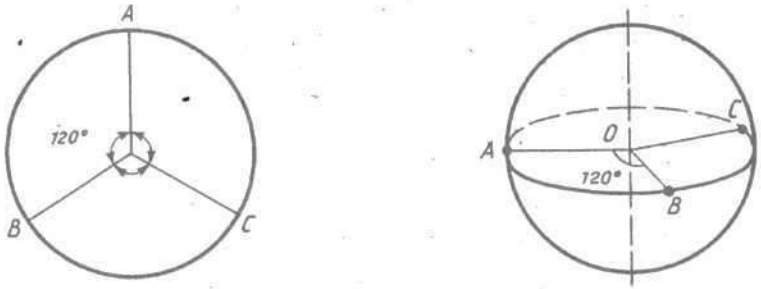


Рис. 106

лучим, что A лежит между B и C и C лежит между A и B, что противоречит аксиоме II_2. Здесь основным моментом является то, что поворот вокруг оси, перпендикулярной к плоскости ABC, является s-движением, а s-движение сохраняет предикат между.

Аналогичное рассуждение годится для e-плоскости. Поэтому в s- и e-плоскостях понятие отрезка между двумя точками приходится вводить особо, не используя понятия «лежать между». Ограничимся случаем s-плоскости, в e-плоскости определение отрезка аналогично.

Определение 3. **Пусть A и B — две различные s-точки. Если A и B диаметрально противоположны, то s-отрезком $[A, B]_s$ называется пересечение S^2 с любой замкнутой полуплоскостью, проходящей через O, границей которой является прямая (AB). Если A и B не диаметрально противоположные точки, то s-отрезком $[A, B]_s$ называется пересечение S^2 с тем геометрическим углом $\angle AOB$, радианная мера которого меньше π.**

Упражнение 2. *Если в s-плоскости ввести понятие «точка C лежит между точками A и B», как $C \in [A, B]_s$, то*

а) аксиомы II_2 и II_3 выполняются;

б) аксиома II_1 не выполняется;

в) любая s-точка лежит между любыми диаметрально противоположными s-точками. Докажите.

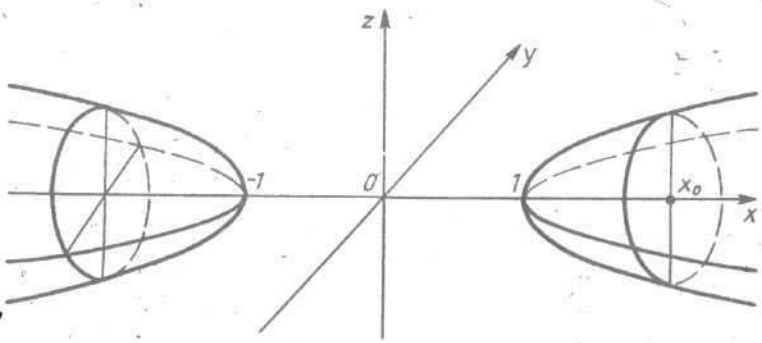


Рис. 107

3. Определение 4. **а) Гиперболической точкой (h-точкой) называется любая точка $\langle x, y, z\rangle$ из R^3, для которой $-x^2+y^2+z^2=-1$ и $x>0$. Множество всех h-точек назовем h-плоскостью (рис. 107).**

б) Гиперболической прямой (h-прямой) называется непустое множество всех h-точек, лежащих в некоторой плоскости, проходящей через начало координат O (рис. 108).

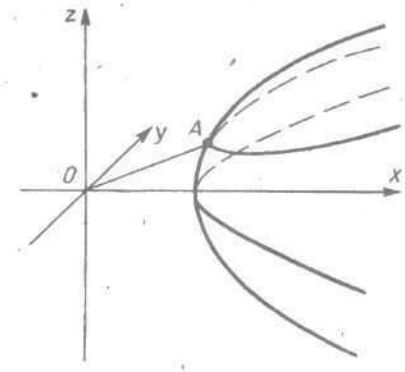


Рис. 108

Из курса аналитической геометрии хорошо известно, что поверхность $S_1^2 \rightleftarrows \{\langle x, y, z\rangle \in \boldsymbol{R}^3 \mid -x^2+y^2+z^2=-1\}$ является двуполостным гиперболоидом с вершинами в точках $\langle 1, 0, 0\rangle$ и $\langle -1, 0, 0\rangle$. В пересечении с плоскостями $y=0$ или $z=0$ этот гиперболоид образует гиперболы $-x^2+z^2=-1$ или $-x^2+y^2=-1$ с асимптотами $x=\pm z$ или $x=\pm y$. В пересечениях с вертикальными плоскостями $x=x_0$, где $|x_0|\geqslant 1$, гиперболоид S_1^2 образует окружности $\begin{cases} x=x_0 \\ y^2+z^2=x_0^2-1 \end{cases}$, (рис. 107).

Упражнение 3. *Докажите, что:*

а) функция $\langle x, y, z\rangle \mapsto -x^2+y^2+z^2$ непрерывна;

б) множество S_1^2 замкнуто и неограничено в $\boldsymbol{R}^3$;

в) множество значений непрерывной функции $\langle x, y, z\rangle \mapsto x$, определенной на S_1^2, состоит из двух лучей $]-\infty, -1]$ и $[1, +\infty[$;

г) в топологии, индуцированной пространством $\boldsymbol{R}^3$, множество S_1^2 разбивается на два непересекающихся замкнутых и связных подмножества ${}^+S_1^2 \rightleftarrows \{\langle x, y, z\rangle \in S_1^2 \mid x \geqslant 1\}$ и ${}^-S_1^2 \rightleftarrows S_1^2 \mid {}^+S_1^2$.

Часто гиперболоид S_1^2 называют гиперболической сферой. В дальнейшем будем обозначать S_1^2, ${}^+S_1^2$ и ${}^-S_1^2$ соответственно S_h, $S_{h,+}$ и $S_{h,-}$.

Упражнение 3 показывает, что гиперболическая сфера является *дизъюнктным* объединением своих «полусфер» $S_{h,+}$ и $S_{h,-}$ с «вершинами» в $\langle 1, 0, 0\rangle$ и $\langle -1, 0, 0\rangle$ соответственно.

Так как полусферы $S_{h,+}$ и $S_{h,-}$ не пересекаются, то рассмотрение пар диаметрально противоположных точек по аналогии с переходом от сферической планиметрии к эллиптической эквивалентно рассмотрению точек какой-нибудь одной из полусфер. Согласно определению 4 h-плоскость совпадает с полусферой $S_{h,+}$. В $S_{h,+}$ всегда рассматривается топология, индуцированная множеством $\boldsymbol{R}^3$.

Часто удобно следующее параметрическое задание h-плоскости $S_{h,+}$:

$$(*)\quad \begin{cases} x=x(t,\ \varphi)=\operatorname{ch} t, \\ y=y(t,\ \varphi)=\operatorname{sh} t\cdot\cos\varphi, \\ z=z(t,\ \varphi)=\operatorname{sh} t\cdot\sin\varphi, \end{cases}$$

где $0 \leqslant t < \infty$, $0 \leqslant \varphi < 2\pi$, и гиперболические косинус и синус определяются равенствами

$$\operatorname{ch} t = \frac{e^t + e^{-t}}{2}, \quad \operatorname{sh} t = \frac{e^t - e^{-t}}{2}.$$

Упражнение 4. *а) Проверьте, что функции x, y, z, определенные в $(*)$, удовлетворяют уравнению $-x^2+y^2+z^2=-1$.*

б) Докажите, что $(*)$ задает гомеоморфизм полосы $]0, +\infty[\times$ $\times[0, 2\pi[$ и $S_{h,+}\setminus\{\langle 1, 0, 0\rangle\}$.

в) Используя $(*)$, докажите, что для любых различных h-точек $A_1=\langle x_1, y_1, z_1\rangle$ и $A_2=\langle x_2, y_2, z_2\rangle$ число $x_1\cdot x_2+y_1\cdot y_2+z_1\cdot z_2$ меньше -1.

г) Используя п. в), докажите, что для любых h-точек A_1 и A_2 и для любой точки B' из обычного интервала $]A_1, A_2[$ интервал $]0, B'[$ пересечет h-плоскость в некоторой h-точке B и B' — не h-точка.

Из курса аналитической геометрии известно, что пересечение плоскости с поверхностью второго порядка в $\boldsymbol{R}^3$ либо пусто, либо является кривой в этой плоскости, порядка не выше двух, т. е. является одним из следующих множеств в плоскости: точка, прямая, пара пересекающихся прямых, пара параллельных прямых, эллипс, гипербола, парабола.

Предложение 1. ***Любая h-прямая является ветвью некоторой гиперболы.***

▷ Пусть h-прямая l есть пересечение $S_{h,+}$ и некоторой плоскости π, проходящей через точку $\bar{O}$, а l' есть пересечение $S_{h,-}$ с той же плоскостью π, т. е. $l \cup l'$ есть пересечение плоскости π с поверхностью

$$S_h = \{\langle x, y, z\rangle \in \boldsymbol{R}^3 \mid -x^2+y^2+z^2=-1\}$$

второго порядка в $\boldsymbol{R}^3$. Тогда $l \cup l'$ есть кривая в плоскости порядка не выше двух. Так как плоскость π и двуполостный гиперболоид S_h симметричны относительно начала координат $\bar{O}$, то и кривая $l \cup l'$ в плоскости π симметрична относительно $\bar{O}$. Так как поверхность S_h не связна, то и кривая $l \cup l'$ не связна. Поэтому $l \cup l'$ есть либо пара параллельных прямых, либо гипербола. Докажем, что первое невозможно. Допустим противное и на одной из этих двух параллельных прямых возьмем три различные точки A, B, C. Какие-то две из этих точек принадлежат одной из «половинок» $S_h : S_{h,+}$ или $S_{h,-}$. Применив, если надо, центральную симметрию, получим, что они лежат на $S_{h,+}$. Тогда по упражнению 4,г середина D отрезка между этими двумя точками не лежит в $S_{h,+}$. Точка D не может лежать и в $S_{h,-}$, так как ее первая координата не меньше единицы. Следовательно, $D \notin S_h$. Противоречие.

Итак, кривая $l \cup l'$ является гиперболой в плоскости π, а $l = (S_{h,+} \cap \pi)$ и $l' = (S_{h,-} \cap \pi)$ центрально симметричны. Следовательно, l является ветвью гиперболы $l \cup l'$. ■

Для h-точки и h-прямой предикат «лежит» определяется как теоретико-множественная принадлежность.

Понятие «между» для h-точек сводится к аналогичному понятию для лучей с началом в точке $\bar{O}$, лежащих в одной полуплоскости. А именно: пусть l — прямая, проходящая через точку $\bar{O}$, и L — открытая полуплоскость, границей которой является прямая l. Для любых лучей l_1' и l_2 с началом в точке $\bar{O}$ таких, что $l_1 \setminus \{\bar{O}\} \subset L$ $l_2 \setminus \{\bar{O}\} \subset L$, определим геометрический угол $l_1 \widehat{O} l_2$ как множество всех сумм векторов $\overline{O, A} + \overline{O, B}$, где $A \in l_1$, $B \in l_2$. Будем говорить, что луч l_3 в полуплоскости L с началом в точке $\bar{O}$ лежит между l_1 и l_2, если $l_3 \subset l_1 \widehat{O} l_2$. Пусть теперь h-точки A, B, C лежит на h-прямой $(S_{h,+} \cap$ $\cap \pi)$, где π — некоторая плоскость в $\boldsymbol{R}^3$, проходящая через точку $\bar{O}$. По предложению 1 h-прямая $(S_{h,+} \cap \pi)$ — ветвь некоторой гиперболы в плоскости π. Рассмотрим лучи l_1 и l_2 в плоскости π, являющиеся асимптотами этой ветви гиперболы. Тогда, как известно из курса аналитической геометрии, угол $l_1 \widehat{O} l_2$ меньше развернутого угла. Поэтому во множестве всех лучей $l_D \rightrightarrows [O, D)$, где $D \in (S_{h,+} \cap$ $\cap \pi)$, по указанному выше способу вводится понятие «между» и ровно один из лучей l_A, l_B, l_C лежит между двумя другими. По определению будем говорить, что соответствующая h-точка лежит между двумя другими h-точками.

Предложение 2. ***Множества всех h-точек и всех h-прямых в h-плоскости удовлетворяют аксиомам первой и второй групп аксиоматики Евклида — Гильберта.***

▷ Пусть A и B — различные h-точки и $A = \langle x, y, z\rangle$. Если точки $\bar{O}$, A, B лежат на одной прямой, то для некоторого $\lambda \in \boldsymbol{R}$ имеем: $B = \langle \lambda x, \lambda y, \lambda z\rangle$. Так как B есть h-точка, то $-\lambda^2 x^2 + \lambda^2 y^2 + \lambda^2 z^2 =$ $= -1$ и, следовательно, $\lambda^2 = 1$. Так как первые координаты h-точек не меньше единицы, то $\lambda \neq -1$.

Следовательно, $\lambda = 1$ и $A = B$, что противоречит предположению о различности A и B. Поэтому точки $\bar{O}$, A, B не лежат на одной прямой и через эти точки можно провести единственную плоскость π, которая в пересечении с $S_{h,+}$ даст единственную h-прямую, проходящую через h-точки A и B. Итак, аксиома I_1 проверена. Аксиомы I_2 и I_3 очевидно выполняются в h-плоскости.

Акскомы порядка II_1 и II_2 для h-точек выводятся из теоретико-множественных свойств предиката $\subseteq$. Проверим аксиому Π_3.

Пусть A, B, C — три различные h-точки, не лежащие на одной h-прямой, π — плоскость, проходящая через начало координат, и $l = (S_{h,+} \cap \pi)$ — соответствующая h-прямая. Допустим, что l пересекает во внутренней точке одну из сторон h-треугольника ABC. Пусть, например, l пересекает h-отрезок $[A, B]_h$ в h-точке D. Так как h-точка D лежит между h-точками A и B, то луч l_D лежит между лучами l_A и l_B, и поэтому он пересечет обычный отрезок $[A, B]$ в некоторой внутренней точке D'. Рассмотрим плоскости π и (ABC). Во-первых, они пересекаются в точке D'. Во-вторых, они не совпадают, так как $O \notin (ABC)$. Действительно, включение $O \in (ABC)$ означало бы, что h-точки A, B, C лежат на одной h-прямой $(S_{h,+} \cap$ $\cap (ABC))$, что противоречит условию. Следовательно, плоскости π и

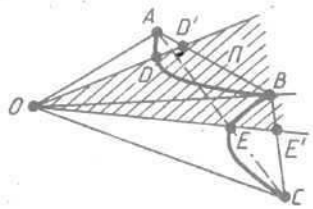


Рис. 109

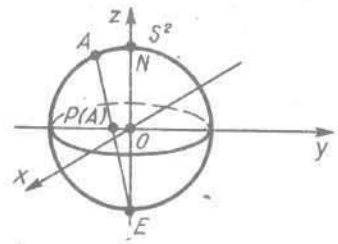


Рис. 110

(ABC) пересекаются по некоторой прямой, которая лежит в плоскости (ABC) и пересекает сторону $[A, B]$ треугольника ABC во внутренней точке D'. По аксиоме II_3, верной для евклидовой плоскости (ABC), эта прямая пересечет еще одну сторону треугольника ABC. Пусть, например, она пересекает отрезок $[B, C]$ во внутренней точке E'. Так как точка E' лежит на отрезке $[B, C]$ то, см. упражнение 4,г, луч $[O, E')$ пересечет h-отрезок $[B, C]_h$ в некоторой h-точке E. По построению $E \in S_{h,+} \cap \pi$. Следовательно, h-прямая l пересекает еще одну сторону h-треугольника с вершинами A, B, C (рис. 109). ■

4. В этом пункте рассмотрим стереографические проекции s-, e- и h-плоскостей в арифметическую плоскость $\boldsymbol{R}^2$. В каждом из этих трех случаев за центр стереографической проекции возьмем полюс соответствующей сферы. Получатся три стереографические проекции, которые позволяют «перенести» s-, e- и h-точки и s-, e- и h-прямые из соответствующих плоскостей в арифметическую плоскость $\boldsymbol{R}^2$. Возникающие в $\boldsymbol{R}^2$ после переноса образы соответствующих точек и прямых можно описать в терминах плоскости $\boldsymbol{R}^2$ без явного упоминания s-, e- и h-точек и прямых и соответствующих стереографических проекций. Часто s-, e- и h-плоскостями называют возникающие таким образом структуры в $\boldsymbol{R}^2$. Начнем с s-плоскости.

Пусть $E \rightleftarrows \langle 0, 0, -1 \rangle$ — южный полюс сферы

$$S^2 = \{\langle x, y, z \rangle \in \boldsymbol{R}^3 \mid x^2 + y^2 + z^2 = 1\}.$$

Для любой точки $A \in S^2$, отличной от E, определим $P(A)$ как точку пересечения луча $[E, A)$ с горизонтальной плоскостью (Oxy). Тогда отображение P является биекцией $S^2 \setminus \{E\}$ на (Oxy). Эта биекция называется *стереографической проекцией* «проколотой» сферы $S^2 \setminus \{E\}$ на плоскость (Oxy) (рис. 110).

Предложение 3. ***а) Пусть $A = \langle x, y, z \rangle$ — любая s-точка, отличная от E. Тогда***

$$P(A) = \left\langle \frac{x}{1+z}, \frac{y}{1+z}, 0 \right\rangle.$$

б) Пусть l — любая s-прямая, проходящая через точку E. Тогда $P(l)$ — прямая в плоскости (Oxy), проходящая через точку O.

в) Пусть l — любая s-прямая, не проходящая через точку E. Тогда $P(l)$ — окружность в плоскости (Oxy), пересекающая окружность $S^1 \rightleftarrows \{\langle x, y \rangle \mid x^2 + y^2 = 1\}$ в двух диаметрально противоположных точках.

▷ а) Пусть $B = \left\langle \frac{x}{1+z}, \frac{y}{1+z}, 0 \right\rangle$. Так как $A \neq E$, то $z \neq -1$, и поэтому векторы $\overrightarrow{E, A} = \langle x, y, 1+z \rangle$ и $\overrightarrow{E, B} = \left\langle \frac{x}{1+z}, \frac{y}{1+z}, 1 \right\rangle$ коллинеарны, т. е. точка B лежит на прямой (EA). Ясно также, что $B \in (Oxy)$. Следовательно, $P(A) = B$.

б) Если s-прямая l проходит через E, то по определению I она проходит и через северный полюс $N \rightleftarrows \langle 0, 0, 1 \rangle$, т. е. плоскость π, которая в пересечении с S^2 дает l, содержит прямую (Oz). Тогда нетрудно видеть, что $P(l)$ является прямой, по которой пересекаются плоскости π и (Oxy). Действительно, $A \in l \Rightarrow A \in \pi \cap S^2 \Rightarrow P(A) \in \in \pi \cap (Oxy)$ и, наоборот, $P(A) \in (\pi \cap (Oxy)) \Rightarrow A = P^{-1}(P(A)) \in \pi \cap S^2 \Rightarrow \Rightarrow A \in l$.

в) Пусть l — s-прямая, полученная в пересечении S^2 и плоскости π, проходящей через O. Так как $E \notin l$, то можно считать, что плоскость π задается уравнением $ax + by + z = 0$. Пусть $A = \langle x, y, z \rangle \in l = S^2 \cap \pi$ и точка $P(A)$ имеет в плоскости Oxy координаты x' и y', т. е. $x' = \frac{x}{1+z}$ и $y' = \frac{y}{1+z}$. Тогда $(x')^2 + (y')^2 = \frac{1-z^2}{(1+z)^2} = \frac{1-(ax+by)^2}{(1+z)^2} = \frac{1}{(1+z)^2} - (ax' + by')^2$, т. е.

$$(x')^2 + (y')^2 + (ax' + by')^2 = \frac{1}{(1+z)^2}.$$

С другой стороны, $x = x'(1+z)$ и $y = y'(1+z)$, т. е.

$$ax'(1+z) + by'(1+z) + z + 1 = 1,$$

и поэтому

$$ax' + by' + 1 = \frac{1}{1+z}.$$

Окончательно получаем, что

$$(x')^2 + (y')^2 + (ax' + by')^2 = (ax' + by' + 1)^2 = (ax' + by')^2 + 2ax' + 2by' + 1$$

и

$$(x' - a)^2 + (y' - b)^2 = 1 + a^2 + b^2,$$

т. е. $P(l)$ — окружность в плоскости Oxy радиуса $\sqrt{1 + a^2 + b^2}$ с центром в точке $\langle a, b \rangle$. Эта окружность пересекается с горизонтальным экватором в двух диаметрально противоположных точках, в которых прямая $\pi \cap (Oxy)$ т. е. прямая $ax' + by' = 0$ пересекает окружность $S^1 \rightleftarrows \{\langle x', y' \rangle \mid (x')^2 + (y')^2 = 1\}$ ■

Упражнение 5. *Докажите, что:*

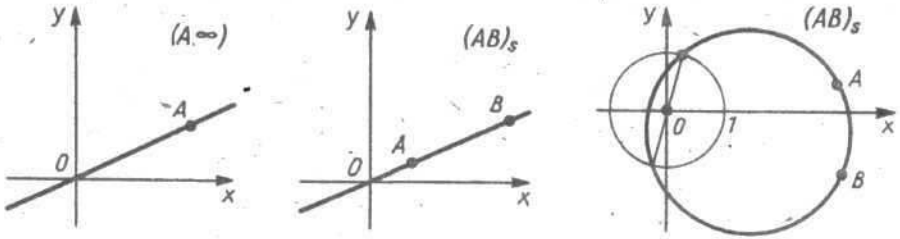


Рис. 111

а) если $B=\langle x, y\rangle$ *— точка в плоскости* Oxy, *то* $P^{-1}(B)=$ $=\left\langle \frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, \frac{1-x^2-y^2}{1+x^2+y^2} \right\rangle$;

б) если m *— прямая в плоскости* (Oxy), *проходящая через точку* O, *то* $P^{-1}(m)$ *есть* s*-прямая, проходящая через точку* E;

в) если m *— окружность в плоскости* (Oxy), *пересекающая окружность* $S^1=\{\langle x, y\rangle \mid x^2+y^2=1\}$ *в двух диаметрально противоположных точках, то* $P^{-1}(m)$ *есть* s*-прямая, не проходящая через точку* E.

Результаты предложения 3 и упражнения 5 показывают, что при указанной стереографической проекции s-точкам в s-плоскости соответствуют все точки арифметической плоскости $\boldsymbol{R}^2$ и еще одна точка, отличная от всех точек плоскости $\boldsymbol{R}^2$, которую мы назовем бесконечно удаленной точкой и обозначим ∞. Не существенно, как представлять себе эту бесконечно удаленную точку. При той же стереографической проекции s-прямым в s-плоскости соответствуют все обычные прямые в $\boldsymbol{R}^2$, проходящие через начало координат, а также все окружности в $\boldsymbol{R}^2$, пересекающие единичную окружность в диаметрально противоположных точках, рисунок 111. С помощью этой стереографической проекции s-плоскость отождествляется с множеством $\boldsymbol{R}^2\cup\{\infty\}$, а s-прямые в s-плоскости отождествляются с указанными выше прямыми и окружностями в $\boldsymbol{R}^2\cup\{\infty\}$. Поэтому s-плоскостью иногда называют само множество $\boldsymbol{R}^2\cup\{\infty\}$ вместе с указанными в нем прямыми и окружностями; иногда элементы этого множества и эти прямые и окружности называют соответственно s-точками и s-прямыми.

Теперь рассмотрим e-плоскость. Ограничим стереографическую проекцию P на верхнюю полусферу $S^2_+ \rightleftarrows \{\langle x, y\rangle \in S^2 \mid x \geqslant 0\}$. Такое ограничение задает биекцию множества S^2_+ на замкнутый единичный круг D в арифметической плоскости $\boldsymbol{R}^2$.

При таком ограничении стереографической проекции e-точкам в e-плоскости соответствуют либо точки открытого круга $D \rightleftarrows \{\langle x, y\rangle \mid x^2+y^2<1\}$, либо пары диаметрально противоположных точек единичной окружности S^1 (рис. 112). Поэтому e-плоскости соответствует круг D, в котором каждая граничная точка a заменена на (неупорядоченную) пару $\{a, -a\}$ диаметрально противоположных точек. При этом точка $-a$ заменяется на ту же пару $\{a, -a\}$. Рассмотрим две различные e-точки, которым соответствуют точки A и B открытого круга D. Если точки $\bar{O}$, A, B лежат на одной прямой, то e-прямой, проходящей через две данные e-точки, соответствует диаметр круга D, на котором лежат точки A и B. При этом концы диаметра отождествляются с одной и той же парой, состоящей из этих концов. Если точки $\bar{O}$, A, B не лежат на одной прямой, то соответствующая e-прямая в e-плоскости перейдет в ту часть s-прямой, проходящей через точки A и B, которая лежит в круге D. Другими словами, эта e-прямая перейдет в ту дугу окружности, проходящей через A и B и пересекающей S^1 в двух диаметрально противоположных точках, которая лежит в круге D. Как и выше, концы этой дуги отождествляются, рисунок 113. Иногда e-плоскостью называют множество всех точек круга D, объединенное с множеством всех пар диаметрально противоположных точек окружности S^1, вместе с указанными выше «диаметрами» и «дугами» в этом множестве. Элементы этого множества и эти «диаметры» и «дуги» иногда называют соответственно e-точками и e-прямыми. Рассмотренные структуры, состоящие из s-точек и s-прямых, e-точек и e-прямых в арифметической плоскости $\boldsymbol{R}^2$, назовем плоскими моделями s- и e-планиметрий. Теперь укажем две плоские модели h-планиметрии.

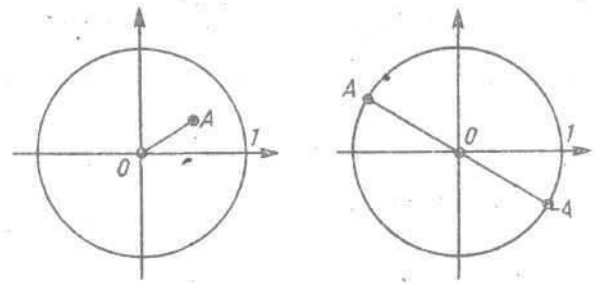


Рис. 112

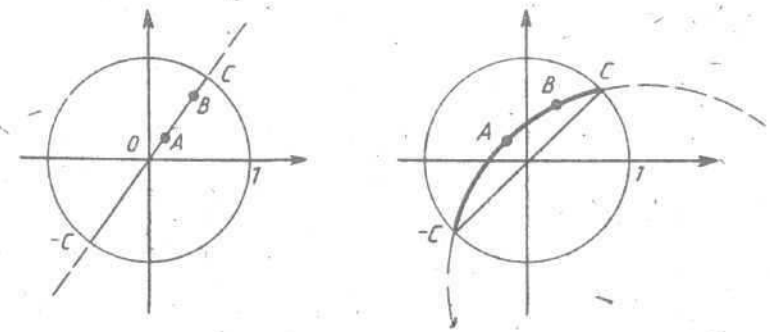


Рис. 113

Построим плоскую модель h-планиметрии, используя стереографическую проекцию из южного полюса $F=\langle -1, 0, 0\rangle$ гиперболической сферы S_h. Эта модель носит название первой модели Пуанкаре плоскости Лобачевского ($=h$-плоскости).

Для любой точки A положим по определению $P(A)$ равным точке пересечения луча $[F, A[$ и плоскости Oyz.

Предложение 4. ***а) Пусть $A=\langle x, y, z\rangle$ — любая h-точка. Тогда $P(A)=\left\langle 0, \frac{y}{1+x}, \frac{z}{1+x}\right\rangle$.***

б) Отображение P является биекцией множества $S_{h,+}$ на открытый круг $\mathring{D}\rightleftarrows\{\langle 0, y, z\rangle \mid y^2+z^2<1\}$ в плоскости (Oyz).

в) Пусть l — h-прямая, проходящая через северный полюс $N\rightleftarrows\langle 1; 0, 0\rangle$. Тогда $P(l)$ — диаметр круга $\mathring{D}$.

г) Пусть l — h-прямая, не проходящая через точку N. Тогда $P(l)$ — дуга окружности, пересекающей границу круга $\mathring{D}$ под прямым углом.

▷ а) Пусть $B\rightleftarrows\left\langle 0, \frac{y}{1+x}, \frac{z}{1+x}\right\rangle$. Так как A — h-точка, то $x\neq -1$ и векторы $\overrightarrow{E, A}=\langle 1+x, y, z\rangle$ и $\overrightarrow{E, B}=\left\langle 1, \frac{y}{1+x}, \frac{z}{1+x}\right\rangle$ коллинеарны, т. е. точка B лежит на прямой (E, A). Ясно, что $B\in(Oyz)$. Следовательно, $B=P(A)$ (рис. 114).

б) Пусть A — h-точка и $P(A)=\langle 0, Y, Z\rangle$. Тогда по п. а.

$$Y^2+Z^2=\frac{y^2+z^2}{(1+x)^2}=\frac{x^2-1}{(1+x)^2}=\frac{x^2+2x+1-2x-2}{(1+x)^2}=1-\frac{2}{1+x}<1,$$

т. е. $P(A)\in\mathring{D}$. Прямым подсчетом нетрудно проверить, что если

$$B=\langle 0, Y, Z\rangle\in\mathring{D}, \text{ то } A\rightleftarrows\left\langle \frac{1+Y^2+Z^2}{1-Y^2-Z^2}, \frac{2Y}{1-Y^2-Z^2}, \frac{2Z}{1+Y^2-Z^2}\right\rangle$$

является h-точкой и $P(A)=B$. Отсюда следует инъективность и сюръективность отображения P.

в) Пусть l — h-прямая, проходящая через N и $l=S_{h,+}\cap\pi$, где π — плоскость, проходящая через O и, тем самым, задаваемая уравнением $ax+by+cz=0$. Так как $N\in\pi$, то $a=0$. Поэтому

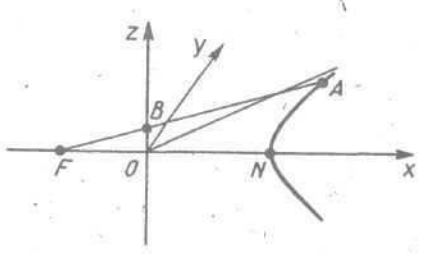


Рис. 114

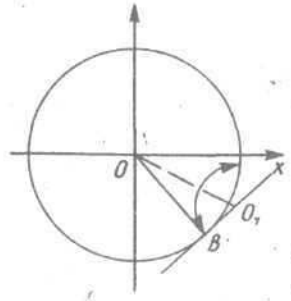


Рис. 115

для координат Y и Z точки $P(A)$, где $A\in l=S_{h,+}\cap\pi$ и $A=\langle x, y, z\rangle$, получаем, что

$$bY+cZ=\frac{by+cz}{1+x}=0.$$

Но уравнение $bY+cZ=0$ как раз и задает диаметр круга D.

г) Подставим выражения для x, y, z через Y, Z, полученные в п. б, в уравнение $ax+by+cz=0$ плоскости π, проходящей через начало координат. Получим

$$a\cdot\frac{1+Y^2+Z^2}{1-Y^2-Z^2}+b\cdot\frac{2Y}{1-Y^2-Z^2}+c\cdot\frac{2Z}{1-Y^2-Z^2}=0$$

или $a\cdot(1+Y^2+Z^2)+2bY+2cZ=0$.

Разделив на $a\neq 0$ и выделив полные квадраты, получаем

$$\left(Y+\frac{b}{a}\right)^2+\left(Z+\frac{c}{a}\right)^2=\frac{b^2+c^2-a^2}{a^2}.$$

Поэтому образом h-прямой $\pi\cap S_{h,+}$ при стереографической проекции P является дуга окружности радиуса $\frac{\sqrt{b^2+c^2-a^2}}{|a|}$ с центром в точке $O_1\rightleftarrows\left\langle -\frac{b}{a}, -\frac{c}{a}\right\rangle$. Пусть B — одна из точек пересечения этой окружности с границей круга D (рис. 115). Тогда $|OB|^2=1$, $|O_1B|^2=\frac{b^2+c^2-a^2}{a^2}$ и $|OO_1|^2=\frac{b^2}{a^2}+\frac{c^2}{a^2}$, т. е. $|OO_1|^2=|OB|^2+|O_1B|^2$, и по обратной теореме Пифагора получаем, что угол $\angle OBO_1$ — прямой. ■

Упражнение 6. *Докажите, что*

а) если $B=\langle O, Y, Z\rangle\in\mathring{D}$, то $P^{-1}(B)=\left\langle \frac{1+Y^2+Z^2}{1-Y^2-Z^2}, \frac{2Y}{1-Y^2-Z^2}, \frac{2Z}{1-Y^2-Z^2}\right\rangle$;

б) если m — диаметр круга $\mathring{D}$, то $P^{-1}(m)$ — h-прямая, проходящая через северный полюс N;

в) если m — дуга окружности, пересекающей границу круга D под прямым углом, которая лежит внутри круга, то $P^{-1}(m)$ — h-прямая, не проходящая через N.

Вторая плоская модель h-планиметрии получается из первой плоской модели h-планиметрии переходом от открытого единичного круга $\mathring{D}$ к открытой верхней полуплоскости в арифметической плоскости $\boldsymbol{R}^2$. При этом $\boldsymbol{R}^2$ и $\boldsymbol{C}$ отождествляются, а переход выполняется с помощью дробно-линейного отображения

$$z\mapsto\frac{z-1}{iz+i}, \qquad (*)$$

где $z\in\mathring{D}$. В анализе отображение $(*)$ рассматривают и для любых z из $\boldsymbol{C}\cup\{\infty\}$.

Прямое вычисление показывает, что при отображении $(*)$:

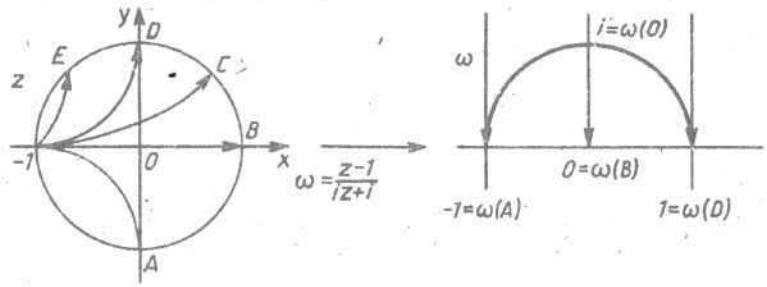


Рис. 116

а) число 0 переходит в число i, и вообще все вещественные числа, кроме -1, переходят в чисто мнимые числа;

б) число -1 переходит в ∞, а окружность $|z|=1$ переходит в вещественную ось;

в) прямые в первой модели Пуанкаре h-планиметрии перейдут в полуокружности или лучи в открытой верхней полуплоскости, которые пересекают вещественную ось под прямым углом.

Наглядно действие отображения $(*)$ можно описать следующим образом. Представим себе, что граница единичного круга D сделана из тонкой резины, а внутрь круга D налита жидкость. Мы разрываем резину в точке -1 и растягиваем границу в вещественную ось. При этом жидкость равномерно и без перемешивания выливается и заполняет всю верхнюю полуплоскость (рис. 116).

Итак, h-точками во второй плоской модели Пуанкаре h-планиметрии (плоскости Лобачевского) называются точки верхней открытой полуплоскости, а h-прямыми называются полуокружности или лучи, перпендикулярные к ее границе.

Сферическую (эллиптическую, гиперболическую) плоскость, т. е. множество всех s-точек (e-точек, h-точек) и множество всех s-прямых (e-прямых, h-прямых), далее будем обозначать $S^2(E^2, H^2)$. Построенные в п. 4 этого параграфа плоские модели будем обозначать S^2P, E^2P, H^2P_1 (первая модель Пуанкаре h-планиметрии) и H^2P_2 (вторая модель Пуанкаре h-планиметрии).

§ 2. Метрики в сферической, эллиптической и гиперболической плоскостях

1. Определение 5. **Пусть A и B — две s-точки. Радианная мера одного из четырех углов, образованных прямыми (OA) и (OB), не превосходит π. Сферическим расстоянием (s-расстоянием) между A и B называется радианная мера этого угла.**

Это s-расстояние между s-точками A и B можно также определить, используя понятие длины кривой. Напомним, что длиной кривой

$\varphi : [a, b] \to \boldsymbol{R}^3$ называется число $l_\varphi = \int_a^b \sqrt{(x'(t))^2+(y'(t))^2+(z'(t))^2}\,dt$, где x', y', z' — производные функций $x(t)$, $y(t)$, $z(t)$ (см. приложение 1, пункт 2). Так как радианная мера центрального угла равна длине дуги единичной окружности, на которую он опирается, то получим, что s-расстояние между диаметрально противоположными s-точками равно π, а для других пар точек оно равно длине меньшей дуги, соединяющей эти точки в плоскости, проходящей через них и точку $\bar{O}$.

Обозначим $\rho_s(A, B)$ — s-расстояние между A и B. В этом обозначении индекс s иногда опускается.

Упражнение 7. *Докажите, что:*

а) $\rho_s(\cdot, \cdot)$ *является метрикой на множестве* S^2.

Указание. Используйте стереометрическую теорему: сумма радианных мер двух плоских углов любого трехгранного угла не меньше радианной меры третьего плоского угла;

б) $\rho_s(A, B) = 2\arcsin\frac{|A, B|}{2}$, *где* $|A, B|$ — *обычное евклидово расстояние между точками* A *и* B;

в) $|A, B| \leqslant \rho_s(A, B)$;

г) $\rho_s(A, B) \leqslant \frac{1}{\cos\frac{\rho_s(A, B)}{2}}|A, B|$.

Указание. Рассмотрите треугольник AOB и гомотетию с центром в точке $\bar{O}$ и коэффициентом $\frac{1}{\cos\frac{\rho_s(A, B)}{2}}$.

Метрика $\rho_s(\cdot, \cdot)$, как и всякая метрика, задает на S^2 топологию. Оказывается, что эта топология совпадает с топологией, задаваемой евклидовой метрикой $|\cdot, \cdot|$, т. е. совпадает с топологией, индуцируемой пространством $\boldsymbol{R}^3$.

Предложение 5. ***Топология на S^2, задаваемая сферическим расстоянием, совпадает с топологией, задаваемой евклидовым расстоянием.***

▷ Пусть $|A, B| \leqslant \sqrt{2}$. Тогда, по п. б) и г) упражнения 7

$$\rho_s(A, B) \leqslant \frac{1}{\sqrt{1-\left(\frac{|A, B|}{2}\right)^2}}|A, B| \leqslant \sqrt{2}|A, B| < 3|A, B|.$$

Если же $|A, B| > \sqrt{2}$, то

$$\rho_s(A, B) \leqslant \pi = \frac{\pi}{|A, B|}\cdot|A, B| < \frac{\pi}{\sqrt{2}}|A, B| < 3|A, B|.$$

Используя упражнение 7 в), получаем, что для любых s-точек A и B

$$|A, B| \leqslant \rho_s(A, B) < 3|A, B|. \qquad (*)$$

Рассмотрим евклидову окрестность s-точки A радиуса $\varepsilon>0$, т. е. $U(A,\ \varepsilon)\rightleftarrows\{B\in S^2\,|\,|A,\ B|<\varepsilon\}$. Из $(*)$ получаем, что $U(A,\ \varepsilon)$ содержит ρ — окрестность s-точки A радиуса ε, т. е. если

$$U_\rho(A,\ \varepsilon)=\{B\in S^2\,|\,\rho(A,\ B)<\varepsilon\},\ \text{то}\ U_\rho(A,\ \varepsilon)\subset U(A,\ \varepsilon).$$

Но тогда, если $U\subset S^2$ и U — открыто в топологии, задаваемой метрикой $|\cdot,\ \cdot|$, то всякая точка $A\in U$ входит в U с некоторой евклидовой окрестностью $U(A,\ \varepsilon)$ и поэтому входит в U с ρ-окрестностью $U_\rho(A,\ \varepsilon)$, т. е. U — открыто и в топологии, задаваемой метрикой $\rho(\cdot,\ \cdot)$.

Обратное утверждение аналогичным образом выводится из $(*)$ с использованием включения $U\left(A,\ \frac{\varepsilon}{3}\right)\subset U_\rho(A,\ \varepsilon)$.

Следующая теорема 1 показывает, что сферический отрезок между s-точками A и B является кратчайшим путем среди всех кусочно-гладких путей, соединяющих A и B. Напомним (§ 1, п. 2), что понятие s-отрезка не может быть определено на основе понятия «между», так как последнее отсутствует в s-планиметрии. Тем не менее, с метрической точки зрения s-отрезок удовлетворяет естественному требованию к отрезку, как пути, имеющему наименьшую длину.

Напомним: если $\tau:[0,\ 1]\to[a,\ b]$ — гладкая функция из $[0,\ 1]$ на $[a,\ b]$, для которой τ' везде отлична от нуля, а $\varphi:[a,\ b]\to\boldsymbol{R}^3$ — кривая в $\boldsymbol{R}^3$, то длина кривой φ равна длине кривой $\varphi\circ\tau$. Другими словами, длина гладкой кривой не зависит от способа параметризации этой кривой (см. приложение 1, пункт 2, предложения 1, 3).

Теорема 1. ***а) Длина сферического отрезка с концами в s-точках A и B равна $\rho_s(A,\ B)$.***

б) Пусть $\varphi:[0,\ 1]\to S^2$ — любая кусочно-гладкая кривая на S^2, у которой $\varphi(0)=A$ и $\varphi(1)=B$. Тогда длина l_φ кривой φ больше или равна $\rho_s(A,\ B)$.

▷ а) Это утверждение доказано в начале этого параграфа, так как сферический отрезок совпадает с дугой, на которую опирается соответствующий угол. Сейчас по существу проверим, что радианная мера угла равна длине дуги, на которую он опирается. Введем ортонормированную систему координат в $\boldsymbol{R}^3$ таким образом, чтобы точка A имела координаты $\langle 1,\ 0,\ 0\rangle$, а у точки B вторая координата была неотрицательна и третья координата равна нулю. Тогда параметрическое задание s-отрезка между A и B имеет вид:

$$x(t)=\cos t,\ y(t)=\sin t,\ z(t)=0,\ t\in[0,\ \alpha].$$

Здесь α — радианная мера того геометрического угла, образованного лучами $[OA)$ и $[OB)$, который не превосходит развернутого угла. Поэтому длина отрезка в этой системе координат равна

$$\int_0^\alpha\sqrt{\sin^2 t+\cos^2 t}\,dt=\alpha=\rho_s(A,\ B),$$ рисунок 117.

Для завершения доказательства остается напомнить, что длина кривой в $\boldsymbol{R}^3$ не зависит от того, в какой из ортонормированных систем координат вычисляется соответствующий интеграл.

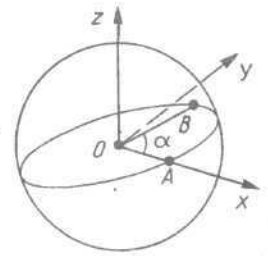


Рис. 117

б) Сначала рассмотрим случай, когда образ φ состоит из конечного числа s-отрезков. Доказательство проведем индукцией по числу n звеньев такой сферической ломаной (s-ломаной). Если $n=2$, то образ φ состоит из s-отрезков $[AC]_s$ и $[CB]_s$ и по п. а) этой теоремы, а также по неравенству треугольника для метрики ρ_s получим, что $l_\varphi=\rho(A,\ C)+\rho(C,\ B)\geqslant\rho(A,\ B)$.

Если s-ломаная Γ состоит из $(n+1)$-го звена, то, соединив s-отрезком s-точку B с началом n-го звена этой s-ломаной, получим новую s-ломаную Γ', состоящую уже из n звеньев. Длина Γ' по предположению индукции не меньше $\rho(A,\ B)$ и по неравенству треугольника для метрики ρ не больше длины Γ. Поэтому и длина Γ не меньше $\rho(A,\ B)$.

Пусть теперь $\varphi:[0,\ 1]\to S^2$ — любая кусочно-гладкая кривая, у которой $\varphi(0)=A$ и $\varphi(1)=B$. Рассмотрим конечное разбиение T отрезка $[0,\ 1]$ и соответствующее ему конечное разбиение T_φ образа φ. Пусть Γ — ломаная, состоящая из прямолинейных отрезков с вершинами в точках разбиения T_φ, а Γ_s s-ломаная с вершинами в этих же точках. Длина $l(\Gamma_s)$ s-ломаной Γ_s, как доказано выше, не меньше $\rho(A,\ B)$, а длина $l(\Gamma)$, как это доказывается в курсе анализа, не больше длины l_φ кривой φ, т. е. $\rho(A,\ B)\leqslant l(\Gamma_s)$ и $l(\Gamma)\leqslant l_\varphi$. Допустим от противного, что l_φ меньше $\rho(A,\ B)$, т. е. при некотором $\varepsilon>0$ выполняется $\rho(A,\ B)=l_\varphi+\varepsilon$.

Но тогда $l(\Gamma_s)-l(\Gamma)\geqslant\rho(A,\ B)-l_\varphi=\varepsilon$, т. е. при любом разбиении T длина s-ломаной Γ_s превышает длину ломаной Γ не меньше чем на некоторое постоянное число $\varepsilon>0$, что противоречит тому, что $l(\Gamma_s)-l(\Gamma)$ стремится к нулю при диаметре разбиения T, стремящемся к нулю. ■

2. Определение 6. **Пусть $A_e=\{A,\ -A\}$ и $B_e=\{B,\ -B\}$ — две e-точки. Рассмотрим два геометрических угла, образованных лучами $[O,\ A)$ и $[O,\ B)$ и лучами $[O,\ A)$ и $[O,\ (-B))$. Радианная мера одного из этих углов не превосходит $\pi/2$. Эллиптическим расстоянием (e-расстоянием) между A_e и B_e называется радианная мера этого угла (рис. 118).**

Отметим, что e-расстояние между A_e и B_e равно меньшей из длин тех четырех дуг, на которые точки A, $-A$, B, $-B$ разбивают окружность, проходящую через них. Если же все эти четыре дуги имеют одинаковую длину, то прямые (OA) и (OB) перпендикулярны и в этом случае e-расстояние равно $\pi/2$.

Для иллюстрации понятия e-расстояния рассмотрим e-плоскость как верхнюю полусферу S^2_+ с отождествленными диаметрально противоположными точками ее границы. Пусть две s-точки A и B не

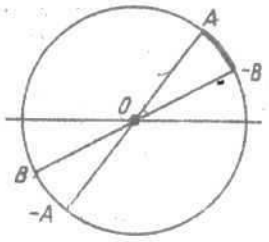
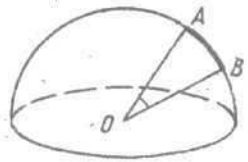


Рис. 118



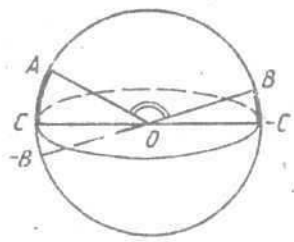


Рис. 119

лежат на границе S^2_+. Если меньший из углов $\angle AOB$ не больше прямого, то e-расстояние между A_e и B_e равно длине соответствующей дуги $\overset{\smile}{AB}$ экватора, проходящего через A и B. Если же угол $\angle AOB$ тупой, то e-расстояние между A_e и B_e равно сумме длин двух дуг $\overset{\smile}{AC}$ и $\overset{\smile}{B,(-C)}$, где C и $(-C)$ — точки пересечения границы полусферы S^2_+ и экватора, проходящего через A и B; причем углы $\angle AOC$ и $\angle BO,(-C)$ — острые, рисунок 119.

e-Расстояние между e-точками A_e и B_e обозначим $\rho_e(A_e, B_e)$. Индекс e иногда опускают. Отметим, что

$$\rho_e(\{A, -A\}, \{B, -B\}) = \min\{\rho_s(A, B), \pi - \rho_s(A, B)\}.$$

Теорема 2. ***а) Функция ρ_e является метрикой в e-плоскости.***
б) Отображение s-плоскости в e-плоскость вида

$$P : A \mapsto \{A, -A\} = A_e$$

является локальным гомеоморфизмом, т. е. для любой s-точки A найдутся такие окрестности $V(A)$ и $V(A_e)$ у соответственно точек A и A_e, что сужение отображения P является гомеоморфизмом $V(A)$ на $V(A_e)$.

▷ а) В доказательстве нуждается лишь неравенство треугольника для функции ρ_e. Пусть A_e, B_e, C_e три e-точки. Выберем s-точки A и B так, что $A_e = \{A, -A\}$, $B_e = \{B, -B\}$ и угол $\angle AOB$ не превосходит прямого, т. е. $\rho(A_e, B_e) = \rho_s(A, B)$. Введем в $\boldsymbol{R}^3$ систему координат таким образом, чтобы s-точка A была северным полюсом сферы S^2, а у s-точки B первая координата равнялась нулю и вторая и третья были неотрицательны (рис. 120).

Тогда можно считать, что s-точка C, где $C_e = \{C, -C\}$, лежит в верхней полусфере. Если C лежит в первом или втором октанте, то все три угла $\angle AOB$, $\angle AOC$, $\angle BOC$ не превосходят прямого, т. е. s-расстояния между A, B, C совпадут с e-расстояниями между A_e, B_e, C_e и для таких A_e, B_e, C_e неравенство треугольника для e-расстояния следует из неравенства треугольника для s-расстояния. Пусть C лежит в четвертом октанте и угол $\angle BOC$ — тупой. Тогда

$$\rho(B_e, C_e) = \rho_s(-B, C) < \rho_s(B, C) \leqslant \rho_s(A, B) + \rho_s(A, C) = \rho(A_e, B_e) + \rho(A_e, C_e).$$

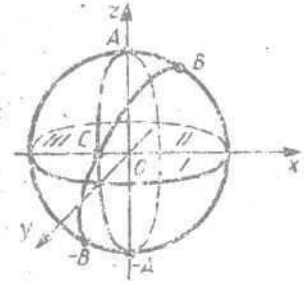


Рис. 120

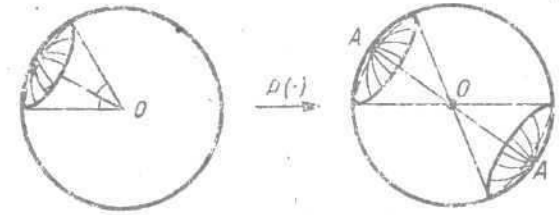


Рис. 121

Если $\angle BOC$ — не тупой, то

$$\rho(B_e, C_e) = \rho_s(B, C) \leqslant \rho_s(A, B) + \rho_s(A, C) = \rho(A_e, B_e) + \rho(A_e, C_e).$$

Случай, когда C лежит в третьем октанте, разбирается аналогично.

б) Пусть A — любая s-точка, $V(A)$ — ее открытая окрестность радиуса $\frac{\pi}{4}$ и $V(A_e)$ — открытая окрестность $A_e = \{A, -A\}$ радиуса $\frac{\pi}{4}$ (рис. 121).

Так как $V(A)$ не содержит диаметральных s-точек, то отображение P инъективно на $V(A)$. Так как для любых s-точек B и C из $V(A)$ угол $\angle BOC$ является острым, то $\rho_s(B, C) = \rho_e(P(B), P(C))$, т. е. отображение P сохраняет расстояния. Следовательно, $P(V(A)) = V(A_e)$, т. е. P — биекция $V(A)$ на $V(A_e)$. Кроме того, $P : V(A) \to V(A_e)$ и $P^{-1} : V(A_e) \to V(A)$ непрерывны как отображения, сохраняющие расстояние. ■

Замечание. В доказательстве п. б) теоремы 2 точка A не обязательно должна быть центром окрестности $V(A)$.

Предложение 6. ***Пусть A есть s-точка и φ — непрерывное отображение отрезка $[0, 1]$ в e-плоскость E^2 такое, что $\varphi(0) = A_e$, где $A_e = \{A, -A\}$. Тогда найдется единственное непрерывное отображение $\tilde{\varphi}$ отрезка $[0, 1]$ в s-плоскость S^2 такое, что $\tilde{\varphi}(0) = A$ и $P \circ \tilde{\varphi}(t) = \varphi(t)$ для всех $t \in [0, 1]$.***

▷ Для каждой e-точки $\varphi(t)$ выберем открытую ρ_e-окрестность радиуса $\frac{\pi}{4}$. Так как $\varphi[0, 1]$ — компактно (непрерывный образ компакта — компакт), то из такого открытого покрытия $\varphi[0, 1]$ можно выбрать конечное подпокрытие $V_1, \ldots, V_n$. При этом можно считать, что V_1 совпадает с открытой ρ_e-окрестностью e-точки A_e радиуса $\frac{\pi}{4}$. По теореме 2 б) ограничение отображения P на P_s-окрестность $V(A)$ s-точки A радиуса $\frac{\pi}{4}$ является гомеоморфизмом.

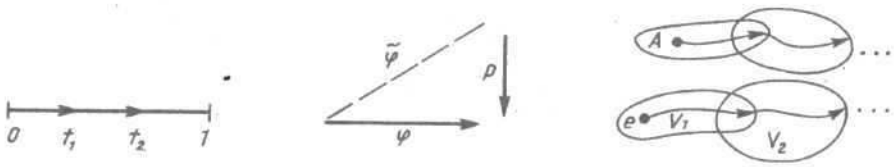


Рис. 122

Пусть

$$t_1 = \sup\{t \geqslant 0 \mid \forall \tau \leqslant t,\ \varphi(\tau) \in V_1\}.$$

Тогда $t_1 > 0$ и $\varphi([0,\ t_1)) \subset V_1$. Определим непрерывное отображение $\tilde{\varphi}$ полуинтервала $[0,\ t_1)$ в s-плоскость равенством $\tilde{\varphi}(t) = P^{-1} \circ \varphi(t)$, где P^{-1} — обратное отображение к гомеоморфизму $P \upharpoonright V(A)$. В результате мы «подняли» отображение $\varphi \upharpoonright [0,\ t_1)$ до отображения $\tilde{\varphi}$ так, что $\tilde{\varphi}(0) = A$ и $P \circ \tilde{\varphi}(t) = \varphi(t)$ (рис. 122).

Возьмем e-точку на кривой $\varphi(t)$, лежащую в $V_1 \cap V_2$, и, используя п. б) теоремы 2, аналогичным образом «поднимем» в s-плоскость окрестность V_2, а вместе с ней и участок кривой $\varphi([0,\ 1))$, где $t_2 > t_1$. Через n таких «поднятий» мы получим искомое отображение $\tilde{\varphi}$. ■

Замечание. Нетрудно видеть, что если поднимать кривую φ в кривую $\tilde{\varphi}$ с началом в s-точке $-A$, то это поднятие будет отличаться от первого лишь знаком.

Отображение $\tilde{\varphi}$ называется поднятием отображения φ.

Определение 7. а) **Непрерывная кривая φ в e-плоскости E^2 называется кусочно-гладкой, если кусочно-гладко ее поднятие $\tilde{\varphi}$ в s-плоскости S^2.**

б) Длиной кусочно-гладкой кривой φ в e-плоскости E^2 называется длина поднятия $\tilde{\varphi}$ в s-плоскости S^2.

Ясно, что определение 7 не зависит от того, в какую из точек A и $-A$ поднимается начало $A_e = [A,\ -A)$ кривой φ. Используя указанную связь между s-кривыми и e-кривыми, нетрудно показать, что e-отрезок имеет наименьшую длину из длин всевозможных кусочно-гладких e-кривых, соединяющих концы отрезка. При этом основным моментом является тот факт, что поднятие e-отрезка является s-отрезком. После этого доказательство по существу сводится к аналогичному факту для s-отрезков (теорема Ч).

3. Определение 8. а) **Гиперболическим скалярным произведением в R^3 называется следующая функция**

$$(\langle x_1,\ y_1,\ z_1\rangle,\ \langle x_2,\ y_2,\ z_2\rangle)_h = -x_1x_2 + y_1y_2 + z_1z_2.$$

б) Гиперболическим расстоянием (h-расстоянием) между точками A и B называется квадратный корень из числа $(A-B,\ A-B)_h$.

в) Пусть $\varphi : [a,\ b] \to R^3$ — гладкая кривая. Ее гиперболической длиной (h-длиной) называется

$$\int_a^b \sqrt{-\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\,dt, \text{ где}$$

$$\varphi(t) = \langle x(t),\ y(t),\ z(t)\rangle.$$

Замечание. В пунктах б) и в) этого определения под квадратным корнем $\sqrt{a}$ при $a \geqslant 0$ понимается арифметический корень, а при $a < 0$ понимается чисто мнимое число $i\sqrt{-a}$, где $\sqrt{-a}$ арифметический корень.

Упражнение 8. *а) Докажите, что h-скалярное произведение симметрично, билинейно, но не положительно определено.*

б) Найдите множество точек в R^3, находящихся на нулевом h-расстоянии от точки $\bar{O}$.

Арифметическое пространство R^3, рассматриваемое вместе с гиперболическим скалярным произведением, называют псевдоевклидовым пространством и обозначают R_1^3; при этом h-расстояния между точками R_1^3 могут быть положительными, нулевыми и чисто мнимыми числами, а h-длины кривых в R^3 могут быть комплексными числами.

Замечательно, что гладкие кривые, целиком лежащие в h-плоскости H^2, имеют вещественную положительную h-длину или, как говорят, метрика на H^2, индуцированная из псевдоевклидова пространства R_1^3, является евклидовой.

Для доказательства этого факта достаточно проверить, что h-длина любого касательного к гиперболической сфере H^2 вектора положительна.

Рассмотрим параметрическое задание h-плоскости

$$\begin{cases} x = \operatorname{ch} t, \\ y = \operatorname{sh} t \cos\varphi, \\ z = \operatorname{sh} t \sin\varphi, \end{cases} \quad 0 \leqslant t < \infty,\ 0 \leqslant \varphi < 2\pi,$$

см. § 1, п. 3. Фиксируем число $t \neq 0$ и рассмотрим векторы

$$e_1 = \left\langle \frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t} \right\rangle, \quad e_2 = \left\langle \frac{\partial x}{\partial \varphi}, \frac{\partial y}{\partial \varphi}, \frac{\partial z}{\partial \varphi} \right\rangle,$$

образующие базис в касательной плоскости к H^2 в точке $\langle x,\ y,\ z\rangle$. Тогда

$$e_1 = \langle \operatorname{sh} t,\ \operatorname{ch} t \cos\varphi,\ \operatorname{ch} t \sin\varphi\rangle,$$
$$e_2 = \langle 0,\ -\operatorname{sh} t \sin\varphi,\ \operatorname{sh} t \cos\varphi\rangle.$$

Ясно, что $(e_1,\ e_2)_h = 0$ и для всех $\alpha,\ \beta \in R$ таких, что $\alpha^2 + \beta^2 \neq 0$, получаем

$$(\alpha e_1 + \beta e_2,\ \alpha e_1 + \beta e_2)_h = \alpha^2 (e_1 e_1)_h^2 + \beta^2 (e_2 e_2)_h^2 =$$
$$= \alpha^2(-\operatorname{sh}^2 t + \operatorname{ch}^2 t(\cos^2\varphi + \sin^2\varphi)) + \beta^2 \operatorname{sh}^2 t(\sin^2\varphi + \cos^2\varphi) =$$
$$= \alpha^2 + \beta^2 \cdot \operatorname{sh}^2 t > 0.$$

При $t = 0$ касательная плоскость в точке $\langle x,\ y,\ z\rangle = \langle 1,\ 0,\ 0\rangle$ задается уравнением $x = 1$ и наше утверждение легко проверяется непосредственным вычислением.

Пусть $A = (a_{ij})$, где $1 \leqslant i,\ j \leqslant 3$, такая матрица, что

$$(*)\quad \begin{cases} -a_{11}^2+a_{21}^2+a_{31}^2=-1 \\ -a_{12}^2+a_{22}^2+a_{32}^2=-a_{13}^2+a_{23}^2+a_{33}^2=1 \\ -a_{11}a_{12}+a_{21}a_{22}+a_{31}a_{32}=-a_{11}a_{13}+a_{21}a_{23}+a_{31}a_{33}= \\ =-a_{12}a_{13}+a_{22}a_{23}+a_{32}a_{33}=0. \end{cases}$$

Обычный базис $e_1=\langle 1, 0, 0\rangle$, $e_2=\langle 0, 1, 0\rangle$, $e_3=\langle 0, 0, 1\rangle$ при линейном отображении $\boldsymbol{R}^3$ на себя, задаваемом матрицей A, перейдет в базис e_1', e_2', e_3'. В последнем базисе кривая φ имеет параметрическое задание:

$$\begin{cases} x'(t)=a_{11}x(t)+a_{12}y(t)+a_{13}z(t) \\ y'(t)=a_{21}x(t)+a_{22}y(t)+a_{23}z(t) \\ z'(t)=a_{31}x(t)+a_{32}y(t)+a_{33}z(t), \end{cases}$$

где $\varphi(t)=\langle x(t), y(t), z(t)\rangle$ — параметрическое задание φ в базисе e_1, e_2, e_3. Тогда прямое вычисление с учетом $(*)$ показывает, что

$$-\left(\frac{dx'}{dt}\right)^2+\left(\frac{dy'}{dt}\right)^2+\left(\frac{dz'}{dt}\right)^2=-\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2.$$

Поэтому длина кривой φ при переходе от координат x, y, z к координатам x', y', z' не изменится, если этот переход задается матрицей A с условием $(*)$. Нетрудно также видеть, что линейное отображение, задаваемое матрицей A, сохраняет гиперболическое скалярное произведение (проверьте). Линейные отображения, задаваемые матрицами с условием $(*)$, называются движениями псевдоевклидова пространства $\boldsymbol{R}_1^3$.

Для дальнейшего важно уметь задавать h-прямые параметрически. Поступим следующим образом. Сначала параметрически зададим l_0 — h-прямую $H^2\cap Oxy$. Учитывая параметрическое задание всей h-плоскости H^2 (см. § 1), получаем параметрическое задание h-прямой l_0

$$\begin{cases} x=\operatorname{ch} t \\ y=\operatorname{sh} t \\ z=0, \ t\in\boldsymbol{R}. \end{cases}$$

Для любой h-прямой $l=H^2\cap\pi$ найдем движение T псевдоевклидова пространства $\boldsymbol{R}_1^3$, при котором h-прямая l_0 перейдет в h-прямую l. Композиция параметрического задания h-прямой l_0 с движением T даст тогда параметрическое задание h-прямой l.

Предложение 7. ***Для любой плоскости π, проходящей через начало координат и пересекающей H^2, найдется линейное движение T псевдоевклидова пространства $\boldsymbol{R}_1^3$, которое отображает плоскость Oxy на плоскость π. При этом в базисе $e_1'=Te_1$, $e_2'=Te_2$, $e_3'=Te_3$ — h-прямой $H^2\cap\pi$ имеет вид — $(x')^2+(y')^2=-1$, $x'\geqslant 1$ и поэтому задается параметрически***

$$\begin{cases} x'=\operatorname{ch} t \\ y'=\operatorname{sh} t \\ z'=0, \ t\in\boldsymbol{R}. \end{cases}$$

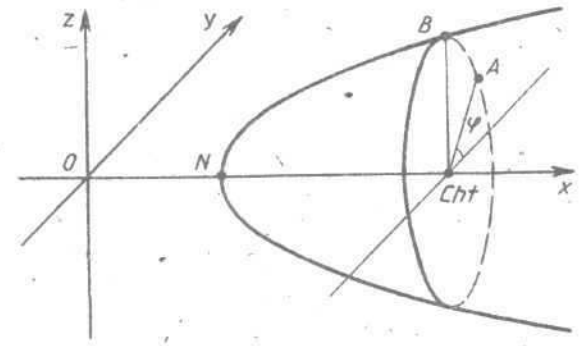


Рис. 123

▷ Пусть $A=\langle \operatorname{ch} t, \ \cos\varphi \operatorname{sh} t, \ \sin\varphi \operatorname{sh} t\rangle$ — вершина h-прямой $H^2\cap\pi$, т. е. вершина ветви гиперболы в плоскости π. Рассмотрим h-точку $B=\langle \operatorname{ch} t, 0, \operatorname{sh} t\rangle$, которая получается из точки A в результате поворота на угол $\alpha=\left(\frac{\pi}{2}-\varphi\right)$ вокруг оси Ox, рисунок 123. Такой поворот задается матрицей:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

и поэтому является движением пространства $\boldsymbol{R}_1^3$. Обозначим этот поворот R_α. При повороте R_α h-плоскость H^2 перейдет на себя, плоскость π перейдет в плоскость π', проходящую через прямые (OB) и Oy, а h-прямая $H^2\cap\pi$ перейдет в h-прямую $H^2\cap\pi'$ с вершиной в h-точке B.

Рассмотрим движение T_1 пространства $\boldsymbol{R}_1^3$, задаваемое матрицей

$$\begin{pmatrix} \operatorname{ch} t & 0 & \operatorname{sh} t \\ 0 & 1 & 0 \\ \operatorname{sh} t & 0 & \operatorname{ch} t \end{pmatrix}$$

При этом движении h-точка $N=\langle 1, 0, 0\rangle$ перейдет в h-точку $B=\langle \operatorname{ch} t, 0, \operatorname{sh} t\rangle$, а ось Oy останется на месте. Поэтому движение T_1 переводит плоскость Oxy на плоскость π', положим $T=R_\alpha^{-1}\circ T_1$. Тогда T есть движение, как композиция движений, и T переводит плоскость Oxy на плоскость π.

Так как T сохраняет гиперболическое скалярное произведение $(\cdot, \cdot)_h$, то T переводит h-плоскость H^2 на себя. Следовательно, T переводит h-прямую $H^2\cap Oxy$ на h-прямую $H^2\cap\pi$. Так как z — координаты для точек h-прямой $H^2\cap Oxy$ равны нулю, то z'-координаты (третьи координаты в базисе e_1', e_2', e_3') для точек h-прямой $H^2\cap\pi$ также равны нулю. В плоскости π выполняется $(e_1', e_1')_h=-1$, $(e_2', e_2')_h=1$, $(e_1', e_2')_h=0$ и h-прямая $H^2\cap\pi$ в этой плоскости имеет вид $\{C\in\pi|(C, C)_h=-1\}$. Если $C=xe_1'+y'e_2'$,

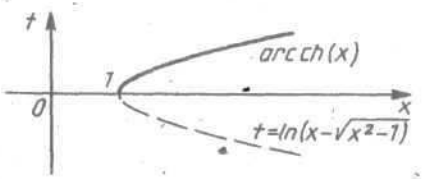


Рис. 124

то получаем, что $-1=(C,\ C)_h=-(x')^2+(y')^2$. Так как при этом $e_1'\in H^2\cap\pi$, то $x'\geqslant 1$ для всех точек $C\in H^2\cap\pi$. ■

Определение 9. **Пусть A и B — две h-точки и $[A,\ B]_h$ есть h-отрезок с концами в A и B, т. е. отрезок ветви некоторой гиперболы. Гиперболическим расстоянием между h-точками A и B называется h-длина кривой $[A,\ B]_h$. Гиперболическое расстояние обозначим $\rho_h(\cdot,\ \cdot)$.**

Теорема 3. ***а) Для любых h-точек A_1 и A_2 выполняется***

$$\rho(A_1,\ A_2)=\operatorname{arcch}\,(-(A_1,\ A_2)_h).$$

б) Функция ρ_h является метрикой в h-плоскости H^2.

в) h-Отрезок с концами в h-точках A_1 и A_2 имеет длину, наименьшую из h-длин всех кусочно-гладких h-кривых с концами в A_1 и A_2.

▷ а) Напомним, что $\operatorname{arcch}(\cdot)$ — гиперболический арккосинус — является положительной ветвью отображения, обратного к $\operatorname{ch}(\cdot)$. Более точно, решая уравнение $x=\frac{e^t+e^{-t}}{2}$ относительно $t\geqslant 0$, получим $t=\ln(x\pm\sqrt{x^2-1})$, т. е. $\operatorname{arcch} x=\ln(x+\sqrt{x^2-1})$, $x\geqslant 1$ (рис. 124).

Пусть $x=\operatorname{ch}$ и $y=\operatorname{sh} t$. Если $y\geqslant 0$, то и $t\geqslant 0$, и поэтому

$$t=\operatorname{arcch} x=\ln(x+\sqrt{x^2-1})=\ln(x+y).$$

Если $y<0$, то и $t<0$, и поэтому

$$t=-\operatorname{arcch} x=\ln(x-\sqrt{x^2-1})=\ln(x+y).$$

Докажем теперь требуемое равенство для h-точек A_1 и A_2, лежащих на h-прямой $(H^2\cap Oxy)$. Зададим эту h-прямую параметрически $\begin{cases}x=\operatorname{ch} t,\\ y=\operatorname{sh} t,\ t\in\boldsymbol{R},\\ z=0.\end{cases}$

Согласно определению 9,

$$\rho_h(A_1,\ A_2)_h=\rho_h(\langle x_1,\ y_1,\ 0\rangle,\ \langle x_2,\ y_2,\ 0\rangle)=$$

$$=\int\limits_{t_1}^{t_2}\sqrt{-\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}\,dt=$$

$$=\int\limits_{t_1}^{t_2}\sqrt{-\operatorname{sh}^2 t+\operatorname{ch}^2 t}\,dt=\int\limits_{t_1}^{t_2}dt=t_2-t_1.$$

Здесь мы предположили, что $t_1\leqslant t_2$. В общем случае h-расстояние равно $|t_2-t_1|$.

Пусть $x_1\cdot y_2-x_2\cdot y_1\geqslant 0$. Тогда

$$\rho_h(A_1,\ A_2)=|t_2-t_1|=\left|\ln\frac{x_2+y_2}{x_1+y_1}\right|=|\ln(x_2+y_2)(x_1-y_1)|=$$
$$=|\ln((x_1x_2-y_1y_2)+(x_1y_2-y_1x_2))|.$$

Так как $x_1y_2-x_2y_1\geqslant 0$, то для доказательства равенства $\rho_h(A_1,\ A_2)=\operatorname{arcch}(x_1x_2-y_1y_2)$ достаточно проверить, что

$$(x_1y_2-x_2y_1)^2=(x_1x_2-y_1y_2)^2-1$$

или, что $x_1^2(x_2^2-1)+x_2^2(x_1^2-1)=x_1^2\cdot x_2^2+(x_1^2-1)(x_2^2-1)-1$.

Но последнее равенство, очевидно, верно.

Случай $x_1y_2-x_2y_1<0$ разбирается аналогично.

Пусть теперь A_1 и A_2 — произвольные h-точки. Если $A_1=A_2$, то $(A_1,\ A_2)_h=-1$ и равенство $\rho_h(A_1,\ A_2)=\operatorname{arcch}(-(A_1,\ A_2)_h)$, очевидно, верно. Если $A_1\neq A_2$, то проведем через точки O, A_1, A_2 плоскость π и согласно предложению 7 выберем такой базис e_1', e_2', e_3', в котором h-прямая $H^2\cap\pi$, проходящая через точки A_1 и A_2, имеет вид $-(x')^2+(y')^2=1$, $x'\geqslant 1$. Тогда в координатах x', y', z' можно провести аналогичные разобранному выше случаю вычисления и получить, что в этих координатах выполняется требуемое равенство. Так как переход от координат x, y, z к координатам x', y', z' осуществляется с помощью движения пространства $\boldsymbol{R}_1^3$ (см. предложение 7), т. е. не меняется $\rho_h(\cdot,\ \cdot)$ и $(\cdot,\ \cdot)_h$, то требуемое равенство верно и в первоначальных координатах.

б) В доказательстве нуждается лишь неравенство треугольника, т. е. неравенство

$$\operatorname{arcch}(-(A_1,\ A_2)_h)\leqslant\operatorname{arcch}(-(A_1,\ A_3)_h)+\operatorname{arcch}(-(A_2,\ A_3)_h).$$

Так как обе части этого неравенства неотрицательны, то, в силу монотонности функции $\operatorname{ch}(\cdot)$ на $[0;\ +\infty)$, это неравенство эквивалентно неравенству

$$-(A_1,\ A_2)\leqslant\operatorname{ch}(\operatorname{arcch}(-(A_1,\ A_3))+$$
$$+\operatorname{arcch}(-(A_2,\ A_3))=(A_1,\ A_3)(A_2,\ A_3)+$$
$$+\sqrt{(A_1,\ A_3)^2-1}\,\sqrt{(A_2,\ A_3)^2-1}\,.$$

Или

$$-(A_1,\ A_2)-(A_1,\ A_3)(A_2,\ A_3)\leqslant\sqrt{(A_1,\ A_3)^2-1}\,\sqrt{(A_2,\ A_3)^2-1}\,.$$

Если левая часть отрицательна, то неравенство, очевидно, верно. Если левая часть неотрицательна, то неравенство при возведении в квадрат перейдет в эквивалентное неравенство

$$(A_1,\ A_2)^2+(A_1,\ A_3)^2(A_2,\ A_3)^2+2(A_1,\ A_2)(A_1,\ A_3)(A_2,\ A_3)\leqslant$$
$$\leqslant 1-(A_1,\ A_3)^2-(A_2,\ A_3)^2+(A_1,\ A_3)^2(A_2,\ A_3)^2$$

или же

$$-1+2(A_1,\ A_2)(A_1,\ A_3)(A_2,\ A_3)+(A_1,\ A_2)^2+(A_1,\ A_3)^2+(A_2,\ A_3)^2\leqslant 0.$$

Оказывается, что в левой части последнего неравенства стоит определитель матрицы $A=(a_{ij})$, где $a_{ij}=(A_i,\ A_j)$. Это проверяется прямым подсчетом с использованием того, что $a_{11}=a_{22}=a_{33}=-1$.

Пусть векторы $\overline{O, A_1}$, $\overline{O, A_2}$, $\overline{O, A_3}$ линейно независимы и C — матрица линейного отображения, переводящего эти векторы в $e_1=\langle 1, 0, 0\rangle$, $e_2=\langle 0, 1, 0\rangle$, $e_3=\langle 0, 0, 1\rangle$. Как известно из курса алгебры, матрица E, составленная из h-скалярных произведений (e_i, e_j), равна C^tAC, где C^t — матрица, транспонированная к матрице C. Поэтому $\det E=\det(C^tAC)=(\det C)^2\det A$.

Но так как $E=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}$, то $\det E=-1<0$, и, следовательно, $\det A<0$.

Если один из векторов $\overline{O, A_i}$ есть линейная комбинация двух других, то одна из строк матрицы A, в силу билинейности формы $(\cdot, \cdot)$, есть линейная комбинация двух других, и поэтому $\det A=0$.

в) Доказательство в точности повторяет схему доказательства теоремы 1 б). ■

4. Метрику в плоских моделях неевклидовых плоскостей S^2, E^2 и H^2 определим следующим образом. Для любых двух точек A и B из какой-либо плоской модели сначала отобразим эти точки в точки A' и B' соответствующей неевклидовой плоскости с помощью биекции, обратной к той биекции, по которой строилась эта плоская модель (см. § 1, п. 4). Затем найдем расстояние между точками A' и B' в соответствующей неевклидовой плоскости (см. пп. 1—3 этого параграфа). По определению положим расстояние между точками A и B в плоской модели равным расстоянию между точками A' и B' в соответствующей неевклидовой плоскости. Хотелось бы, однако, уметь вычислять расстояния в плоских моделях неевклидовых плоскостей в терминах самих плоских моделей, т. е. без перехода к расстояниям в S^2, E^2 или H^2.

Рассмотрим плоскую модель S^2P сферической плоскости S^2.

Теорема 4. ***В плоской модели S^2P сферической плоскости S^2 для любых конечных s-точек A и B***

$$\rho_s(A, B)=2\arcsin\frac{|A, B|}{\sqrt{1+|O, A|^2}\sqrt{1+|O, B|^2}}$$

и

$$\rho_s(\infty, B)=2\arcsin\frac{1}{\sqrt{1+|O, B|^2}}$$

где $|A, B|$ — евклидово расстояние между точками A и B плоскости Oxy.

▷ Пусть $B=\langle X, Y, 0\rangle$. Тогда $P^{-1}(B)=\left\langle\frac{2X}{1+|O, B|^2}, \frac{2Y}{1+|O, B|^2}, \frac{1-|O, B|^2}{1+|O, B|^2}\right\rangle$ (см. § 1). Найдем евклидово расстояние $|E, P^{-1}(B)|$. Так как $E=\langle 0, 0, -1\rangle$ и $|O, B|^2=X^2+Y^2$, то $|E, P^{-1}(B)|^2=$

$$=\frac{4X^2}{(1+|O, B|^2)^2}+\frac{4Y^2}{(1+|O, B|^2)^2}+\left(\frac{1-|O, B|^2}{1+|O, B|^2}+1\right)^2=\frac{4}{1+|O, B|^2}.$$

Учитывая связь (см. упражнение 7) между s-расстоянием $\rho_s(E, P^{-1}(B))$ и евклидовым расстоянием $|E, P^{-1}(B)|$, получаем

$$\rho_s(\infty, B)=\rho_s(E, P^{-1}(B))=2\arcsin\frac{|E, P^{-1}(B)|}{2}=2\arcsin\frac{1}{\sqrt{1+|O, B|^2}}.$$

Пусть теперь A и B — конечные s-точки в S^2P и $A'\neq P^{-1}(A)$, $B'=P^{-1}(B)$. Тогда $|A, E|^2=1+|O, A|^2$, $|B, E|^2=1+|O, B|^2$ и $\operatorname{tg}\angle OEA=|O, A|$, $\operatorname{tg}\angle OEB=|O, B|$.

Так как $|O, E|=|O, A'|=1$, то $|A', E|=2\cos\angle OEA=$

$$=\frac{2}{\sqrt{1+\operatorname{tg}^2\angle OEA}}=\frac{2}{\sqrt{1+|O, A|^2}}$$ и аналогичным образом $|B', E|=$

$$=\frac{2}{\sqrt{1+|O, B|^2}}.$$

Следовательно,

$$|A', B'|^2=|A', E|^2+|B', E|^2-2|A', E|\,|B', E|\cdot\cos\angle AEB=$$

$$=\frac{4}{1+|O, A|^2}+\frac{4}{1+|O, B|^2}-\frac{8}{\sqrt{1+|O, A|^2}\sqrt{1+|O, B|^2}}\times$$

$$\times\frac{(1+|O, A|^2)+(1+|O, B|^2)-|A, B|^2}{2\sqrt{1+|O, A|^2}\sqrt{1+|O, B|^2}}=\frac{4|A, B|^2}{(1+|O, A|^2)(1+|O, B|^2)}.$$

Для получения конечного результата остается воспользоваться формулой: $\rho(A', B')=2\arcsin\frac{|A', B'|}{2}$. ■

Упражнение 9. *Используя результат теоремы 4, проведите аналогичные вычисления для e-расстояния в плоской модели E^2P эллиптической плоскости E^2.*

Выясним, каким образом вычисляется расстояние между двумя точками в модели H^2P_1. С этой целью удобно плоскость Oyz, в которой лежит открытый единичный круг D, точки которого являются точками H^2P_1, рассматривать как комплексную плоскость C.

Напомним, что сложным (двойным) отношением упорядоченной четверки комплексных чисел z_1, z_2, z_3, z_4 называется комплексное число

$$((z_1, z_2)(z_3, z_4))\rightleftarrows\frac{z_1-z_3}{z_1-z_4}:\frac{z_2-z_3}{z_2-z_4}.$$

Сложное отношение можно определить и в расширенной комплексной плоскости $\boldsymbol{C}\cup\{\infty\}$. Например, если $z_2=\infty$, то следует рассмотреть $\lim\limits_{z_2\to\infty}\frac{z_1-z_3}{z_1-z_4}:\frac{z_2-z_3}{z_2-z_4}$, который равен $\frac{z_1-z_3}{z_1-z_4}$ и который по определению называется сложным отношением четверки $((z_1, \infty)(z_3, z_4))$.

Теорема 5. ***Пусть z_1 и z_2 — две точки из H^2P_1 и (z_1, z_2) — прямая в H^2P_1, проходящая через z_1 и z_2, т. е. дуга окружности, проходящей через z_1 и z_2 перпендикулярно к границе круга D. Обозначим u и v — точки пересечения этой окружности с границей***

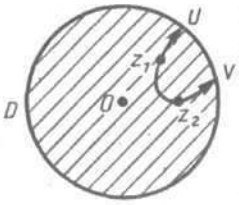


Рис. 125

круга D (рис. 125). Тогда расстояние между z_1 и z_2 в модели H^2P_1 равно

$$|\ln((z_1, z_2)(u, v))|.$$

▷ Как и при доказательстве теоремы 3, рассмотрим сначала частный случай. А именно, пусть $z_1 = P(A_1)$, $z_2 = P(A_2)$, где h-точки $A_1 = \langle \operatorname{ch} t_1, \operatorname{sh} t_1, 0\rangle$, $A_2 = \langle \operatorname{ch} t_2, \operatorname{sh} t_2, 0\rangle$ лежат на h-прямой $H^2 \cap Oxy$ и $P : H^2 \to \mathring{D}$ — стереографическая проекция из южного полюса $F = \langle -1, 0, 0\rangle$, см. § 1, п. 4. Тогда комплексные числа z_1 и z_2 являются вещественными и из подобия треугольников следует, что

$$z_1 = \frac{\operatorname{sh} t_1}{1+\operatorname{ch} t_1},\quad z_2 = \frac{\operatorname{sh} t_2}{1+\operatorname{ch} t_2}.$$

В рассматриваемом случае $u = 1$, $v = -1$. Поэтому

$$\frac{z_1-u}{z_1-v} = \frac{z_1-1}{z_1+1} = \frac{\operatorname{sh} t_1 - 1 - \operatorname{ch} t_1}{\operatorname{sh} t_1 + 1 + \operatorname{ch} t_1} = \frac{-e^{-t_1}-1}{e^{t_1}+1} = -\frac{1+\frac{1}{e^{t_1}}}{e^{t_1}+1} = -e^{-t_1}.$$

Аналогично $\frac{z_2-u}{z_2-v} = -e^{-t_2}$, и поэтому $|\ln((z_1, z_2), (u, v))| = |\ln\frac{-e^{-t_1}}{-e^{-t_2}}| = |t_2 - t_1|$, что по вычислениям из теоремы 3 равно $\rho_h(A_1, A_2)$.

Пусть z_1 и z_2 — произвольные различные точки из H^2P_1 и $z_1 = P(A_1)$, $z_2 = P(A_2)$. Через точки O, A_1, A_2 проведем плоскость π и пусть T — движение пространства $\boldsymbol{R}^3_1$, переводящее плоскость Oxy на плоскость π (см. предложение 7). Тогда h-точки $A_1' \rightleftarrows T^{-1}(A_1)$ и $A_2' \rightleftarrows T^{-1}(A_2)$ лежат на h-прямой $H^2 \cap Oxy$ и $\rho_h(A_1', A_2') = \rho_h(A_1, A_2)$.

Так как для точек $z_1' = P(A_1')$, $z_2' = P(A_2')$ равенство $|\ln((z_1', z_2'), (-1, 1))| = \rho_h(A_1', A_2')$ уже доказано, что достаточно проверить равенство двух сложных отношений $((z_1, z_2), (u, v))$ и $((z_1', z_2'), (u', v'))$, где $z_1 = PTP^{-1}(z_1')$, $z_2 = PTP^{-1}(z_2')$, $u = PTP^{-1}(u')$, $v = PTP^{-1}(v')$, $u' = 1$, $v' = -1$.

По предложению 7 $T = R_\alpha^{-1}T_1$. Поэтому $PTP^{-1} = (PR_\alpha^{-1}P^{-1}) \times (PT_1P^{-1})$ и достаточно проверить, что отображения

$$PT_1P^{-1}, \text{ где } T_1 = \begin{pmatrix} \operatorname{ch} t & 0 & \operatorname{sh} t \\ 0 & 1 & 0 \\ \operatorname{sh} t & 0 & \operatorname{ch} t \end{pmatrix} \quad \text{и } PR_\alpha^{-1}P^{-1}, \text{ где } R_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

открытого круга $\mathring{D}$ в себя сохраняют сложное отношение.

Отображение $PR_\alpha P^{-1}$ является поворотом круга $\mathring{D}$ вокруг начала координат на угол α, т. е. отображением вида $z \mapsto e^{i\alpha}z$, которое, очевидно, сохраняет сложное отношение.

Из курса теории аналитических функций известно, что при дробно-линейных отображениях, т. е. при отображениях вида

$$z \mapsto \frac{az+b}{cz+d},\ a, b, c, d \in \boldsymbol{C},\ ad - bc \neq 0,$$

сложное отношение сохраняется. Мы не будем полностью доказывать, что отображение PT_1P^{-1} является дробно-линейным. Укажем лишь, как можно найти формулу, задающую это отображение.

Пусть $B = \langle \operatorname{ch} t, 0, \operatorname{sh} t\rangle$. Тогда

$$P(B) = i\frac{\operatorname{sh} t}{1+\operatorname{ch} t} \rightleftarrows ia.$$

При стереографической проекции P h-прямая $H^2 \cap Oxy$ перейдет в диаметр $]-1, 1[$ круга D, а h-прямая с вершиной в h-точке B перейдет в дугу окружности, проходящую через точку ia и пересекающую под прямым углом границу круга D. Нетрудно найти радиус R этой окружности. Действительно, $1 + R^2 = (R+a)^2$, поэтому

$$R = \frac{1-a^2}{2a} = \frac{1-\frac{\operatorname{sh}^2 t}{(1+\operatorname{ch} t)^2}}{\frac{2\operatorname{sh} t}{1+\operatorname{ch} t}} = \frac{1+\operatorname{ch}^2 t + 2\operatorname{ch} t - \operatorname{sh}^2 t}{2(1+\operatorname{ch} t)\cdot\operatorname{sh} t} = \frac{2(1+\operatorname{ch} t)}{2(1+\operatorname{ch} t)\operatorname{sh} t} = \frac{1}{\operatorname{sh} t}.$$

Следовательно, при отображении PT_1P^{-1} точка O перейдет в точку $i\frac{\operatorname{sh} t}{1+\operatorname{ch} t}$, точка 1 перейдет в точку $e^{i\left(\frac{\pi}{2} - \operatorname{arctg}\frac{1}{\operatorname{sh} t}\right)}$, а точка -1 перейдет в точку $e^{i\left(\frac{\pi}{2} - \operatorname{arctg}\frac{1}{\operatorname{sh} t}\right)}$. По трем парам соответствующих точек дробно-линейное отображение однозначно определяется. ■

Упражнение 10. *Докажите, что:*

а) $((z_1, z_2)(z_3, z_4)) = ((z_1, z_2)(z_4, z_3))^{-1} = ((z_2, z_1)(z_3, z_4))^{-1} = ((z_3, z_4)(z_1, z_2))$;

б) в модели H^2P_1 выполняются аксиомы непрерывности (V группа);

в) расстояние между точками 0 *и* $\frac{1}{2}$ *равно* $\ln 3$, *а между* 0 *и* 0,99 *равно* $\ln 199$.

Перейдем ко второй модели Пуанкаре H^2P_2 h-планиметрии. Переход от первой модели H^2P_1 Пуанкаре ко второй модели H^2P_2 осуществляется при помощи дробно-линейного отображения вида $z \mapsto \frac{z-1}{iz+i}$, а при дробно-линейных отображениях сложное отношение сохраняется. Поэтому расстояние между точками ω_1 и ω_2 в H^2P_2 определяется по формуле

$$\rho_h(\omega_1, \omega_2) = |\ln((\omega_1, \omega_2), (u, v))|,$$

где u и v — точки пересечения вещественной оси Ox окружностью, проходящей через точки ω_1 и ω_2 верхней полуплоскости и перпендикулярно вещественной оси. Напомним, что $Re\omega$ и $Im\omega$ — вещественная и, соответственно, мнимая часть комплексного числа ω.

Предложение 8. ***а) Если*** $\operatorname{Re}\omega_1=\operatorname{Re}\omega_2$, ***т. е.*** ω_1 ***и*** ω_2 ***лежат на луче, перпендикулярном к оси*** Ox, ***то расстояние между*** ω_1 ***и*** ω_2 ***равно*** $\left|\ln\left(\frac{Im\omega_2}{Im\omega_1}\right)\right|$;

б) если $\operatorname{Re}\omega_1\neq\operatorname{Re}\omega_2$, ***то расстояние между*** ω_1 ***и*** ω_2 ***равно*** $\left|\ln\left(\frac{\operatorname{tg}\varphi_2/2}{\operatorname{tg}\varphi_1/2}\right)\right|$, ***где*** φ_1 ***и*** φ_2 ***— аргументы комплексных чисел*** $\omega_1-\omega_0$ ***и*** $\omega_2-\omega_0$ ***и*** ω_0 ***— центр окружности, проходящей через*** ω_1 ***и*** ω_2 ***перпендикулярно оси*** Ox.

▷ а) Пусть $\omega_1=x+iy_1$ и $\omega_2=x+iy_2$. Тогда $u=x$ и $v=\infty$, и поэтому

$$|\ln((\omega_1,\ \omega_2)(u,\ v))|=\left|\ln\frac{x+iy_1-x}{x+iy_1-\infty}:\right.$$

$$\left.:\frac{x+iy_2-x}{x+iy_2-\infty}\right|=\lim_{v\to\infty}\left|\ln\frac{y_1}{y_2}\cdot\frac{x+iy_2-v}{x+iy_1-v}\right|=\left|\ln\frac{y_2}{y_1}\right|.$$

б) Пусть точка ω лежит на верхней полуокружности с центром в вещественной точке ω и u и v — точки пересечения полуокружности с осью Ox (рис. 126).

Можно считать, что $u<v$. Найдем отношение $\frac{\omega-u}{\omega-v}$.

Пусть $|\omega-u|$ и $|\omega-v|$ модули соответствующих комплексных чисел, т. е. длины отрезков $[\omega,\ u]$ и $[\omega,\ v]$. Обозначив за α и β аргументы комплексных чисел $\omega-u$ и $\omega-v$, получаем, что

$$\frac{\omega-u}{\omega-v}=\frac{|\omega-u|e^{i\alpha}}{|\omega-v|e^{i\beta}}=\frac{|\omega-u|e^{i(\alpha-\beta)}}{|\omega-v|}=\frac{|\omega-u|}{|\omega-v|}e^{-i\frac{\pi}{2}}=$$

$$=-i\frac{|\omega-u|}{|\omega-v|},\ \text{так как}\ \beta=\alpha+\frac{\pi}{2}.$$

Поэтому $\frac{\omega-u}{\omega-v}=-i\operatorname{ctg}\alpha=-i\operatorname{ctg}\frac{\varphi}{2}$, где φ-аргумент комплексного числа $\omega-\omega_0$ и $\varphi=2\alpha$, так как φ — центральный угол, опирающийся на ту же дугу, что и вписанный угол α.

Следовательно, расстояние между ω_1 и ω_2 (рис. 127) равно

$$|\ln((\omega_1,\ \omega_2)(u,\ v))|=\left|\ln\frac{\omega_1-u}{\omega_1-v}:\frac{\omega_2-u}{\omega_2-v}\right|=\left|\ln\frac{-i\operatorname{ctg}\frac{\varphi_1}{2}}{-i\operatorname{ctg}\frac{\varphi_2}{2}}\right|=$$

$$=\left|\ln\frac{\operatorname{tg}\frac{\varphi_2}{2}}{\operatorname{tg}\frac{\varphi_1}{2}}\right|.$$

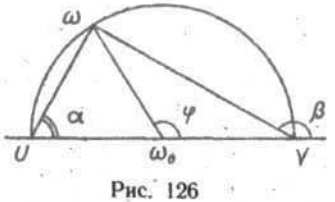


Рис. 126

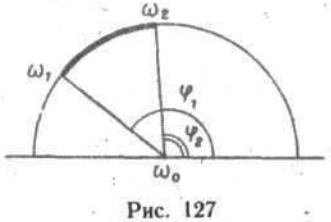


Рис. 127

Следствие. а) Всякий сдвиг вдоль вещественной оси является движением в модели H^2P_2;

б) всякая гомотетия с вещественным центром и положительным коэффициентом является движением в модели H^2P_2. (Напомним, что гомотетией с центром в точке a и коэффициентом λ называется преобразование вида $H_{a,\lambda}(x)\rightleftarrows a+\lambda\cdot(x-a):\boldsymbol{R}^2\to\boldsymbol{R}^2$).

§ 3. Группы движений и измерения углов в сферической, эллиптической и гиперболической плоскостях

1. В любом метрическом пространстве M с метрикой ρ в нем определены понятия движения и конгруэнтности.

Определение 10. **а) Движением (изометрией) метрического пространства** $\langle M,\ \rho\rangle$ **называется любое отображение** f **пространства** M **на себя, для которого** $\rho(x,\ y)=\rho(f(x),\ f(y))$, $\forall x,\ y\in M$. **Множество всех движений пространства** $\langle M,\ \rho\rangle$ **обозначим** $\operatorname{Iso}M$.

б) Два подмножества X_1 **и** X_2 **метрического пространства** $\langle M,\ \rho\rangle$ **называются конгруэнтными, если существует движение в пространстве** M, **переводящее** X_1 **на** X_2. **Конгруэнтность множеств** X_1 **и** X_2 **записывается** $X_1\sim X_2$.

Упражнение 11. *Докажите, что:*

а) множество $\operatorname{Iso}M$ *относительно операции композиции является группой;*

б) обозначим $\operatorname{Iso}(M,\ x_0)\rightleftarrows\{f\in\operatorname{Iso}M\mid f(x_0)=x_0\}$; *тогда* $\operatorname{Iso}(M,\ x_0)$ *— подгруппа группы* $\operatorname{Iso}M$;

в) отношение конгруэнтности в семействе всех подмножеств метрического пространства $\langle M,\ \rho\rangle$ *является отношением эквивалентности;*

г) в сферической, эллиптической и гиперболической плоскостях с определенными в них метриками (см. § 2) выполняется аксиома III_2 *аксиоматики Евклида — Гильберта.*

Цель этого параграфа — рассмотреть группы движений и понятие конгруэнтности в s-, e- и h-плоскостях. Сразу отметим, что группы $\operatorname{Iso}S^2$, $\operatorname{Iso}E^2$ и $\operatorname{Iso}H^2$ попарно не изоморфны. Замечательно, что тем не менее подгруппы $\operatorname{Iso}(S^2,\ x_0)$, $\operatorname{Iso}(E^2,\ x_0)$ и $\operatorname{Iso}(H^2,\ x_0)$ этих групп, т. е. группы движений с неподвижной точкой x_0, изоморфны между собой (для любой точки x_0) и все они изоморфны хорошо известной группе — группе всех движений $\operatorname{Iso}(\boldsymbol{R}^2,\ \bar{O})$ арифметической плоскости $\boldsymbol{R}^2$ с неподвижной точкой $\bar{O}$.

Начнем с рассмотрения движений в сферической плоскости. Для подмножества X множества M и отображения $f:X\to Y$ назовем отображение $g:M\to Y$ продолжением отображения f на M, если $g\restriction X=f$. Одно и то же отображение имеет, вообще говоря, много различных продолжений.

Теорема 6. ***а) Существует такой изоморфизм*** ψ ***группы*** $\operatorname{Iso}S^2$ ***на группу*** $O_3\rightleftarrows\operatorname{Iso}(\boldsymbol{R}^3,\ \boldsymbol{O})$, ***что для всех*** $f\in\operatorname{Iso}S^2$ ***движение*** $\psi(f)\in O_3$ ***есть продолжение движения*** f.

б) Для любой s-точки A подгруппа $\mathrm{Iso}(S^2, A)$ *группы* $\mathrm{Iso}\,S^2$ *всех движений s-плоскости изоморфна группе* O_2.

▷ а) Как всегда, в арифметическом пространстве $\boldsymbol{R}^3$ рассматривается метрика

$$\rho(A, B)=\|A-B\|\rightleftarrows\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2},$$

где $A=\langle x_1, y_1, z_1\rangle$ и $B=\langle x_2, y_2, z_2\rangle$. Построим отображение $\psi: \mathrm{Iso}\,S^2\to O_3$. Если $f\in\mathrm{Iso}\,S^2$, то для всех $A\in\boldsymbol{R}^3$ положим

$$\psi(f)(A)=\begin{cases}\bar{O}, & \text{если } A=\bar{O}\\ \|A\|\cdot f\left(\frac{A}{\|A\|}\right), & \text{если } A\neq\bar{O}.\end{cases}$$

Докажем, что $\psi(f)\in\mathrm{Iso}(\boldsymbol{R}^3)$. Если одна из точек A или B совпадает с $\bar{O}$, то равенство

$$\|\psi(f)(A)-\psi(f)(B)\|=\|A-B\|,$$

очевидно, верно. Пусть A и B — отличные от $\bar{O}$ точки из $\boldsymbol{R}^3$ и $A_1\rightleftarrows\frac{A}{\|A\|}$, $B_1\rightleftarrows\frac{B}{\|B\|}$. Тогда $A_1\in S^2$ и $B_1\in S^2$ и так как по построению отображения ψ отображение $\psi(f)$ является продолжением f, то $f(A_1)\in S^2$, $f(B_1)\in S^2$ и $\|f(A_1)-f(B_1)\|=\|A_1-B_1\|$. Следовательно треугольники $\{\bar{O}, A_1, B_1\}$, $\{\bar{O}, f(A_1), f(B_1)\}$ равны между собой и угол $\angle f(A_1), \bar{O}, f(B_1)$ равен углу $\angle A_1, \bar{O}, B_1$. Рассмотрим треугольники $\{\bar{O}, A, B\}$ и $\{\bar{O}, \psi(f)(A), \psi(f)(B)\}$. В них $\|\psi(f)(A)\|=\|\,\|A\|\cdot f(A_1)\|=\|A\|\cdot\|f(A_1)\|=\|A\|$ и аналогично $\|\psi(f)(B)\|=\|B\|$. Кроме того, угол $\angle\,\psi(f)(A), \bar{O}, \psi(f)(B)$ равен углу $\angle f(A_1), \bar{O}, f(B_1)$, а угол $\angle A, \bar{O}, B$ равен углу $\angle A_1, \bar{O}, B_1$ (рис. 128). Следовательно, эти треугольники равны между собой, и поэтому

$$\|\psi(f)(A)-\psi(f)(B)\|=\|A-B\|.$$

Итак, построено такое отображение ψ из $\mathrm{Iso}(S^2)$ в O_3, что $\psi(f)$ есть продолжение f на $\boldsymbol{R}^3$ для всех $f\in\mathrm{Iso}\,S^2$.

Проверим гомоморфность отображения ψ. Пусть f и g из $\mathrm{Iso}(S^2)$ и $A\in\boldsymbol{R}^3$. Так как

$$\left\|\,\|A\|\cdot g\left(\frac{A}{\|A\|}\right)\right\|=\|A\|\cdot\left\|g\left(\frac{A}{\|A\|}\right)\right\|=\|A\|,$$

то

$$\psi(f)\circ\psi(g)(A)=\psi(f)\left(\|A\|\cdot g\left(\frac{A}{\|A\|}\right)\right)=\left\|\,\|A\|\cdot g\left(\frac{A}{\|A\|}\right)\right\|\times$$

$$\times\ f\left(\frac{\|A\|\cdot g\left(\frac{A}{\|A\|}\right)}{\left\|\,\|A\|\cdot g\left(\frac{A}{\|A\|}\right)\right\|}\right)=\|A\|\cdot f\left(g\left(\frac{A}{\|A\|}\right)\right)=\psi(f\circ g)(A).$$

Следовательно, $\psi(f)\circ\psi(g)=\psi(f\circ g)$. Проверим инъективность отображения ψ, т. е. проверим, что $\mathrm{Ker}\,\psi=\mathrm{id}\restriction S^2$. Пусть $\psi(f)=\mathrm{id}\restriction\boldsymbol{R}^3$. Тогда $\psi(f)(A)=\|A\|\times$ $\times f\left(\frac{A}{\|A\|}\right)=A$ для всех точек $A\neq\bar{O}$. Если $\|A\|=1$, т. е. $A\in S^2$, то $f(A)=A$, т. е. $f=\mathrm{id}\restriction S^2$.

Наконец, проверим сюръективность отображения ψ. Пусть W любое линейное движение в $\boldsymbol{R}^3$, для которого $W(\bar{O})=\bar{O}$. Положим $f(B)\rightleftarrows W(B)$ для любых $B\in S^2$. Тогда $f\in\mathrm{Iso}\,S^2$ и $\psi(f)=W$, так как в силу линейности W имеем, что $\psi(f)(A)=\|A\|\cdot f\left(\frac{A}{\|A\|}\right)=\|A\|\times$ $\times W\left(\frac{A}{\|A\|}\right)=\|A\|\cdot\frac{1}{\|A\|}W(A)=W(A)$ для любых $A\in\boldsymbol{R}^3$, $A\neq\bar{0}$.

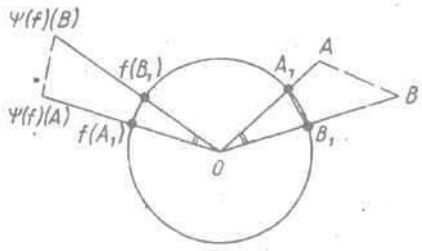


Рис. 128

Итак, ψ — изоморфизм групп $\mathrm{Iso}(S^2)$ и O_3.

б) Фиксируем s-точку A и пусть $A_\perp$ — плоскость, проходящая через начало координат $\bar{O}$ ортогонально прямой $(\bar{O}A)$. Если $f\in\mathrm{Iso}(S^2, A)$ то, по построению из п. а,

$$\psi(f)\in\mathrm{Iso}(\boldsymbol{R}^3, \bar{O})\text{ и }\psi(f)(A)=f(A)=A$$

Так как $\psi(f)$ есть движение $\boldsymbol{R}^3$ с неподвижными точками $\bar{O}$ и A, то для любой точки B из плоскости $A_\perp$ точка $\psi(f)(B)$ также лежит в плоскости $A_\perp$, т. е. $\psi(f)\restriction A_\perp$ — движение плоскости $A_\perp$ с неподвижной точкой $\bar{O}$. Изоморфность отображения ψ означает, что отображение

$$f\mapsto\psi(f)\restriction A_\perp$$

является изоморфизмом группы $\mathrm{Iso}(S^2, A)$ на группу $\mathrm{Iso}(A_\perp, \bar{O})$. Остается доказать, что последняя группа изоморфна группе $O_2\rightleftarrows$ $\rightleftarrows\mathrm{Iso}(\boldsymbol{R}^2, \bar{O})$.

В плоскости $A_\perp$ выберем (см. глава II, § 4) базис $\{e_1', e_2'\}$ ортонормированный относительно ограничения на $A_\perp$ скалярного произведения из $\boldsymbol{R}^3$ и определим изоморфизм W из $A_\perp$ в $\boldsymbol{R}^2$ равенством

$$W(x\cdot e_1'+y\cdot e_2')\rightleftarrows x\cdot e_1+y\cdot e_2$$

где $\{e_1, e_2\}$ — обычный базис в $\boldsymbol{R}^2$. Тогда для любых точек B и C из $A_\perp$

$$\|B-C\|_{A_\perp}=\|W(B)-W(C)\|_{\boldsymbol{R}^2}$$

и соответствие $T\mapsto W\circ T\circ W^{-1}$, где $T\in\mathrm{Iso}(A_\perp, \bar{O})$, задает изоморфизм группы $\mathrm{Iso}(A_\perp, \bar{O})$ на группу $\mathrm{Iso}(\boldsymbol{R}^2, \bar{O})$. ■

Предложение 9. ***Всякое движение пространства*** $\boldsymbol{R}^3$ ***с неподвижной точкой*** $\bar{O}$ ***является линейным отображением*** $\boldsymbol{R}^3$ ***в себя.***

▷ Пусть $W\in\mathrm{Iso}(\boldsymbol{R}^3, \bar{O})$ и C — середина некоторого отрезка $[A, B]$, где $A, B\in\boldsymbol{R}^3$. Тогда

$$\|W(C)-W(A)\|=\|W(C)-W(B)\|=\frac{1}{2}\|W(A)-W(B)\|.$$

Следовательно, в треугольнике $\{W(A), W(C), W(B)\}$ сумма длин двух сторон равна длине третьей стороны. Поэтому точка $W(C)$ лежит на прямой $(W(A), W(B))$, и так как она равноудалена от его концов, то $W(C)$ — середина отрезка $[W(A), W(B)]$. Итак, для любых $A, B \in \boldsymbol{R}^3$

$$W\left(\frac{A+B}{2}\right)=\frac{W(A)+W(B)}{2}.$$

В частности, при $B=\bar{O}$ в силу условия $W(\bar{O})=\bar{O}$ получаем, что $W\left(\frac{A}{2}\right)=\frac{1}{2}W(A)$ для всех $A \in \boldsymbol{R}^3$. Поэтому для любых $A, B \in \boldsymbol{R}^3$

$$W(A+B)=W\left(\frac{2\cdot A+2\cdot B}{2}\right)=\frac{W(2\cdot A)+W(2\cdot B)}{2}=W(A)+W(B),$$

так как $\frac{1}{2}W(2\cdot A)=W\left(\frac{2\cdot A}{2}\right)=W(A)$ и аналогично $\frac{1}{2}W(2\cdot B)=W(B)$.

Теперь проверим, что для всех $\lambda \in \boldsymbol{R}$ и $A \in \boldsymbol{R}^3$

$$W(\lambda\cdot A)=\lambda\cdot W(A)$$

Если $\lambda=\frac{1}{2^n}$, $n \in \boldsymbol{N}$, то это равенство получается по индукции из равенства $W\left(\frac{A}{2}\right)=\frac{1}{2}W(A)$. Если $\lambda=\frac{m}{2^n}$, $n, m \in \boldsymbol{N}$, то

$$W(\lambda\cdot A)=\underbrace{W\left(\frac{A}{2^n}+\ldots+\frac{A}{2^n}\right)}_{m \text{ раз}}=\underbrace{W\left(\frac{A}{2^n}\right)+\ldots+W\left(\frac{A}{2^n}\right)}_{m \text{ раз}}=$$

$$=\frac{m}{2^n}\cdot W(A)=\lambda\cdot W(A).$$

Если $\lambda>0$, то $\lambda=\lim\limits_{k\to\infty}\lambda_k$, где λ_k — двоично-рациональные числа, т. е. числа вида $\frac{m}{2^n}$. Так как отображение W сохраняет расстояние, то W — непрерывное отображение $\boldsymbol{R}^3$ в себя, и поэтому

$$W(\lambda\cdot A)=W(\lim_{k\to\infty}\lambda_k\cdot A)=W(\lim_{k\to\infty}(\lambda_k\cdot A))=\lim_{k\to\infty}W(\lambda_k\cdot A)=$$
$$=\lim_{k\to\infty}(\lambda_k\cdot W(A))=(\lim_{k\to\infty}\lambda_k)\cdot W(A)=\lambda\cdot W(A).$$

Наконец, если $\lambda<0$, то

$$W(\lambda\cdot A)+W((-\lambda)\cdot A)=W(\lambda\cdot A+(-\lambda)\cdot A)=W(\bar{O})=\bar{O},$$

и поэтому $W(\lambda\cdot A)=-W((-\lambda)\cdot A)=-(-\lambda)\cdot W(A)=\lambda\cdot W(A)$. ■

Из предложения 9 следует, что любой элемент группы $\mathrm{Iso}(\boldsymbol{R}^3, \bar{O})$ задается матрицей размера (3×3). Как и в главе II, мы не будем делать различия между линейным отображением пространства $\boldsymbol{R}^3$ в себя и той матрицей, которой это линейное отображение задается.

Обозначим W^t матрицу, транспонированную к матрице W. Нетрудно видеть, что при указанном отождествлении

$$O_3=\{W \in \mathrm{Mat}(3\times 3) \mid W\cdot W^t=W^t\cdot W=E\},$$

где

$$E=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

Упражнение 12. *Докажите, что:*

а) если $W \in O_3$, *то* $\det W=\pm 1$;

б) множество $\mathrm{SO}_3 \rightleftarrows \{W \in O_3 \mid \det W=1\}$ — *подгруппа группы* O_3;

в) множество $\{E, -E\}$ — *нормальный делитель группы* O_3, *а фактор-группа* $O_3/\{E, -E\}$ *изоморфна группе* SO_3;

г) любой элемент W *из* SO_3 *задает линейное отображение* $\boldsymbol{R}^3$ *в себя, являющееся поворотом вокруг некоторой прямой, проходящей через начало координат* $\bar{O}$;

д) если $\psi : O_3 \to G$ — *такой сюръективный гомоморфизм группы* O_3 *на группу* G, *что* $\mathrm{Ker}\,\psi=\{E, -E\}$, *то существует изоморфизм* $\varphi : G \to \mathrm{SO}_3$ *группы* G *на группу* SO_3, *для которого* $\psi\circ\varphi=\mathrm{id}_G$.

2. Перейдем к движениям в эллиптической плоскости E^2. Для каждого движения $f \in \mathrm{Iso}\,S^2$ определим отображение $f_\pm$ e-плоскости E^2 в себя вида

$$f_\pm(\{A, -A\}) \rightleftarrows \{f(A), -f(A)\}.$$

Предложение 10. ***Отображение*** $\varphi : f \mapsto f_\pm$ ***является гомоморфизмом группы*** $\mathrm{Iso}\,S^2$ ***в группу*** $\mathrm{Iso}\,E^2$, ***ядро которого состоит из двух элементов*** id_{S^2} ***и*** $-\mathrm{id}_{S^2}$.

▷ Пусть $A_e=\{A, -A\}$ и $B_e=\{B, -B\}$ — две e-точки. Не уменьшая общности, можно считать, что B — точка, для которой радианная мера угла $\angle AOB$ не больше $\pi/2$, т. е. число $\rho(A_e, B_e)$ равно длине дуги $\overset{\smile}{A, B}$ (рис. 129). Так как $f \in \mathrm{Iso}\,S^2$, то длина дуги $f(A), f(B)$ равна длине дуги $\overset{\smile}{A, B}$ и поэтому не превосходит $\pi/2$. Но тогда

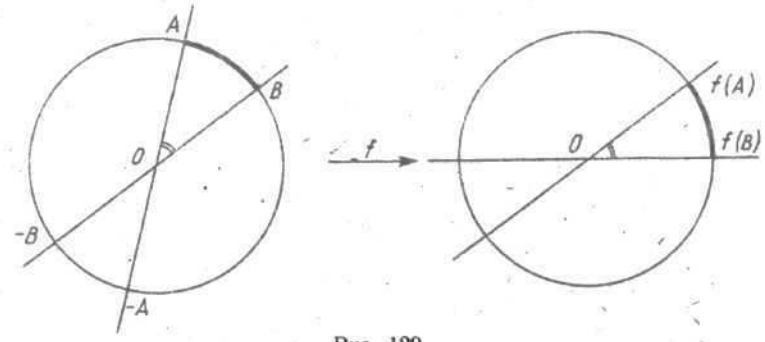


Рис. 129

$\rho(\{f(A), -f(A)\}, \{f(B), -f(B)\})$ равно длине дуги $f(A)$, $f(B)$, т. е. $\rho(f_\pm(A_e), f_\pm(B_e)) = \rho(A_e, B_e)$. Следовательно, $f_\pm \in \mathrm{Iso}(E^2)$.

Пусть $f, g \in \mathrm{Iso}\, S^2$. Тогда для любой e-точки $A_e = \{A, -A\}$ выполняется

$$(f_\pm \circ g_\pm)(A_e) = f_\pm(g_\pm(A_e)) = f_\pm(\{g(A), -g(A)\}) = \{f(g(A)), -f(g(A))\} = (f \circ g)_\pm(A_e),$$

т. е. отображение $\varphi : f \mapsto f_\pm$ является гомоморфизмом группы $\mathrm{Iso}\, S^2$ в группу $\mathrm{Iso}\, E^2$.

Пусть $f \in \mathrm{Ker}\, \varphi$, т. е. для любой e-точки $A_e = \{A, -A\}$ выполняется $f_\pm(A_e) = A_e$. Тогда для любой s-точки A

$$f(A) = A \text{ или } f(A) = -A.$$

Рассмотрим такие s-точки A_1, A_2, A_3, A_4, A_5, для которых векторы $\overline{O, A_i}$, $\overline{O, A_j}$, $\overline{O, A_k}$ линейно независимы для любых попарно различных i, j, k из множества $\{1, 2, 3, 4, 5\}$. Тогда для каких-то трех из этих точек, например, для A_1, A_2, A_3 выполняется $f(A_i) = A_i$ или $f(A_i) = -A_i$, $i = 1, 2, 3$. Так как вектора $\overline{O, A_1}$, $\overline{O, A_2}$, $\overline{O, A_3}$ образуют базис пространства $\boldsymbol{R}^3$, то продолжение $\psi(f)$ движения f на все пространство $\boldsymbol{R}^3$ (см. теорема 6,а) совпадает с $\mathrm{id}_{\boldsymbol{R}^3}$ или с $-\mathrm{id}_{\boldsymbol{R}^3}$. Следовательно, и $f = \psi(f) \upharpoonright S^2$ совпадает с id_{S^2} или с $-\mathrm{id}_{S^2}$. ■

Теорема 7. ***а) Для любой e-точки A_e группа*** $\mathrm{Iso}\,(E^2, A_e)$ ***изоморфна группе O_2.***

б) Группа $\mathrm{Iso}\,(E^2)$ ***изоморфна группе*** SO_3.

▷ а) Пусть $T \in \mathrm{Iso}\,(E^2, A_e)$. Для каждого $\alpha \in \left]0, \frac{\pi}{2}\right[$ рассмотрим множество S_α всех s-точек, удаленных от s-точки A на s-расстояние α; S_α есть окружность, являющаяся пересечением сферы S^2 с плоскостью, перпендикулярной прямой $(\overline{OA})$ и удаленной от точки A на евклидово расстояние $1 - \cos\alpha$ (рис. 130).

Для любой s-точки $B \in S_\alpha$ соответствующая e-точка $B_e = \{B, -B\}$ при движении T перейдет в некоторую e-точку C_e; при этом ровно одна из точек C или $-C$, задающих e-точку $C_e = \{C, -C\}$, лежит на окружности S_α. Положим по определению $T_\alpha(B)$ равной именно этой s-точке.

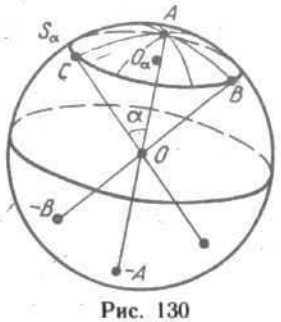


Рис. 130

Упражнение 13. *Докажите, что при фиксированном* $\alpha \in \left]0, \frac{\pi}{2}\right[$;

а) отображение T_α есть движение окружности S_α;

б) соответствие $T \mapsto T_\alpha$ задает гомоморфизм группы $\mathrm{Iso}\,(E^2, A_e)$ *в группу* $\mathrm{Iso}\,(S_\alpha)$;

в) если $B \in S_\alpha$ и точка C есть пересечение окружности S_β, $\beta \in \left]0, \frac{\pi}{2}\right[$ с полуплоскостью, определяемой прямой $(\overline{OA})$ и точкой B, то точка $T_\beta(C)$ есть пересечение окружности S_β с полуплоскостью, определяемой прямой $(\overline{OA})$ и точкой B.

Упражнение 13,в показывает, что движение T индуцирует во всех окружностях S_α, $\alpha \in \left]0, \frac{\pi}{2}\right[$ «одно и то же» движение. В теореме 6,а доказано, что группа $\mathrm{Iso}\,(S^2)$ движений сферы изоморфна группе O_3. Аналогичный результат верен и для группы $\mathrm{Iso}\,(S_\alpha)$ движений окружности S_α: она изоморфна группе O_2. Поэтому для доказательства п. а этой теоремы осталось проверить, что гомоморфизм вида $T \mapsto T_\alpha$ есть изоморфизм группы $\mathrm{Iso}\,(S^2, A_e)$ на группу $\mathrm{Iso}\,(S_\alpha)$ для любого фиксированного $\alpha \in \left]0, \frac{\pi}{2}\right[$.

Пусть $T_\alpha = \mathrm{id}_{S_\alpha}$. По упражнению 13,в получаем, что для всех $\beta \in \left]0, \frac{\pi}{2}\right[$ движение T_β есть id_{S_β}. По построению движений T_β это означает, что движение $T \in \mathrm{Iso}\,(S^2, A_e)$ тождественно на всех e-точках B_e, удаленных от A_e на e-расстояние меньше $\frac{\pi}{2}$. Так как любая e-точка C_e, для которой

$$\rho_e(A_e, C_e) = \frac{\pi}{2}$$

есть предел последовательности e-точек $B_{e,n}$, для которых

$$\rho_e(A_e, B_{e,n}) < \frac{\pi}{2},$$

и так как движение T, как и всякое движение, непрерывно, то

$$T(C_e) = T(\lim_{n\to\infty} B_{e,n}) = \lim_{n\to\infty} T(B_{e,n}) = \lim_{n\to\infty} B_{e,n} = C_e.$$

Следовательно, $T = \mathrm{id}_{E^2}$ и гомоморфизм $T \mapsto T_\alpha$ инъективен.

Пусть $f \in \mathrm{Iso}\,(S_\alpha)$. Докажем, что для некоторого $T \in \mathrm{Iso}\,(E^2, A_e)$ выполняется $T_\alpha = f$. Группа $\mathrm{Iso}\,(S_\alpha)$ изоморфна группе O_2 линейных движений арифметической плоскости $\boldsymbol{R}^2$, а всякое линейное движение плоскости $\boldsymbol{R}^2$ есть композиция ортогональных отражений. Поэтому в силу гомоморфности отображения $T \mapsto T_\beta$ достаточно проверить существование таких $T \in \mathrm{Iso}\,(E^2, A_e)$, что $T_\alpha = f$, лишь для движений f, которые задаются ортогональными отражениями. Обозначим π_α плоскость, содержащую окружность S_α, и пусть O_α центр этой окружности. Пусть отражение f есть отражение в плоскости π_α относительно прямой $(\overline{O, B})$, $B \in S_\alpha$. Пусть g — движение сферы S^2, являющееся ортогональным отражением относительно плоскости $(AO_\alpha B)$, и $g_\pm \rightleftarrows T$ — соответствующее e-движение e-плоскости E^2. Тогда по построению $T \in \mathrm{Iso}\,(E^2, A_e)$ и $T_\alpha = f$. ■

б) В силу результата упражнения 12 д) для доказательства этого пункта достаточно проверить сюръективность гомоморфизма

$\varphi: f \mapsto f_{\pm}$ из группы $\mathrm{Iso}(S^2)$, изоморфной группе O_2, в группу $\mathrm{Iso}(E^2)$. Напомним, что $\mathrm{Ker}\,\varphi = \{\mathrm{id}_{S^2}, -\mathrm{id}_{S^2}\}$, см. предложение 10.

Пусть у движения $T \in \mathrm{Iso}(E^2)$ есть неподвижная точка A_e, т. е. $T \in \mathrm{Iso}(E^2, Ae)$, и пусть на окружности S_α, $\alpha \in \left]0, \frac{\pi}{2}\right[$, движение T действует как некоторое ортогональное отражение. Тогда движение $g \in \mathrm{Iso}(S^2)$, для которого $g_\pm = T$ построено выше в п. а).

Пусть $T \in \mathrm{Iso}(E^2, A_e)$. Тогда по п. а) T однозначно определяется движением $T_\alpha \in \mathrm{Iso}(S_\alpha)$ при некотором $\alpha \in \left]0, \frac{\pi}{2}\right[$. Разложим движение T_α в композицию ортогональных отражений. Тогда движение T есть композиция $T_1 \circ \ldots \circ T_n$, где $T_i \in \mathrm{Iso}(E^2, A_e)$ и движения $T_{i,\alpha}$ в окружности S_α есть отражения, $1 \leqslant i \leqslant n$. Следовательно, для некоторых движений $g_i \in \mathrm{Iso}(S^2)$ выполняется $(g_i)_\pm = T_i$. В силу гомоморфности $\varphi: f \mapsto f_\pm$ для движения $g \rightleftarrows g_1 \circ \ldots \circ g_n$ выполняется

$$g_\pm = (g_1 \circ \ldots \circ g_n)_\pm = (g_1)_\pm \circ \ldots \circ (g_n)_\pm = T_1 \circ \ldots \circ T_n = T_n.$$

Пусть движение $T \in \mathrm{Iso}(E^2)$ не имеет неподвижных точек. Тогда возьмем любую s-точку A и любую из диаметрально противоположных s-точек, задающих e-точку $T(\{A, -A\})$. Рассмотрим движение $g \in \mathrm{Iso}(S^2)$, переводящее эту s-точку в s-точку A. Тогда композиция e-движений $(g)_\pm \circ T$ оставляет неподвижной e-точку $A_e = \{A, -A\}$ и по доказанному выше $(g)_\pm \circ T = f_\pm$ для некоторого движения $f \in \mathrm{Iso}(E^2)$. Отсюда получаем, что $T = (g)_\pm^{-1} \circ f_\pm = (g^{-1} \circ f)_\pm$, где $g^{-1} \circ f \in \mathrm{Iso}(S^2)$. ■

3. Группа движений гиперболической плоскости сложнее групп, разобранных в пунктах 1, 2, т. е. групп $\mathrm{Iso}(S^2)$ и $\mathrm{Iso}(E^2)$.

Для упрощения технических деталей будем рассматривать группу движений второй модели Пуанкаре H^2P_2. Нам потребуется следующий факт из курса теории аналитических функций: всякое отображение f комплексной плоскости $\boldsymbol{C}$ на себя, сохраняющее сложное отношение комплексных чисел, представимо в виде

$$\omega = f(z) = \frac{az+b}{cz+d} \text{ или } \omega = f(z) = \frac{a\bar{z}+b}{c\bar{z}+d},$$

где $a, b, c, d \in \boldsymbol{C}$ и $ad - bc \neq 0$. Напомним, что $\bar{z}$ — комплексное число, сопряженное с числом z.

Так как согласно § 2 h-расстояние в H^2P_2 как раз задается сложным отношением, то движение в H^2P_2 следует искать среди отображений вида $\omega = \frac{az+b}{cz+d}$ или $\omega = \frac{a\bar{z}+b}{c\bar{z}+d}$, которые называются дробно-линейными отображениями соответственно первого и второго рода.

Упражнение 14. *Докажите, что:*

а) множество всех дробно-линейных отображений первого рода, т. е. множество всех отображений вида $\omega = \frac{az+b}{cz+d}$, *где* $a, b, c, d \in \boldsymbol{C}$ *и* $ad - bc \neq 0$, *относительно операции композиции образует группу;*

б) всякое отображение вида $\omega = \frac{az+b}{cz+d}$, *где* $a, b, c, d \in \boldsymbol{C}$ *и* $ad - bc \neq 0$, *сохраняет сложное отношение комплексных чисел;*

в) отображение $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \omega(\cdot)$, *где* $\omega(z) \rightleftarrows \frac{az+b}{cz+d}$, *является гомоморфизмом группы* $\mathrm{GL}_2(\boldsymbol{C})$ *всех невырожденных комплексных матриц размером* 2×2 *на группу всех дробно-линейных отображений первого рода;*

г) множество всех дробно-линейных отображений вида $\omega = \frac{az+b}{cz+d}$, *где* $a, b, c, d \in \boldsymbol{R}$ *и* $ad - bc = 1$ *относительно операции композиции образует группу;*

д) множество всех дробно-линейных отображений второго рода вида $\omega = \frac{a\bar{z}+b}{c\bar{z}+d}$, *где* $a, b, c, d \in \boldsymbol{R}$ *и* $ad - bc = -1$, *не образует группы; однако в объединении с множеством всех дробно-линейных отображений* из п. г *оно (относительно операции композиции) образует группу.*

Теорема 8. *а) Любое дробно-линейное отображение вида*

$$\omega = \frac{az+b}{cz+d}, \text{ где } a, b, c, d \in R \text{ и } ad - bc = 1,$$

или вида $\psi = \frac{a\bar{z}+b}{c\bar{z}+d}$, *где* $a, b, c, d \in R$ *и* $ad - bc = -1$ *является движением во второй модели Пуанкаре геометрии Лобачевского.*

б) Любое движение во второй модели Пуанкаре геометрии Лобачевского имеет один из двух видов, указанных в пункте а.

▷ а) Пусть $\omega(z) = \frac{az+b}{cz+d}$, где $a, b, c, d \in \boldsymbol{R}$ и $ad - bc = 1$. Тогда

$$\omega(z) = \frac{(az+b)(\overline{cz+d})}{|cz+d|^2} = \frac{(az+b)(c\bar{z}+d)}{|cz+d|^2} = \frac{ac|z|^2+bd}{|cz+d|^2} + \frac{adz+bc\bar{z}}{|cz+d|^2} =$$

$$= \frac{ac|z|^2+bd}{|cz+d|^2} + \frac{ad+bc}{|cz+d|^2}\operatorname{Re} z + i\frac{ad-bc}{|cz+d|^2}\operatorname{Im} z, \text{ т. е.}$$

$$\operatorname{Im}\omega(z) = \frac{1}{|cz+d|^2}\operatorname{Im} z.$$

Поэтому из условия $\operatorname{Im} z>0$ следует условие $\operatorname{Im}\omega(z)>0$, т. е. отображение $\omega(\cdot)$ переводит верхнюю полуплоскость на себя. Кроме того, $\omega(\cdot)$, как и всякое дробно-линейное отображение первого рода, сохраняет сложное отношение четверок комплексных чисел (упражнение 14,б). Следовательно, $\omega(\cdot)$ отображает H^2P_2 на себя и при этом сохраняет h-расстояния, т. е. является движением во второй модели Пуанкаре.

Пусть $\omega(z)=\frac{a\bar{z}+b}{a\bar{z}+d}$, где $a, b, c, d\in\boldsymbol{R}$ и $ad-bc=-1$, и $\omega_0(\cdot)$ — отображение, задаваемое формулой $\omega_0(z)\rightleftarrows-\bar{z}$, т. е. $\omega_0(\cdot)$ является отражением относительно оси ординат. Нетрудно видеть, что $\omega_0(\cdot)\in\operatorname{Iso}(H^2P_2)$, и, если $\omega_1(z)\rightleftarrows\frac{(-a)z+b}{(-c)z+d}$, то $\omega(z)=\omega_1(\omega_0(z))$. Как показано выше, отображение $\omega_1(\cdot)$ является движением в H^2P_2, так как $(-a)$, b, $(-c)$, $d\in\boldsymbol{R}$ и $(-a)d+b(c)=1$. Следовательно, отображение $\omega(\cdot)$, как композиция двух движений, является движением в H^2P_2.

б) Пусть движение во второй модели Пуанкаре является дробно-линейным отображением первого рода, т. е. $\omega=\frac{az+b}{cz+d}$, где a, b, c, $d\in\boldsymbol{C}$ и $ad-bc\neq0$. Тогда $\omega(\cdot)$ переводит верхнюю полуплоскость в себя и переводит ось абсцисс в себя, т. е. для всех $x\in\boldsymbol{R}$ выполняется

$$\overline{\omega(x)}=\overline{\left(\frac{ax+b}{cx+d}\right)}=\frac{\bar{a}x+\bar{b}}{\bar{c}x+\bar{d}}=\omega(x)=\frac{ax+b}{cx+d}.$$

Пусть $c\neq0$ и $d\neq0$. Тогда при $x\to\infty$ получим $\frac{\bar{a}}{\bar{c}}=\frac{a}{c}$ или $\left(\frac{\bar{a}}{c}\right)=\frac{a}{c}=\lambda\in\boldsymbol{R}$, а при $x=0$ получим: $\left(\frac{\bar{b}}{d}\right)=\frac{b}{d}=\mu\in\boldsymbol{R}$ т. е. $a=\lambda c$ и $b=\mu d$, где $\lambda, \mu\in\boldsymbol{R}$. При $x=1$ получаем:

$$\frac{\lambda c+\mu d}{c+d}\in\boldsymbol{R},\ \text{т. е.}\ \frac{c+\frac{\mu}{\lambda}d}{c+d}=r\in\boldsymbol{R}.$$

Допустим, что комплексные числа 0, c, d не лежат на одной прямой. Тогда множество $\{c+td\,|\,t\in\boldsymbol{R}\}$ — прямая, проходящая через точки c и $c+d$, а множество $\{r(c+d)\,|\,r\in\boldsymbol{R}\}$ — прямая, проходящая через точки 0 и $c+d$, и эти прямые пересекаются только в точке $c+d$. Поэтому из условия $c+\frac{\mu}{\lambda}d=r(c+d)$ следует, что $\frac{\mu}{\lambda}=1$ или $\lambda=\mu$ и, следовательно, $\begin{vmatrix} a & b\\ c & d\end{vmatrix}=\begin{vmatrix}\lambda c & \mu d\\ c & d\end{vmatrix}=cd(\lambda-\mu)=0$. Получаем противоречие с условием $ad-bc\neq0$. Итак, числа 0, c, d, а вместе с ними и числа $a=\lambda c$, $b=\mu d$ лежат на одной прямой. Сокращая числитель и знаменатель дроби $\frac{az+b}{cz+d}$ на $c\neq0$, получим $\omega(z)=\frac{a'z+b'}{z+d'}$, где

$$a', b', d'\in\boldsymbol{R}\ \text{и}\ \begin{vmatrix} a' & b'\\ 1 & d'\end{vmatrix}\rightleftarrows t\neq0.$$

Из условия

$$0<(\operatorname{Im}\omega(i))=\operatorname{Im}\frac{a'i+b'}{i+d'}=\operatorname{Im}\frac{(a'i+b')(d'-i)}{1+(d')^2}=\frac{a'd'-b'}{1+(d')^2}$$

следует, что $t=a'd'-b'>0$, и, полагая $\alpha\rightleftarrows\frac{a'}{\sqrt{t}}$, $\beta\rightleftarrows\frac{b'}{\sqrt{t}}$, $\gamma\rightleftarrows\frac{1}{\sqrt{t}}$, $\delta\rightleftarrows\frac{d'}{\sqrt{t}}$, получим для всех $z\in\boldsymbol{C}$ равенство $\omega(t)=\frac{\alpha z+\beta}{\gamma z+\delta}$, где $\alpha, \beta, \gamma, \delta\in\boldsymbol{R}$ и $\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}=\frac{1}{t}\begin{vmatrix} a' & b'\\ 1 & d'\end{vmatrix}=1$.

Случай, когда $c=0$ или $d=0$, а также случай, когда $\omega=\frac{a\bar{z}+b}{c\bar{z}+d}$, разбирается аналогично. ■

Замечание. Если $a, b, c, d\in\boldsymbol{R}$ и $ad-bc>0$, то равенство $\omega(z)\rightleftarrows\frac{az+b}{cz+d}$ также задает движение в H^2P_2, так как

$$\omega(z)=\frac{\frac{a}{\sqrt{t}}z+\frac{b}{\sqrt{t}}}{\frac{c}{\sqrt{t}}z+\frac{d}{\sqrt{t}}}\ \text{и}\ \begin{vmatrix}\frac{a}{\sqrt{t}} & \frac{b}{\sqrt{t}}\\ \frac{c}{\sqrt{t}} & \frac{d}{\sqrt{t}}\end{vmatrix}=1.$$

Аналогично, если $a, b, c, d\in\boldsymbol{R}$ и $-t\rightleftarrows ad-bc<0$, равенство $\omega(z)=\frac{a\bar{z}+b}{c\bar{z}+d}$ также задает движение в H^2P_2.

Теорема 8 позволяет установить тесную связь между группой $\operatorname{Iso}(H^2P_2)$ движений в плоскости Лобачевского и группой L_2, с которой мы уже встречались в главе III. Напомним, что

$$L_2\rightleftarrows\left\{\begin{pmatrix} a & b\\ c & d\end{pmatrix}\,\middle|\, a, b, c, d\in\boldsymbol{R},\ ad-bc=\pm1\right\}$$

и наибольшая группа G_s биекций плоскости $\boldsymbol{R}^2$, относительно которой инвариантна площадь, есть полупрямое произведение группы L_2 на группу $(\boldsymbol{R}^2)^+$ всех параллельных переносов плоскости.

Следствие 1. Группа $\operatorname{Iso}(H^2P_2)$ изоморфна фактор-группе группы L_2 по подгруппе, состоящей из двух элементов $E=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ и $-E=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$.

▷ Пусть $T=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in L_2$. Определим отображение $\psi: L_2\to\operatorname{Iso}(H^2P_2)$ равенством $\psi_T(z)\rightleftarrows\frac{az+b}{cz+d}$, если $ad-bc=1$, и равенством $\psi_T(z)\rightleftarrows\frac{a\bar{z}+b}{c\bar{z}+d}$, если $ad-bc=-1$. Тогда ψ — гомоморфизм (см. упражнение 14) и по теореме 8,б отображение ψ — сюръективно.

Пусть $T\in\operatorname{Ker}\psi$, т. е. $\psi_T=\frac{az+b}{cz+d}=z$ для всех $z\in\boldsymbol{C}$. Тогда при всех $z\in\boldsymbol{C}$, $z\neq -d/c$ выполняется $cz^2+dz=az+d$ или $cz^2+(d-a)z+b=0$, т. е. квадратное уравнение имеет бесконечно много решений. Это возможно только в случае $c=b=0$ и $a=d$. Так как $ad-bc=1$, то $T=E$ или $T=-E$.

Пусть $T=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in L_2$ и $ad-bc=-1$. Допустим, что $T\in\operatorname{Ker}\psi$, т. е. $\psi_T(z)=\frac{a\bar z+b}{c\bar z+d}=z$ для всех $z\in\boldsymbol{C}$. Для вещественных чисел z, т. е. когда $z=\bar z$, получаем, что квадратное уравнение $cz^2+(d-a)z+b=0$ имеет бесконечно много решений: все числа кроме $-\frac{d}{c}$. Поэтому $c=b=0$, $a=d$, т. е. $\psi_T(z)=z$ для всех $z\in\boldsymbol{C}$. В частности $\psi_T(i)=-i\neq i$, т. е. $T\notin\operatorname{Ker}\psi$.

Итак, $\operatorname{Ker}\psi=\{E,\ -E\}$, а группа $\operatorname{Iso}(H^2P_2)$ изоморфна фактор-группе L_2 по ее подгруппе $\operatorname{Ker}\psi$. ■

Следствие 2. Группа движений $\operatorname{Iso}(H^2P_2,\ z_0)$ с неподвижной точкой z_0 изоморфна фактор-группе группы O_2 по подгруппе, состоящей из двух элементов E и $-E$.

▷ Пусть $z_0=i$. Рассмотрим ограничение гомоморфизма $\psi: L_2\to\operatorname{Iso}(H^2P_2)$ (из следствия 1) на подгруппу O_2 группы L_2. Докажем, что $\psi(O_2)=\operatorname{Iso}(H^2P_2,\ i)$. Пусть $\omega(\cdot)\in\operatorname{Iso}(H^2R_2,\ i)$ и $\omega(z)=\frac{az+b}{cz+d}$ для некоторых чисел a, b, c, $d\in\boldsymbol{R}$ и $ad-bc=1$ (см. теорему 8,б).

Тогда $\frac{ai+b}{ci+d}=i\Leftrightarrow ai+b=-c+di\Leftrightarrow a=d,\ b=-c$ и $ad-bc=a^2+b^2=1$, т. е. $a=\cos\varphi$ и $b=\sin\varphi$ для некоторого $\varphi\in[0,\ 2\pi[$. Следовательно, выполняется тождество $\omega(\cdot)\equiv\psi_T(\cdot)$ для $T\rightleftarrows\begin{pmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{pmatrix}\in\mathrm{SO}_2$.

Если $\omega(z)=\frac{a\bar z+b}{c\bar z+d}$, где $ad-bc=-1$, то аналогично из равенства $\omega(i)=i$ получаем тождество $\omega(\cdot)=\psi_T(\cdot)$, где $T\rightleftarrows\begin{pmatrix}\cos\varphi & \sin\varphi\\ \sin\varphi & -\cos\varphi\end{pmatrix}\in O_2\setminus\mathrm{SO}_2$. Так как по следствию 1 $\operatorname{Ker}\psi\subset O_2$, то $\operatorname{Ker}(\psi|\ O_2)=\operatorname{Ker}\psi=\{E,\ -E\}$. Поэтому следствие 2 доказано для числа $z_0=i$.

Пусть $z_0=x_0+iy_0$, $y_0>0$ и $\omega_0(z)\rightleftarrows\frac{x_0z-y_0}{z}$. Тогда $\omega_0(\cdot)\in\operatorname{Iso}(H^2P_2)$, $\omega_0(i)=z_0$ и, как легко проверить, соответствие $\omega\mapsto\omega_0\cdot\omega\cdot\omega_0^{-1}$ является изоморфизмом групп $\operatorname{Iso}(H^2P_2,\ z_0)$ и $\operatorname{Iso}(H^2P_2,\ i)$. ■

Следствие 3. Группа движений $\operatorname{Iso}(H^2P_2,\ z_0)$ с неподвижной точкой z_0 изоморфна O_2.

4. Перейдем теперь к свойствам понятия «конгруэнтность». Мы подробно остановимся на проверке аксиом III группы аксиоматики Евклида — Гильберта во второй модели H^2P_2 Пуанкаре.

Обозначим a_0 h-прямую, проходящую через h-точки i и $\frac{1+i\sqrt3}{2}$, т. е. полуокружность в верхней полуплоскости, проходящую через точки -1, i, 1, а через a_0^+ обозначим пересечение h-прямой a_0 с первой четвертью, т. е. h-луч $[i,\ 1)$.

Предложение 11. ***Для любого числа $d>0$ на h-луче a_0^+ найдется единственная h-точка $z(d)$, для которой***

$$\rho_h(i,\ z(d))=d.$$

▷ По формуле для h-расстояния из предложения 8 получаем для любой точки $z\in a_0^+$

$$\rho_h(i,\ z)=\left|\ln\left(\frac{\operatorname{tg}\frac{\pi}{4}}{\operatorname{tg}\frac{\varphi}{2}}\right)\right|,$$

где φ — величина угла между лучом $[0,\ z)$ и положительным направлением оси абсцисс.

Поэтому

$$(\rho(i,\ z)=d)\Leftrightarrow\left(e^d=\frac{1}{\operatorname{tg}\frac{\varphi}{2}}\right)\Leftrightarrow(\varphi=2\operatorname{arctg}(e^{-d})).$$

Итак, искомая точка $z(d)$ находится по формуле

$$z(d)=\cos(2\operatorname{arctg}(e^{-d}))+i\sin(2\operatorname{arctg}e^{-d})=\frac{1-e^{-2d}}{1+e^{-2d}}+i\frac{2e^{-d}}{1+e^{-2d}}.$$ ■

Теорема 9. ***Два h-отрезка конгруэнтны между собой тогда и только тогда, когда их h-длины равны между собой.***

▷ Так как h-длина h-отрезка равна h-расстоянию между его концами (см. определение 9) и так как при h-движениях h-расстояния сохраняются, то h-длины конгруэнтных h-отрезков равны между собой. Докажем обратное утверждение.

Пусть $[z_1,\ z_2]_h$ — произвольный h-отрезок, h-длина которого равна $d>0$, т. е. $\rho_h(z_1,\ z_2)=d$. Докажем, что он конгруэнтен h-отрезку $[i,\ z(d)]_h$, где $z(d)$ — h-точка на луче a_0^+, построенная в предложении 11.

Рассмотрим сначала случай, когда $\operatorname{Re}z_1\neq\operatorname{Re}z_2$, т. е. когда h-точки z_1 и z_2 не лежат на одной вертикали. Параллельным переносом вдоль оси абсцисс переведем h-прямую $(z_1,\ z_2)$, т. е. полуокружность, перпендикулярную оси абсцисс, проходящую через точки z_1 и z_2, в полуокружность, перпендикулярную оси абсцисс с центром в точке $\bar O$. Затем подходящей гомотетией с центром в точке $\bar O$ переведем последнюю полуокружность в полуокружность a_0. Так как по следствию из предложения 8 указанные отображения верхней полуплоскости в себя являются h-движениями, то, тем самым, на h-прямой a_0 найден h-отрезок $[z_3,\ z_4]_h$, конгруэнтный h-отрезку $[z_1,\ z_2]_h$.

Найдем движение $\omega(\cdot)\in\operatorname{Iso}(H^2P_2)$, переводящее h-точку z_3 в i и

переводящее все точки h-прямой a_0 в точки этой же h-прямой. Для этого нужно записать $\omega(\cdot)$ в виде $\omega(z)=\frac{az+b}{cz+d}$ и решить относительно a, b, c, d систему уравнений:

$$\omega(1)=\frac{a+b}{c+d}=1$$
$$\omega(-1)=\frac{-a+b}{-c+d}=-1$$
$$\omega(z_3)=\frac{az_3+b}{cz_3+b}=i$$

Обозначив $z_3=\cos\varphi+i\sin\varphi$, $0<\varphi<\pi$, получим

$$\omega(z)=\frac{z\cos\varphi+\sin\varphi-1}{z(\sin\varphi-1)+\cos\varphi}$$

Отметим, что $\begin{vmatrix}\cos\varphi & \sin\varphi-1\\ \sin\varphi-1 & \cos\varphi\end{vmatrix}=2\sin\varphi(1-\sin\varphi)>0$, так как $1\geqslant\sin\varphi>0$, и если $\sin\varphi=1$, то $\varphi=\frac{\pi}{2}$, что означает совпадение z_3 с i.

При h-движении $\omega(\cdot)$ h-отрезок $[z_3, z_4]_h$ перейдет в конгруэнтный ему h-отрезок $[i, z]_h$. Если z лежит в первой четверти, то $z=z(d)$. Если z лежит во второй четверти, то при отражении относительно оси ординат, являющемся h-движением, h-отрезок $[i, z]_h$ перейдет в h-отрезок $[i, z(d)]_h$.

Итак, искомое h-движение, переводящее $[z_3, z_4]_h$ на $[i, z(d)]_h$, построено как композиция некоторых h-движений.

Случай, когда $\operatorname{Re}z_1=\operatorname{Re}z_2$, сводится к разобранному случаю указанием произвольного h-движения, переводящего вертикаль $\{z|\operatorname{Re}z=\operatorname{Re}z_1,\ \operatorname{Im}z>0\}$ в какую-нибудь полуокружность в верхней полуплоскости, перпендикулярную оси абсцисс. Таким h-движением является, например, дробно-линейное отображение $\omega(z)=\frac{(\operatorname{Re}z_1)^2}{z}$ при $\operatorname{Re}z_1\neq0$ и дробно-линейное отображение

$$\omega(z)=\frac{z}{z+1}\text{ при }\operatorname{Re}z_1=0$$

Пусть теперь $[z_1, z_2]_h$ и $[z_1', z_2']_h$ — h-отрезки, h-длины которых равны между собой и равны числу $d>0$. По доказанному выше каждый из этих h-отрезков конгруэнтен одному и тому же h-отрезку $[i, z(d)]_h$. Следовательно, по аксиоме III_2 (см. упражнение 11,г) эти h-отрезки конгруэнтны. ■

Следствие 1. В h-плоскости выполняется аксиома III_1.

▷ Достаточно доказать выполнимость аксиомы III_1 во второй модели H^2P_2 Пуанкаре h-планиметрии.

Пусть даны h-прямая a, h-точка z на ней и некоторый h-отрезок h-длины $d>0$. Как и в теореме 9 построим движение $\omega(\cdot)\in\operatorname{Iso}(H^2P_2)$, переводящее h-прямую a в h-прямую a_0, а точку z в

i. При этом для фиксированного на h-прямой a луча с началом z можно считать, что при этом движении данный луч перешел в луч a_0^+. По предложению 11 на луче a_0^+ существует единственная h-точка $z(d)$, для которой

$$\rho(i, z(d))=d$$

Тогда $\omega^{-1}(z(d))$ — та единственная h-точка на фиксированном луче с началом в z, которая удалена от точки z на h-расстояние d. Так как h-длина h-отрезка $[z, \omega^{-1}(z(d))]_h$ равна d и равна h-длине фиксированного h-отрезка, то по теореме 9 эти h-отрезки конгруэнтны. ■

Следствие 2. В h-плоскости выполняется аксиома III_3.

▷ Сначала проверим, что для h-точки z, лежащей между h-точками z_1 и z_2, выполняется

$$\rho(z_1, z_2)=\rho(z_1, z)+\rho(z, z_2)$$

Пусть $\operatorname{Re}z_1\neq\operatorname{Re}z_2$, x_0 — центр полуокружности, перпендикулярной оси абсцисс и проходящей через z_1 и z_2, и φ_1, φ_2, φ — величины углов, образованных лучами $[x_0, z_1)$, $[x_0, z_2)$, $[x_0, z)$ с положительным направлением оси абсцисс.

Пусть для определенности $\varphi_1<\varphi_2$. Тогда $0<\varphi_1<\varphi<\varphi_2<\pi$ и в силу монотонности функций $\operatorname{tg}(\cdot)$ на промежутке $\left(0, \frac{\pi}{2}\right)$ и $\ln(\cdot)$ на луче $(0, +\infty)$ получаем

$$\rho(z_1, z)=\left|\ln\left(\frac{\operatorname{tg}\frac{\varphi}{2}}{\operatorname{tg}\frac{\varphi_1}{2}}\right)\right|=\ln\left(\operatorname{tg}\frac{\varphi}{2}\right)-\ln\left(\operatorname{tg}\frac{\varphi_1}{2}\right)$$

$$\rho(z_2, z)=\left|\ln\left(\frac{\operatorname{tg}\frac{\varphi_2}{2}}{\operatorname{tg}\frac{\varphi}{2}}\right)\right|=\ln\left(\operatorname{tg}\frac{\varphi_2}{2}\right)-\ln\left(\operatorname{tg}\frac{\varphi}{2}\right)$$

$$\rho(z_1, z_2)=\ln\left(\operatorname{tg}\frac{\varphi_2}{2}\right)-\ln\left(\operatorname{tg}\frac{\varphi_1}{2}\right)$$

Следовательно, «равенство треугольника» в этом случае выполняется. Аналогично рассматривается случай $\operatorname{Re}z_1=\operatorname{Re}z_2$.

Пусть теперь h-точка z' лежит между h-точками z_1' и z_2' и h-отрезки $[z_1, z]_h$, $[z, z_2]_h$ конгруэнтны h-отрезкам $[z_1', z']_h$, $[z', z_2']_h$. Тогда по теореме 9

$$\rho(z_1, z)=\rho(z_1', z')\text{ и }\rho(z, z_2)=\rho(z', z_2').$$

Поэтому

$$\rho(z_1, z_2)=\rho(z_1, z)+\rho(z, z_2)=\rho(z_1', z')+\rho(z', z_2')=\rho(z_1', z_2').$$

Следовательно, по теореме 9 h-отрезки $[z_1, z_2]_h$ и $[z_1', z_2']_h$ конгруэнтны. ■

Предложение 12. ***Пусть дан геометрический угол с вер-***

шиной в точке i, одна из сторон которого лежит в первой четверти и горизонтальна, т. е. есть луч $\{z \mid \operatorname{Re} z \geqslant 0,\ \operatorname{Im} z = 1\}$. ***Тогда в фиксированной h-полуплоскости с границей*** a_0 ***существует единственный h-луч с вершиной в точке i, касающийся второй стороны этого геометрического угла.***

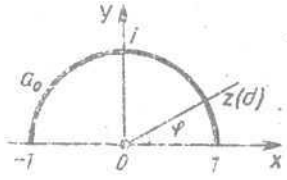


Рис. 131

▷ Если данный геометрический угол прямой, то искомым h-лучом будет являться та часть оси абсцисс, которая лежит в выбранной h-полуплоскости с границей a_0 (рис. 131). Если этот угол не прямой, то проведем из точки i перпендикуляр ко второй стороне этого угла, обозначим z_0 точку его пересечения с осью абсцисс (рис. 132). Тогда искомым h-лучом является та часть полуокружности с центром в точке z_0 и радиусом $\|i,\ z_0\|$, которая лежит в выбранной h-полуплоскости с границей a_0.

Единственность такого h-луча следует из единственности перпендикуляра, проведенного из точки i ко второй стороне угла. ■

Теорема 10. *Два h-угла конгруэнтны между собой тогда и только тогда, когда геометрические углы, образованные касательными, проведенными из вершин этих углов, конгруэнтны как углы в арифметической плоскости* R^2.

▷ Если два h-угла конгруэнтны и $\omega(\cdot)$ — движение в H^2P_2, переводящее один из этих h-углов в другой, то $\omega(\cdot)$ — дробно-линейное отображение первого или второго рода. Поэтому оно сохраняет углы между кривыми, т. е. переводит геометрические углы между касательными, проведенными в точках пересечений кривых, в конгруэнтные геометрические углы.

Рассмотрим любой h-угол с вершиной в h-точке z_0 и фиксируем точки z_1 и z_2 на сторонах этого угла. Пусть $\omega(\cdot)$ — движение в H^2P_2, переводящее z_0 в i и z_1 в $z(d)$, где $d \rightleftarrows \rho_h(z_0, z_1)$. Это h-движение переводит одну из сторон данного h-угла с вершиной в точке z_0 в луч a_0^+. Верхней h-полуплоскостью с границей a_0 назовем множество точек из верхней полуплоскости, удаленных от 0 на расстояние, большее единицы. Если при отображении $\omega(\cdot)$ точка z_2 переходит в нижнюю полуплоскость с границей a_0, то рассмотрим h-движение, при котором точки i и $z(d)$ неподвижны, а нижняя и верхняя h-полуплоскости с границей a_0 поменяются местами. Примером такого h-движения является инверсия с центром в точке 0 относительно единичной окружности.

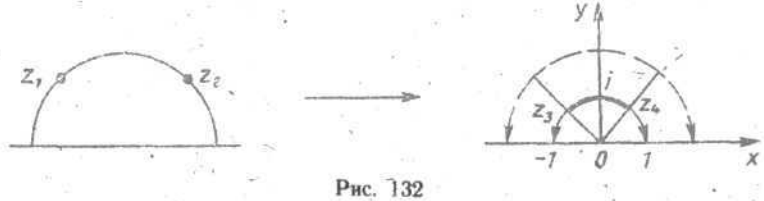


Рис. 132

Композиция этих двух h-движений даст h-движение, при котором h-угол с вершиной в точке z_0 перейдет в конгруэнтный ему h-угол с вершиной в точке i, одна из сторон которого — луч a_0^+, а другая лежит в верхней h-полуплоскости с границей a_0.

Пусть дан h-угол с вершиной в h-точке z_0', геометрический угол которого, образованный касательными к сторонам, конгруэнтен соответствующему геометрическому углу рассмотренного выше h-угла с вершиной в h-точке z_0. Проведя для h-угла с вершиной в z_0' аналогичные построения, получим, что этот угол также конгруэнтен h-углу с вершиной в точке i, одна из сторон которого — луч a_0^+, а другая сторона лежит в верхней полуплоскости с границей a_0. Итак, для двух h-углов с вершинами в точках z_0 и z_0' построены конгруэнтные им h-углы с вершиной в точке i с одной общей стороной — лучом a_0^+. Так как при h-движениях геометрические углы, образованные касательными, переходят в конгруэнтные им углы, то геометрические углы, образованные касательными к сторонам этих двух h-углов с вершиной i, конгруэнтны. По предложению 12 получаем, что эти два h-угла с вершиной в i совпадают между собой. Следовательно, данные h-углы с вершинами в z_0 и z_0' конгруэнтны одному и тому же h-углу с вершиной в i и поэтому конгруэнтны между собой. ■

Следствие 1. В h-плоскости выполняется аксиома III_4.

▷ Пусть дан h-луч с вершиной в точке z_0, фиксирована h-полуплоскость, заданная этим лучом, и дан некоторый h-угол. Аналогично доказательству следствия 1 из теоремы 9, с помощью некоторого h-движения переведем h-луч с вершиной z_0 в h-луч a_0^+ таким образом, чтобы фиксированная h-полуплоскость перешла в верхнюю h-полуплоскость с границей a_0. Затем отложим в верхней h-полуплоскости единственный (по предложению 12) h-луч, образующий с h-лучом a_0^+ h-угол, конгруэнтный данному. С помощью движения $\omega^{-1}(\cdot)$ перенесем этот h-угол в h-угол с вершиной в z_0. Получим h-угол с вершиной в точке z_0, лежащий в фиксированной h-полуплоскости, одна из сторон которого совпадает с данным h-лучом и который конгруэнтен данному h-углу. ■

Следствие 2. В h-плоскости выполняется аксиома III_5.

▷ Пусть для двух h-треугольников $\{z_0, z_1, z_2\}$ и $\{z_0', z_1', z_2'\}$ выполняется $\rho_h(z_0, z_1) = \rho_h(z_0', z_1') = d_1$, $\rho_h(z_0, z_2) = \rho_h(z_0', z_2')$ и h-углы $\angle z_1z_0z_2$ и $\angle z_1'z_0'z_2'$ конгруэнтны. Рассмотрим h-движения $\omega(\cdot)$ и $\omega'(\cdot)$, при которых

$$\omega(z_0) = \omega'(z_0') = i,\ \omega(z_1) = \omega'(z_1') = z(d_1) \text{ и } \omega(z_2),\ \omega'(z_2')$$

лежат в верхней h-полуплоскости с границей a_0. Так как в полуплоскости $\{z \mid \operatorname{Im} z > 1\}$ от луча $\{z \mid \operatorname{Re} z \geqslant 0,\ \operatorname{Im} z = 1\}$ можно отложить ровно один геометрический угол, конгруэнтный данному (аксиома III_4 для евклидовой плоскости), то, используя предложение 12, получим, что точка $\omega(z_2)$ и $\omega'(z_2')$ лежат на одном и том же h-луче с вершиной в точке i.

Так как $\rho_h(\omega(z_2),\ i) = \rho_h(z_2,\ z_0) = \rho_h(z_2',\ z_0') = \rho_h(\omega'(z_2',\ i))$, то по

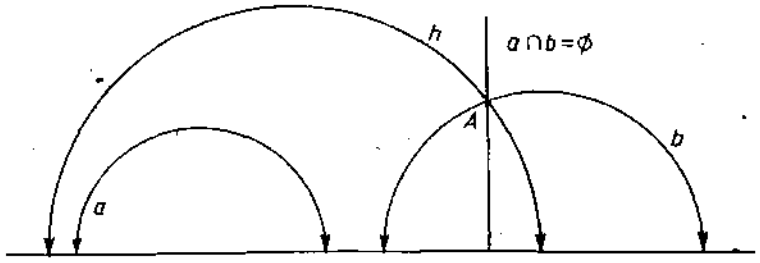


Рис. 133

аксиоме III_1 (следствие 1 из теоремы 9) получим, что $\omega(z_2) = \omega'(z_2')$. Следовательно, оба h-треугольника при h-движениях ω и ω' перейдут в один и тот же h-треугольник $\{i, z(d), \omega(z_2)\}$. Поэтому h-угол $\angle z_0 z z_2$ при h-движении $(\omega')^{-1} \circ \omega$ перейдет в h-угол $\angle z_0' z' z_2'$. Следовательно, эти h-углы конгруэнтны. ■

Упражнение 15. *Докажите, что:*

а) в модели Пуанкаре H^2P_2 плоскости Лобачевского выполняются аксиомы Архимеда и Кантора;

б) в модели Пуанкаре H^2P_2 не выполняется аксиома параллельности (см. рис. 133);

в) в сферической и в эллиптической плоскостях выполняется аксиома Кантора (отметим, что аксиома Архимеда в этих плоскостях не имеет смысла, так как в них отсутствует понятие «между»);

г) в сферической и в эллиптической плоскостях не выполняется аксиома IV (аксиома параллельности), но выполняется аксиома IV,б: любые две различные прямые пересекаются между собой.

1 р. 10 к.