
ШКОЛА ОПОЙЦЕВА

**МАТЕМАТИЧЕСКИЙ
АНАЛИЗ**



URSS

МОСКВА

Опойцев Валерий Иванович

Школа Опойцева: Математический анализ.

М.: ЛЕНАНД, 2016. — 272 с.

Стандартная программа математического анализа (производные, интегралы) расширена добавлением ингредиентов из других дисциплин, чем достигается цельность предмета. Яснее становится «что для чего нужно». При этом изложение отличается краткостью и прозрачностью, вплоть до объяснений на пальцах. Акцент делается на понимании существа дела, причем с заботой о новичках, знакомящихся с предметом либо впервые, либо — по второму кругу, после не вполне удачного первого. Охват материала достаточно широкий, но изложение построено так, что можно ограничиться любым желаемым срезом содержания. Значительное внимание уделяется мотивации результатов. Объяснения даются «человеческим языком». Книга легко читается, самодостаточна и может служить основой при изучении матанализа.

Для студентов, преподавателей, инженеров и научных работников.

Формат 60×90/16. Печ. л. 17. Зак. № АХ-710.

Отпечатано в ООО «ЛЕНАНД».

117312, Москва, пр-т Шестидесятилетия Октября, 11А, стр. 11.

ISBN 978-5-9710-3295-3

© ЛЕНАНД, 2016

19529 ID 213615



9 785971 032953



Все права защищены. Никакая часть настоящей книги не может быть воспроизведена или передана в какой бы то ни было форме и какими бы то ни было средствами, будь то электронные или механические, включая фотокопирование и запись на магнитный носитель, а также размещение в Интернете, если на то нет письменного разрешения владельца.

Оглавление

Предисловие	7
1 Стартовая площадка	8
1.1 Откуда берутся «всякие» числа	9
1.2 Всё ли так просто	10
1.3 Комплексные числа	12
1.4 Об ощущении таинственности	16
1.5 Комбинаторика	18
1.6 Бином Ньютона	20
1.7 Многочлены	21
1.8 Теоретико-множественные операции	23
2 Последовательности и пределы	26
2.1 Сходимость и пределы	26
2.2 Простейшие инструменты	28
2.3 Как работает лемма Вейерштрасса	30
2.4 Предел функции	32
2.5 Непрерывные функции	35
2.6 О теории вещественных чисел	37
2.7 Надумана ли проблема и каковы блага	40
2.8 Равномерная непрерывность	42
2.9 Фундаментальные последовательности	43
2.10 Числовые ряды	44
3 Производная и дифференциал	51
3.1 Производная	51
3.2 Правила дифференцирования	53
3.3 Дифференциалы	55

3.4	Производные элементарных функций	59
3.5	Тропа на вершину Тейлора ,	64
3.6	Разложение Тейлора	67
3.7	Контрпримеры и парадоксы	69
4	Интеграл	75
4.1	Первообразная	75
4.2	Определённый интеграл	77
4.3	Взаимосвязь интегралов	79
4.4	Техника интегрирования	81
4.5	Прикладные задачи	85
4.6	Несобственные интегралы	92
4.7	Дифференциальные уравнения	98
5	Функции нескольких переменных	105
5.1	В двух словах о векторах	105
5.2	Предел и сходимость	107
5.3	Непрерывность	109
5.4	Частные производные	110
5.5	Приращения и дифференциалы	110
5.6	Производные и дифференциалы высших порядков	112
5.7	Градиент	113
5.8	О роли повторных пределов	117
5.9	Интегралы, зависящие от параметра	122
6	Отображения, или операторы	125
6.1	Аргументы и функции как векторы	125
6.2	Линейные отображения	127
6.3	Обратимые преобразования	130
6.4	Детерминанты, или определители	132
6.5	Эквивалентные нормы	133
6.6	Дифференцирование оператора	136
6.7	Принцип сжимающих отображений	137
6.8	Обратные и неявные функции	138

7	Кратные интегралы	141
7.1	Двойные интегралы	141
7.2	Замена переменных	144
7.3	Кратные интегралы	146
7.4	Объёмы n -мерных тел	146
7.5	Сюрпризы n измерений	148
8	Векторный анализ	150
8.1	Координаты и векторы	150
8.2	Скалярное произведение	155
8.3	Векторное произведение	157
8.4	Приложения к механике	160
8.5	Дивергенция	163
8.6	Оператор Гамильтона	167
8.7	Циркуляция	168
9	Гладкая оптимизация	173
9.1	Безусловный экстремум	173
9.2	Достаточные условия	175
9.3	Условный экстремум	176
9.4	Общий случай	180
9.5	Нелинейное программирование	182
9.6	Интерпретация множителей Лагранжа	184
9.7	Двойственные задачи	185
10	Аналитические функции	187
10.1	О роли комплексных чисел	187
10.2	Дифференцируемость	190
10.3	Примеры	193
10.4	Простейшие свойства	194
10.5	Контурные интегралы	196
10.6	Интеграл Коши	200
10.7	Регулярность	202
10.8	Аналитическое продолжение	205
10.9	Теорема Руше	207

11	Функциональные ряды	210
11.1	Равномерная сходимость	211
11.2	Степенные ряды	212
11.3	Ортогональные разложения	214
11.4	Механизм производящих функций	217
11.5	Ряды Фурье	218
11.6	Интеграл Фурье	221
11.7	Преобразование Лапласа	223
11.8	Дельта-функция	227
12	Неподвижные точки	231
12.1	Проблемы существования решения	231
12.2	Вращение векторного поля	234
12.3	Гомотопия векторных полей	237
12.4	Ядро теории	240
12.5	Теоремы существования	243
12.6	О теореме Брауэра	245
12.7	P -отображения	246
12.8	Алгебраическое число нулей	247
12.9	Индексы на бесконечности	248
12.10	Накрытия и гомеоморфизмы	249
12.11	Параметрические уравнения	251
13	Проблемы обучения	253
13.1	Кто мы есть и как мы учимся	254
13.2	О взаимодействии с подсознанием	257
13.3	Гипнотический вирус, будь он неладен	258
	Обозначения	261
	Литература	262
	Предметный указатель	264

Предисловие

*Нельзя по дороге в баню
отлежаться на поездку в Сочи.*

Как учат там, где система образования замешкалась в стороне? Очень просто. Чтобы краска на сапоги не капала, кисточку держи вот так. Потом — что-то насчёт колера. Далее разные «пустячки» уточняются, расширяются, сливаются, — и вырастает маляр. Или, скажем, кого-то взяли в штат. Вот тебе рабочее место, тут буфет, там туалет, а это твои коллеги, Саша-Маша, что не ясно — спрашивай. Всё! Нормальная стартовая площадка. Остальное приложится.

Это хороший путь, правильный. Эффективный для освоения любой науки. Ибо всякое познание начинается с середины, с некой данности, с кога в мешке, а потом развивается по двум направлениям, вглубь и вширь. Так и надо учить и учиться.



Но *Высшая школа* попала в западню. Группа её резидентов двигалась «вглубь и вширь», потом где-то остановилась и решила, что с этого места и надо обучать население. Мол, так будет лучше. В результате никто в изучаемый предмет *войти не может*. Потому что вход там, где богом задумано, а не там, куда резиденты долго шли, а новичок не может попасть даже по шучьему велению.



Данный курс почти везде начинает с «середины». Многое поясняя по ходу дела. Не до такой степени, конечно, что «вот вам буфет-туалет». Поскольку слегка притормаживает, чтобы не перегнуть палку. Но сдвиг в направлении «идёшь в баню — иди» давно назрел. Чем меньше приготовлений, тем успешнее движение.

Глава 1

Стартовая площадка

*Одoleвает камуфляж,
Блажь!
И страх под горлом начеку —
Страж!
А космос в штопор и Мираж
В раж!
Ну, вмажь разок хотя бы, а,
Вмажь!*

В главе приводятся некоторые сведения из элементарной математики. Частично как напоминание, но в большей степени для пастройки в резонанс. Особенно это касается *комплексных чисел*, которые на полную мощь работают в главе 10, а по ходу дела встречаются лишь изредка. Но они являются великолепным камертоном для ощущения полёта орла, необходимого при изучении высшей математики. Не говоря о том, что странствие в эмпиреях *комплексных чисел* неожиданным образом открывает глаза на природу «обыкновенных» чисел.



Рисунки А. Фоменко

1.1 Откуда берутся «всякие» числа

Более высокие ступени математического образования характеризуются не столько расширением, сколько углублением обзора, — позволяя видеть новое в давно известном. Хотя источники появления чисел теряются в начальном образовании, там есть один принципиальный механизм, который лучше видится на расстоянии. Как и зачем, скажем, вводятся отрицательные числа?

Тут легко заикнуться на пустяковых причинах. Дескать, движение вправо будем помечать знаком «+», влево — «-», и тогда «+а» шагов вправо, «-b» влево дают результирующее местоположение $a - b = (+a) + (-b)$, которое при $b > a$ — отрицательно. Оказывается удобно. Объяснение возможное, но от него круги по воде не расходятся. Настоящая роль отрицательных чисел остаётся вне поля зрения.

Главный механизм здесь в другом. После того как некоторое время вода толчётся в ступе вокруг *натурального ряда*



$$\mathbb{N} = \{1, 2, 3, \dots, n, \dots\}, \quad (1.1)$$

начинают вводить *арифметические операции — действия*. Изучают их свойства¹, перенося внимание с существительных (1.1) на глаголы: «+», «-», «×», «÷». При этом *игровая площадка* \mathbb{N} раздражает своей ущербностью. Не всегда действия на ней выполнимы. Тогда возникает идея расширить площадку (1.1), что «под влиянием вычитания» порождает *множество целых чисел*

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

на коем сложение и вычитание никогда не уведут результат за пределы площадки, на которой идёт игра².

Таким образом, если $a, b \in \mathbb{Z}$, то и $a \pm b \in \mathbb{Z}$. Другими словами, складывать и вычитать на \mathbb{Z} можно без опасений, что результат $a \pm b$ уйдёт за пределы \mathbb{Z} . Т. е. \mathbb{Z} оказывается замкнутой игровой площадкой, что даёт образец, стереотип мышления, действуя по которому вводят другие числа и множества.

¹От перестановки слагаемых сумма не меняется и т. д.

²Соответственно, решаются любые уравнения $a + x = b$.

Эта идея пронизывает всю математику. В том смысле, что площадка, *модель*, на которой выполняются операции, должна быть удобной, чтобы действовать, не заботясь о последствиях. Поэтому *модель* надо расширять до тех пор, пока она не *замкнётся по выполняемым операциям*. Причём удобства здесь — даже не самое главное, но об этом позже.

Аналогично \mathbb{Z} появляется в арифметике множество *рациональных чисел* (дробей), обозначаемое символом \mathbb{Q} . Чтобы деление выполнялось и уравнения $a \cdot x = b$ решались³. Существенно, что $\frac{b}{a}$ в итоге оказывается определено и для дробных a и b . Могло быть хуже. Введение дробей $\frac{b}{a}$ для целых a и b , могло бы не обеспечить $\frac{b}{a} \in \mathbb{Q}$ для дробных a и b . И это в некотором роде была бы катастрофа. Пришлось бы вводить ещё дроби дробей и т. д. Слава богу — не потребовалось. К тому же, сумма и произведение дробей — тоже дроби, т. е.

$$\alpha \frac{b}{a} + \beta \frac{p}{q} \quad \text{👉👈}$$

дробь в случае дробных α и β , в том числе отрицательных. Иначе говоря, множество *рациональных чисел* \mathbb{Q} замкнуто в арифметике, т. е. относительно арифметических операций.

1.2 Всё ли так просто

Так ли просто расширяются *игровые площадки*? Когда как. Отрицательные числа получились в результате присоединения к \mathbb{N} всех разностей $a - b$ для всех $a, b \in \mathbb{N}$ с последующим отождествлением одинаковых разностей, и проверкой, что вычитание для $a, b \in \mathbb{Z}$ уже не приводит к дальнейшему расширению модели.

Но если бы мы захотели бесхитростно добавить к \mathbb{N} квадратные корни \sqrt{b} — возникла бы проблема. Во-первых, не все корни из $b \in \mathbb{N}$ рациональны⁴. Во-вторых, добавление всех \sqrt{b} к \mathbb{N} не образует замкну-

³Запрет $a \neq 0$ в $x = \frac{b}{a}$ отчасти портит вкус победы. Но не физикам. Они, как хорошие люди, при малой выгоде большой подлости, конечно, не сделают, но выгода от деления на ноль у них бывает слишком велика.

⁴Скажем, из $\sqrt{5} = \frac{b}{a}$ следует $b^2 = 5a^2$, откуда вытекает, что b^2 делится на 25, а значит a делится на 5, что противоречит несократимости $\frac{b}{a}$.

Далее рассмотрим числовой ряд a_n ,

$$1, 1, 0, -2, -4, -4, 0, 8, 16, 16, 0, -32, \dots, \quad (1.5)$$

устроенный по правилу

$$a_{n+2} = 2a_{n+1} - 2a_n \quad (1.6)$$

со стартовыми условиями $a_0 = a_1 = 1$. Подставляя $a_n = x^n$ в (1.6), приходим к квадратному уравнению

$$x^2 - 2x + 2 = 0, \quad (1.7)$$

у которого дискриминант отрицателен, и потому вещественных корней нет. Попробуем всё же воспользоваться стандартными формулами корней чисто формально. Корни (1.7) оказываются «странными»

$$x_1 = 1 + \sqrt{-1}, \quad x_2 = 1 - \sqrt{-1},$$

и формула для a_n при условии $a_0 = a_1 = 1$ получается такой⁷



$$a_n = \frac{(1 + \sqrt{-1})^n + (1 - \sqrt{-1})^n}{2}. \quad (1.8)$$

Любопытно, что все «нелегальные» $\sqrt{-1}$ в нечётных степенях здесь взаимно уничтожаются⁸, и a_n получаются действительными, причём в точности реализующими последовательность (1.5). Загадочная картина возникает. Манипуляции незаконны, а результат — правильный.

Такого рода чудеса должны мобилизовать внимание. Где-то поблизости может таиться открытие. Здесь тот самый случай.

1.3 Комплексные числа

Попробуем расширить теперь числовую ось \mathbb{R} в нетривиальном направлении. Уравнение $x^2 = b$ при $b < 0$ не решается. Но что

⁷Точнее, $a_n = \frac{(1+i)^n + (1-i)^n}{2}$, где $i^2 = -1$, откуда $a_n = (\sqrt{2})^n \cos \frac{n\pi}{4}$ с учётом тригонометрических ухищрений, см. далее.

⁸По формуле *бинома Ньютона*.

мешает нам объявить, что некое i при возведении в квадрат даёт -1 . Тогда решением $x^2 = -1$ будет $x_{1,2} = \pm i$.

На этом сказка, понятное дело, не кончается. Замкнутость в арифметике требует ещё приплюсовать к \mathbb{R} числа вида⁹ $x + iy$, $x, y \in \mathbb{R}$, которые и называют *комплексными*. Их множество \mathbb{C} называют также *комплексной плоскостью*. Не факт, конечно, что этого хватает. Но этого хватает, как потом выясняется, и для деления¹⁰, и для извлечения корней,

$$\sqrt{i} = \pm \frac{(1+i)}{\sqrt{2}},$$



и даже для извлечения корней любой степени. К тому же решаться начинают ещё и алгебраические уравнения любой степени, не только квадратные! Что свидетельствует о большой удаче и точном попадании.

Комплексные числа, надо согласиться, при беглом знакомстве уведут почву из-под ног. Однако за удивительным фасадом здесь прячется вполне рациональная идея. Где остановиться, — каждый решает сам. Но в любом случае начало пути пролегает через обыкновенное знакомство с объектом.

1.3.1 Определение. *Комплексными числами (КЧ) называются числа вида*

$$z = x + iy,$$

где x, y — обыкновенные вещественные числа, а i — так называемая мнимая единица, $i^2 = -1$. Величину x называют действительной частью, y — мнимой, и пишут

$$x = \operatorname{Re}(z), \quad y = \operatorname{Im}(z)$$

⁹Дабы всё можно было складывать и перемножать.

¹⁰Опять-таки со скидкой на запрет деления на ноль.

Операции сложения и вычитания определяются покомпонентным сложением и вычитанием:

$$z_1 \pm z_2 = (x_1 \pm x_2) + i(y_1 \pm y_2)$$

Два КЧ считаются равными, когда равны их действительные и мнимые части. Понятия *больше/меньше* для КЧ не определены. Правило умножения получается обыкновенным раскрытием скобок. С учётом $i^2 = -1$, это даёт

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + y_1 x_2).$$

Легко проверяется наличие стандартных свойств умножения.

Для деления используется несложный трюк избавления от мнимой единицы в знаменателе, опирающийся на факт вещественности произведения

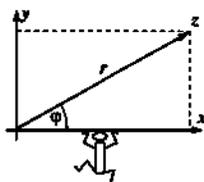
$$z z^* = (x + iy)(x - iy) = x^2 + y^2,$$

где $z^* = x - iy$ называют *сопряжённым числом*. В результате

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*}$$

избавляет от мнимой единицы в знаменателе и сводит деление к умножению $z_1 z_2^*$ в числителе.

Указанные способы умножения и деления на практике используются редко, поскольку есть гораздо более эффективные приемы, основанные на геометрическом представлении КЧ: числу $z = x + iy$ сопоставляется вектор на плоскости $z = \{x, y\}$ (рис. справа). Вот вам и воплощение фикции.



В полярных координатах, $r = |z| = \sqrt{x^2 + y^2}$ — модуль z , φ — аргумент z , в силу

$$x = r \cos \varphi, \quad y = r \sin \varphi,$$

КЧ z записывается в *тригонометрической форме*¹¹

$$z = r(\cos \varphi + i \sin \varphi).$$

Неожиданно обнаруживается, что «неуклюжее» умножение имеет прозрачный геометрический смысл. При умножении z_1 и z_2 модули перемножаются, аргументы складываются. Формула

$$z_1 z_2 = r_1 r_2 [\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)]. \quad (1.9)$$

элементарно проверяется¹², как и формула деления:

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} [\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)]$$

В тригонометрической форме удобно извлекать корни. При $k = 0, 1, \dots, n - 1$ получаются n различных корней n -й степени

$$\sqrt[n]{z} = \sqrt[n]{r} \left(\cos \frac{\varphi + 2k\pi}{n} + i \sin \frac{\varphi + 2k\pi}{n} \right),$$

что проверяется обратным возведением в n -ю степень по очевидной, в силу (1.9), *формуле Муавра*

$$z^n = r^n (\cos n\varphi + i \sin n\varphi)$$

и с учётом периодичности синуса и косинуса. Ко всей этой внутренней кухне мы ещё вернёмся.

Отметим также *формулу Эйлера*

$$e^{i\varphi} = \cos \varphi + i \sin \varphi,$$



истинная природа которой будет выявлена в дальнейшем.

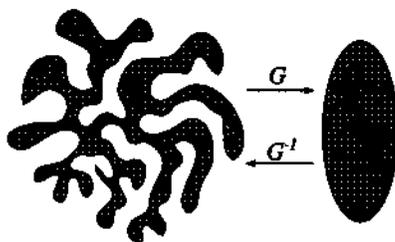
¹¹В силу периодичности тригонометрических функций аргумент φ , вообще говоря, многозначен:

$$\varphi = \text{Arg} z = \{\arg z + 2k\pi : k \in \mathbb{Z}\}.$$

¹²Опорой на формулы синуса и косинуса суммы двух углов.

Удача с находкой *геометрической/тригонометрической формы* комплексного числа — далеко не рядовое событие. Это пример чуда, которое происходит вот по такой схеме.

Большие и малые разделы математики с той или иной долей натяжки можно себе представлять как изучение некоторого множества A объектов с определёнными в этом множестве операциями. Какие-то операции выполняются легко, какие-то — трудно. Естественно выглядит попытка установить взаимнооднозначное соответствие A с каким-либо другим множеством B и посмотреть, какие манипуляции в B соответствуют операциям в A . Если действия в B проще операций в A , то задачи можно решать по схеме, изображенной на рисунке ниже.



Объекты из A трансформируются в B с помощью преобразования G , там выполняются необходимые действия, и результат возвращается в A . Удачный выбор B — всегда событие.

1.4 Об ощущении таинственности

Комплексные числа существуют во Вселенной сами по себе или это мы придумали их для своего ментального удобства? Первая точка зрения облегчает дорогу в сумасшедший дом, где закончили жизненный путь многие творцы *Теории множеств*. Поэтому вторая точка зрения предпочтительна, и мы до поры до времени будем её придерживаться.

«Магическая природа» комплексных чисел даёт о себе знать интуитивно неожиданным образом. Ситуация на *действительной прямой* странным образом почти всегда оказывается зависящей от того, как соответствующие задачи инсталлированы в

комплексную плоскость. Линейные дифференциальные уравнения при попытке ограничиться рассмотрением \mathbb{R}^n из красивой теории превращаются в неудобный клубок загадочных фактов. Электротехника без КЧ блекнет и теряет логическую основу. Трансформаторы, разумеется, продолжают работать, но причины как-то смазываются. Если \mathbb{R} не расширить до \mathbb{C} , почему-то невозможно разобраться в механизме, управляющем сходимостью степенных рядов. Ряд

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - \dots$$

сходится только при $|x| < 1$, а причина на территории \mathbb{R} выглядит странно. У $\frac{1}{1+x^2}$ есть, видите ли, особая точка $x = i$. Но с какой стати точка $i \in \mathbb{C}$ управляет сходимостью ряда на \mathbb{R} ?

В задаче нет даже намека на комплексные числа. Тем не менее, именно на просторах \mathbb{C} достигается ясность. Возникает впечатление, что вспомогательный прием затрагивает какой-то глубинный механизм, работающий на «другом» топливе.

Такого сорта эмоции были широко распространены в «доалгебраические» времена, когда понятия *групп*, *колец* и *полей* только внедрялись в сознание. Когда ещё не сформировалась привычка к абстрактному мышлению алгебраического толка, и уже всё было, но в то же время, ничего не было. Когда изучение действий в отрыве от объектов приложения казалось противоестественным.

Однако постепенно идеология модельных реализаций вошла в плоть и кровь, и свободомыслие вторглось в мир «чисел». Стало ясно, что обычная единица ничуть не реальнее мнимой. В вопросе «что первично, курица или яйцо» тоже произошёл переворот. Главенствующую роль взяли на себя операции, действия. Числа стали вторичны. Для выполнимости действий требовалось лишь подходящее игровое поле. Поначалу думалось — неважно какое. Потом выяснилось — вариант один: \mathbb{C} .

Иначе говоря, арифметические операции порождают цепную реакцию. Из натурального ряда возникают отрицательные числа, потом дроби, затем мнимая единица и, наконец, комплексные числа. На этом *единственно возможное* [2, т. 8] *расширение* игрового поля заканчивается. Таким образом, комплексную плоскость \mathbb{C} , как неизбежный финал роста натурального ряда, порождает арифметика, на которой стоит вся остальная математика. Именно поэтому КЧ дают о себе знать практически в любой области.

1.5 Комбинаторика

1.5.1 Размещения. Число различных вариантов выбора (с учётом порядка) m предметов из n предметов a_1, a_2, \dots, a_n равно

$$A_n^m = n(n-1) \dots (n-m+1)$$

Действительно, есть n способов выбрать один предмет, т. е. $A_n^1 = n$. На каждый выбор первого предмета существует $n-1$ способов выбора второго (из оставшихся $n-1$ предметов) — поэтому $A_n^2 = n(n-1)$. И так далее.

1.5.2 Перестановки. Число всевозможных перестановок n предметов a_1, a_2, \dots, a_n равно « n факториал»

$$n! = 1 \cdot 2 \cdot \dots \cdot n$$

По соображениям удобства принимается $0! = 1$.

Очевидно, перестановка получается при размещении *всех* предметов, что и даёт $A_n^n = n!$

1.5.3 Сочетания. Если m предметов из a_1, a_2, \dots, a_n выбираются без учёта порядка (складываются в мешок), то число различных вариантов (число сочетаний из n по m) равно

$$C_n^m = \frac{n!}{m!(n-m)!}$$

Всевозможные *размещения* получаются перестановками элементов в сочетаниях. Поэтому

$$A_n^m = C_n^m m!,$$

что даёт формулу для C_n^m , с учётом того, что $A_n^m = n!/(n-m)!$

1.5.4 Перестановки с повторениями. Пусть имеется n предметов k типов

$$\underbrace{a_1 \dots a_1}_{n_1} \underbrace{a_2 \dots a_2}_{n_2} \dots \underbrace{a_k \dots a_k}_{n_k},$$

где

$$n_1 + \dots + n_k = n.$$

Число различных перестановок этих предметов равно

$$P(n_1, n_2, \dots, n_k) = \frac{n!}{n_1! n_2! \dots n_k!}$$

В любой перестановке рассматриваемой совокупности предметов, *ничего внешне не меняя*, можно n_1 элементов a_1 переставить между собой $n_1!$ способами, n_2 элементов a_2 — $n_2!$ способами, ..., n_k элементов a_k — $n_k!$ способами. Поэтому $n_1! n_2! \dots n_k!$ перестановок из $n!$ — неотличимы друг от друга, что приводит к указанной формуле.

Рассмотрим, наконец, ещё одну типичную ситуацию. Имеется k типов предметов, число образцов каждого типа — бесконечно. Число различных способов выбора r предметов в данном случае

$$U_k^r = k^r$$

Стандартный пример — десять (типов) цифр, каждую из которых при записи числа можно использовать в любом количестве экземпляров (шестизначных чисел — миллион, 10^6).

Упражнения

1. Доказать:

- $C_n^m = C_n^{n-m}$,
- $C_n^k + C_n^{k+1} = C_{n+1}^{k+1}$,

- $\frac{n+1}{k+1}C_n^k = C_{n+1}^{k+1}$,
- $kC_n^k = nC_{n-1}^{k-1}$.

2. Сколько различных чисел можно получить перестановкой цифр 1, 3, 7? ($3!$)

3. Сколько есть трехзначных чисел, в записи которых участвуют только цифры 1, 3, 7? (3^3)

4. Сколько есть различных чисел, в записи которых участвуют 1, 1, 2, — т. е. две единицы и одна двойка? (3)

1.6 Бином Ньютона

При перемножении n сомножителей

$$(x+y)(x+y)\dots(x+y)$$

число членов вида $x^{n-k}y^k$ равно C_n^k , поскольку k штук y в n сомножителях можно выбрать числом способов C_n^k . Поэтому

$$(x+y)^n = x^n + C_n^1 x^{n-1}y + C_n^2 x^{n-2}y^2 + \dots + C_n^{n-1}xy^{n-1} + y^n$$

Это формула *бинома Ньютона*, которая часто используется.

Полагая $x = y = 1$, получаем

$$C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n = 2^n.$$

В случае $x = 1$, $y = -1$, имеем

$$C_n^0 - C_n^1 + C_n^2 - \dots + (-1)^n C_n^n = 0.$$

Упражнение

$$(x+y+z)^n = \sum P(k_1, k_2, k_3) x^{k_1} y^{k_2} z^{k_3},$$

где суммирование идет по k_1, k_2, k_3 , удовлетворяющим условию $k_1 + k_2 + k_3 = n$.

1.7 Многочлены

Многочлены важны по двум причинам. С одной стороны, они широко используются, с другой — служат хорошим тренажером, вырабатывая полезные навыки «общематематического толка».

Время от времени приходится иметь дело с делением многочленов «в столбик», например,

$$\begin{array}{r}
 x^3 + 2x^2 - x + 3 \quad |x - 1 \\
 \underline{x^3 - x^2} \quad x^2 + 3x + 2 \\
 3x^2 - x \\
 \underline{3x^2 - 3x} \\
 2x + 3 \\
 \underline{2x - 2} \\
 5
 \end{array}$$



Произведённое деление даёт тождество

$$x^3 + 2x^2 - x + 3 = (x - 1)(x^2 + 3x + 2) + 5.$$

Легко видеть, что в общем случае деление многочлена

$$P_n(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 \quad (1.10)$$

на $(x - c)$ даст в частном некоторый многочлен $Q_{n-1}(x)$ и некоторое число R в остатке, т. е.

$$\begin{array}{r}
 P_n(x) \quad |x - c \\
 \dots \quad Q_{n-1}(x) \\
 R,
 \end{array}$$

что равносильно тождеству

$$P_n(x) = (x - c)Q_{n-1}(x) + R,$$

полагая в котором $x = c$, получаем простой, но важный результат:

1.7.1 Теорема Безу. Остаток R при делении $P_n(x)$ на $(x - c)$ равен $P_n(c)$, т. е.

$$R = P_n(c)$$

Таким образом, если c корень уравнения $P_n(x) = 0$, то $R = 0$. В конечном итоге это соображение приводит к разложению

$$P_n(x) = (x - x_1)(x - x_2) \dots (x - x_n), \quad (1.11)$$

где x_1, \dots, x_n — корни многочлена $P_n(x)$, которые по *основной теореме алгебры* 10.9.2 всегда существуют — по крайней мере, в комплексной плоскости.

Примеры

1. Найти остаток от деления $P_n(x)$ на $x^2 + px + q = (x - a)(x - b)$.

Очевидно, $P_n(x)$ при делении на квадратный трехчлен даст в остатке (в общем случае) многочлен первой степени $\gamma x + \delta$, т. е.

$$P_n(x) = (x - a)(x - b)Q_{n-2}(x) + \gamma x + \delta. \quad (1.12)$$

Подставляя в (1.12) сначала $x = a$, потом $x = b$, получаем систему уравнений

$$\begin{cases} P_n(a) = \gamma a + \delta \\ P_n(b) = \gamma b + \delta \end{cases}$$

для определения γ и δ .

2. Показать, что $x^m - 1$ делится на $x^n - 1$ лишь в том случае, когда m делится на n .

Пусть $m = nk + p$, тогда

$$\frac{x^m - 1}{x^n - 1} = x^p \frac{(x^n)^k - 1}{x^n - 1} + \frac{x^p - 1}{x^n - 1}.$$

Далее надо учесть, что $y^k - 1$ всегда делится на $y - 1$,

$$\frac{y^k - 1}{y - 1} = 1 + y + \dots + y^{k-1}.$$

3. Часто встречается разложение на множители следующего многочлена трёх переменных:

$$\begin{aligned} x^3 + y^3 + z^3 - 3xyz &= \\ &= x^3 + 3xy(x + y) + y^3 + z^3 - 3xyz - 3xy(x + y) = \\ &= (x + y)^3 + z^3 - 3xy(x + y + z) = \\ &= \frac{1}{2}(x + y + z)[(x - y)^2 + (x - z)^2 + (y - z)^2]. \end{aligned}$$

4. Раскрытие скобок и приведение подобных в (1.11), с последующим сопоставлением результата с (1.10), даёт теорему Виета для многочлена n -й степени:

$$\begin{array}{l} x_1 + \dots + x_n = -a_{n-1} \\ \sum_{i,j} x_i x_j = a_{n-2} \\ \dots\dots\dots \\ x_1 \dots x_n = (-1)^n a_0 \end{array}$$

1.8 Теоретико-множественные операции

В пределах стандартных обозначений всегда есть люфт. Поэтому об исходных положениях необходимо договориться.

$x \in X$ ($x \notin X$) — элемент x принадлежит (не принадлежит) множеству X . Включение « A является подмножеством множества B » записывается как¹³ $A \subset B$.

$A \cup B$, $A \cap B$, $A \setminus B$ — *объединение*, *пересечение* и *разность множеств*. Используются также эквивалентные обозначения:

$$A + B = A \cup B, \quad A \cdot B = A \cap B, \quad A - B = A \setminus B.$$

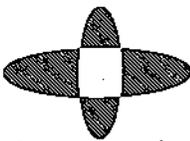
Визуальные пояснения



способствуют наглядному пониманию и запоминанию¹⁴.

¹³ Включение $X \subset Y$, « X подмножество Y », используется далее как эквивалент $X \subseteq Y$. Иначе говоря, у нас $X \subset Y$ означает $X \subseteq Y$. При необходимости оговорить *строгое включение* к $X \subset Y$ добавляется $X \neq Y$.

¹⁴ Идея Гельвеция насчёт знания общих принципов, возмещающего незнание отдельных фактов, — в определённой степени обратима.

Симметрическая разность $A \Delta B$, , состоит

из элементов, которые принадлежат либо только A , либо только B . Операция Δ является коммутативной, ассоциативной и удовлетворяет следующему дистрибутивному закону¹⁵:

$$A \cdot (B \Delta C) = (A \cdot B) \Delta (A \cdot C).$$

$X' = E \setminus X$ — дополнение множества X , определяемое как множество элементов некоторого основного множества E , не принадлежащих X .

$X \times Y$ обозначает декартово произведение множеств X и Y , т. е. множество пар элементов $\{x, y\}$, $x \in X$, $y \in Y$.

\emptyset — пустое множество.

$\bar{\Omega}$ — замыкание Ω . Но «черта над» используется и в других целях, определяемых контекстом.

$\dot{\Omega} = \partial\Omega$ — граница Ω .

$2^\Omega = \mathcal{P}(\Omega)$ — множество всех подмножеств множества Ω .

\mathbb{N} — множество натуральных чисел $\{1, 2, \dots\}$.

\mathbb{Z} — множество целых чисел $\{\dots, -1, 0, 1, \dots\}$.

$\mathbb{R} = (-\infty, \infty)$ — вещественная прямая.

Множество точек $x \in \mathbb{R}$, удовлетворяющих неравенству

$$|x - a| < \varepsilon,$$

¹⁵Любой класс множеств, замкнутый относительно симметрической разности Δ и пересечения $\cdot = \cap$, является коммутативным кольцом, в котором операции Δ и \cap играют роль сложения и умножения. Это замечание для знатоков. На какой-то стадии (за пределами анализа) оно начинает играть интригующими красками. Поначалу на него лучше не обращать внимания.

называется ε -окрестностью точки a и обозначается $I_\varepsilon(a)$. Разумеется, $I_\varepsilon(a)$ — это обыкновенный интервал $(a - \varepsilon, a + \varepsilon)$.

Множество X считается *открытым*, если каждая его точка $x \in X$ *внутренняя*, т. е. входит в X вместе с некоторой своей окрестностью. *Любой интервал — открыт*. Множество X *замкнуто*, если оно содержит все свои *предельные точки*, каковыми считаются точки, в любой окрестности которых есть элементы X . *Отрезок — всегда замкнут*.

Функция $f(x)$, преобразующая элементы множества X в элементы множества Y , обозначается как

$$f : X \rightarrow Y.$$

В случае $f(x) \equiv f_0(x)$ на $X_0 \subset X$, функцию $f : X \rightarrow Y$ называют *продолжением* f_0 на X , а f_0 — *сужением* f на X_0 .

Задание функции $f : X \rightarrow Y$ равносильно заданию ее *графика*, т. е. множества пар $\{x, y\}$, где $x \in X$, $y = f(x) \in Y$.

Глава 2

Последовательности и пределы

90% жизни уходит на приготовления.

Переход к пределу — главный инструмент математического анализа. Все основные понятия — *производные, интегралы, длины, объёмы, ряды* — получаются предельным переходом. Но заикливаться на теме, несмотря на её фундаментальность, сразу не стоит. Лучше пробежаться по диагонали. Потом к недопонятому можно вернуться. Ибо нормальный процесс обучения связан с необходимостью *лождения кругами по одному и тому же месту*. Даже если бы кто-то захотел выучить пределы окончательно, не забегая вперёд, — у него бы не получилось. Не заглядывая в «интегралы и производные», невозможно понять, на какие задачи пределы должны быть заточены. Однако в любой теме есть всё же ядро, которое выглядит более-менее одинаково с любой точки зрения. С него и надо начинать на первом этапе.



2.1 Сходимость и пределы

Числовыми последовательностями a_n называют функции целочисленного аргумента, типа $a_n = n$ или, скажем, $a_n = \frac{1}{n}$. Дабы

подчеркнуть, что речь идёт не об n -м члене, а о всей последовательности, иногда пишут $\{a_n\}$, но чаще употребляют просто a_n , возлагая понимание на контекст.

2.1.1 Определение. Числовая последовательность a_n сходится при $n \rightarrow \infty$ к пределу a

$$(\text{пишут } a_n \rightarrow a, \text{ либо } \lim_{n \rightarrow \infty} a_n = a),$$

если по любому $\varepsilon > 0$ можно указать такое N , что

$$|a_n - a| < \varepsilon \quad \text{для всех } n > N.$$

Определение 2.1.1 обычно плохо укладывается в голове, что больше связано с психологией, чем с математикой, см. раздел 13.3. Устранить психологический барьер помогает иногда взгляд на ситуацию с другого ракурса. Вот эквивалентное определение.

2.1.2 Определение. Последовательность a_n при $n \rightarrow \infty$ сходится к пределу a , если неравенство $|a_n - a| < \varepsilon$ при любом $\varepsilon > 0$ может нарушаться только конечное число раз¹.

Докажите эквивалентность, и чем чёрт не шутит, психологические затруднения уйдут с вашего пути.

Всякая теория в начале пути проходит через какие-нибудь тривиальности, которые всё же настраивают в резонанс, и заслуживают хотя бы беглого взгляда.

- Если $a_n \rightarrow a$, $b_n \rightarrow b$, и $a_n < b_n$, то $a \leq b$.

Здесь и далее «для всех n » и « $n \rightarrow \infty$ » подразумеваются.

- Если $a_n \rightarrow a$, $b_n \rightarrow b$, то

$$\boxed{a_n + b_n \rightarrow a + b}, \quad \boxed{\gamma a_n \rightarrow \gamma a}, \quad \boxed{a_n b_n \rightarrow ab}$$

и $a_n/b_n \rightarrow a/b$ при условии $b_n, b \neq 0$.

¹ Другими словами, a_n не разрешается выпрыгивать из любой фиксированной ε -окрестности $(a - \varepsilon, a + \varepsilon)$ сколько угодно раз.

◀ Пусть дано любое $\varepsilon > 0$ и все a_n принадлежат $\left(\frac{\varepsilon}{2}\right)$ -окрестности точки a для $n > N_1$, все b_n — $\left(\frac{\varepsilon}{2}\right)$ -окрестности точки b для $n > N_2$. В этом случае все $a_n + b_n$ будут находиться в ε -окрестности точки $a + b$ для $n > \max(N_1, N_2)$. ▶ Остальные свойства доказываются аналогично.

Приведённое рассуждение отдаёт казуистикой, но подобному формализму надо потихоньку учиться, чтобы уметь работать не только топором, но и скальпелем. Однако на первом этапе освоения новых понятий такими фокусами не стоит чересчур увлекаться, чтобы не сбиваться с пути.

2.1.3 Числовую последовательность, имеющую нулевой предел, $a_n \rightarrow 0$, называют бесконечно малой величиной.

Сию неприятность надо запомнить. Это неудобный сюрприз. Потому что последовательность — не величина, а функция. Однако без терминологических накладок не обойдётся. Мы же не возражаем против «подъёмного крана» из-за того, что «водопроводный» тоже кран. Здесь, надо полагать, a_n называли поначалу «переменной величиной», затем, для краткости, стали называть просто «величиной».

2.1.4 Последовательность a_n при $n \rightarrow \infty$ расходится ($a_n \rightarrow \infty$), если по любому $M > 0$ можно указать такое N , что

$$|a_n| > M \quad \text{для всех} \quad n > N. \quad (2.1)$$

Терминология здесь плавает. В случае (2.1) говорят о сходимости a_n к бесконечному пределу. Говорят также о стремлении к минус или плюс бесконечности (пишут $a_n \rightarrow \mp\infty$), если по любому $M \leq 0$ можно указать такое N , что $a_n \leq M$ для всех $n > N$. При этом расходящиеся последовательности называют бесконечно большими величинами.

2.2 Простейшие инструменты

2.2.1 Теорема о трёх собачках. Если $a_n < b_n < c_n$, и «крайние» последовательности a_n , c_n сходятся к одному и тому же пределу, то к этому же пределу сходится и b_n .

◀ Если a_n, c_n сходятся к одному пределу ζ , то с какого-то момента они все попадают в ε -окрестность ζ , и b_n , зажатое между a_n и c_n , попадает в ту же ε -окрестность. Вот и все доказательство. «Договориться с формалистом» предоставляется в качестве упражнения. ▶

Приведённый результат чаще всего работает в сценарии

$$0 < b_n - b < c_n, \quad c_n \rightarrow 0, \quad \text{тогда} \quad b_n \rightarrow b. \quad (2.2)$$

Дело в том, что определение предела 2.1.1 имеет изъян, поскольку опирается на сам предел, который бывает неизвестен. Но часто предел угадывается, и тогда (2.2) превращает догадку $b_n \rightarrow b$ в установленный факт.

Вот как это работает.

• В разложении $n = [1 + (\sqrt[n]{n} - 1)]^n$ по формуле бинома Ньютона возьмём лишь одно слагаемое $C_n^2 (\sqrt[n]{n} - 1)^2$, получится неравенство

$$n > \frac{n(n-1)}{2} (\sqrt[n]{n} - 1)^2,$$

откуда $0 < (\sqrt[n]{n} - 1)^2 < \frac{2}{n-1}$, что по *теореме о трёх собачках* даёт

$$\sqrt[n]{n} \rightarrow 1.$$

• Учитывая предыдущий результат и непрерывность функции $\log_a x$ при $x = 1$ (см. далее), получаем

$$\boxed{\frac{\log_a n}{n} = \log_a \sqrt[n]{n} \rightarrow 0}$$



• В данном примере обыгрывается лишь $\frac{1}{n} \rightarrow 0$ и, опять-таки с опережением событий, непрерывность функции $\sqrt{1+x}$ в нуле:

$$\sqrt{n}(\sqrt{n+1} - \sqrt{n}) = \frac{1}{1 + \sqrt{1 + \frac{1}{n}}} \rightarrow \frac{1}{2}.$$

Упражнения

Доказать:

$$\frac{a^n}{n} \rightarrow \infty \quad (a > 1), \quad h_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \rightarrow \infty, \quad n \sin \frac{1}{n} \rightarrow 1.$$

2.2.2 Лемма Вейерштрасса. Если последовательность a_n монотонно возрастает (не убывает) и ограничена сверху², то она сходится, $a_n \rightarrow a < \infty$.

Это уже нетривиальный результат, но он прост в использовании и представляется интуитивно очевидным. Хотя насчёт очевидности интуиция заблуждается.

Что касается обоснования леммы 2.2.2, то всё упирается в теорию вещественных чисел. В зависимости от уровня строгости построения вещественной прямой, соответствующий окрас приобретает и лемма Вейерштрасса. Надлежащий приличным кондициям вариант рассматривается далее. Наивный вариант, дескать, «вещественные числа — это десятичные дроби» — логически несостоятелен, но он приемлем на первом этапе освоения пределов, когда не до «философских глубин». Сначала надо освоиться на *этаже пребывания*. Потом можно будет вернуться назад, спуститься в подвал, зайти к смежникам и т. п.

На территории десятичных дробей, будь они неладны, у растущей и ограниченной последовательности с увеличением n перестаёт меняться все большее число знаков после запятой³. Это последовательно определяет бесконечную дробь a , которая и является пределом a_n , ибо a_n может отличаться от a лишь в более и более высоких разрядах по мере увеличения n , т. е. $a_n - a \rightarrow 0$.

2.3 Как работает лемма Вейерштрасса

Посмотрим, каковы возможности инструмента.

²Т. е. существуют такие M и N , что $a_n < M$ при $n > N$.

³Стабилизация десятичных знаков имеет место в условиях монотонного роста a_n . У сходящейся последовательности $a_n = 1 + (-1/10)^n$ все десятичные знаки «прыгают» все время.

2.3.1 Знаменитое число $e = 2,71\dots$ определяется как предел

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e \quad (2.3)$$

◀ Покажем, что последовательность $x_n = \left(1 + \frac{1}{n}\right)^n$ монотонно возрастает и ограничена. В разложении $\left(1 + \frac{1}{n}\right)^n$ по формуле бинома Ньютона k -й член

$$C_n^k \left(\frac{1}{n}\right)^k = \frac{1}{k!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right),$$

очевидно, растет при увеличении n (для фиксированного k). Кроме того, с увеличением n растет число членов разложения $\left(1 + \frac{1}{n}\right)^n$. Поэтому x_n монотонно возрастает. Кроме того, в силу $C_n^k \left(\frac{1}{n}\right)^k \leq \frac{1}{k!}$, последовательность x_n ограничена, что вытекает из оценки

$$x_n \leq 1 + 1 + \frac{1}{2!} + \dots + \frac{1}{n!} < 2 + \frac{1}{2} + \dots + \frac{1}{2^{n-1}} < 3.$$

Следовательно, лемма Вейерштрасса гарантирует её сходимость к некоторому вещественному числу (2.3), каковое «неожиданно» оказывается одним из столпов мироздания⁴. ▶

2.3.2 Отбрасывая часть слагаемых в разложении

$$x_n = \left(1 + \frac{1}{n}\right)^n = \sum_{k=1}^n C_n^k \left(\frac{1}{n}\right)^k = 1 + \sum_{k=1}^n \frac{1}{k!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right),$$

при $k \leq n$ имеем неравенство

$$x_n \geq 1 + \frac{1}{1!} + \frac{1}{2!} \left(1 - \frac{1}{n}\right) + \dots + \frac{1}{k!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right),$$

переходя к пределу, в котором при $n \rightarrow \infty$ и фиксированном k , получаем оценку

$$e \geq 1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{k!} = y_k,$$

⁴Причины вездесущности числа e вскрываются далее.

справедливую при любом k . А поскольку ещё и $x_n < y_n$, то теорема о трёх собаках приводит к другой полезной формуле для числа e ,



$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \dots$$

2.3.3 Следующий рядовой пример показывает, что Вейерштрасс востребован не только в уникальных ситуациях. В последовательности

$$x_n = \underbrace{\sqrt{p + \sqrt{p + \cdots + \sqrt{p}}}}_{n \text{ радикалов}}$$

на каждом следующем шаге последнее p в записи x_n меняется на $p + \sqrt{p}$, поэтому x_n возрастает. Ограниченность устанавливается по индукции. Неравенство $x_n < \sqrt{p} + 1$ верно при $n = 1$, а в предположении его справедливости для $n - 1$, оно верно и для n :

$$x_n^2 = p + x_{n-1} < p + \sqrt{p} + 1 < (\sqrt{p} + 1)^2.$$

Поэтому Вейерштрасс гарантирует нам $x_n \rightarrow a$. Значение a определяется переходом к пределу в равенстве $x_n^2 = p + x_{n-1}$, что даёт для a квадратное уравнение $a^2 - a - p = 0$.

2.4 Предел функции

Далее рассматриваются функции $y = f(x)$, сопоставляющие вещественным значениям аргумента x вещественные значения y .

2.4.1 Число A называют пределом функции $f(x)$ при $x \rightarrow a$ ($f(x) \rightarrow A$ при $x \rightarrow a$, пишут также $\lim_{x \rightarrow a} f(x) = A$), если по любому $\varepsilon > 0$ можно указать такое δ , что

$$|f(x) - A| < \varepsilon$$

при условии $0 < |x - a| < \delta$.

В случае $x \rightarrow \infty$ конец определения таков: если по любому $\varepsilon > 0$ можно указать такое $M > 0$, что $|f(x) - A| < \varepsilon$ для любого $x > M$.

В связи с переходом от дискретного аргумента n к непрерывному x ситуация, конечно, меняется. Главным образом это связано с возможностью стремления x к конечному пределу, что в случае дискретного аргумента — бессмысленно.

Данное определение $\lim_{x \rightarrow a} f(x) = A$ — это та самая дефиниция на (ε, δ) -языке, которая плохо укладывается в голову. Лучшее средство для преодоления барьера — переиначьте определение сто раз, переосмыслите, выверните наизнанку. Не бросайте крутить педали — и второе дыхание придёт.

2.4.2 Эквивалентное определение $f(x \rightarrow a) \rightarrow A$. Число A называется пределом функции $f(x)$ при $x \rightarrow a$, если

$$f(x_n) \rightarrow A$$

для любой последовательности $x_n \rightarrow a$ ($x_n \neq a$)

◀ Доказательство эквивалентности как будто совсем просто. Следование второго из первого очевидно. Обратное легко устанавливается от противного. Что означает нарушение $f(x) \rightarrow A$? При некотором ε не найдётся нужного δ . Другими словами, для любой последовательности $\delta_n \rightarrow 0$ существуют такие x_n , что

$$|f(x_n) - A| \geq \varepsilon, \quad 0 < |x_n - a| \leq \delta_n, \quad (2.4)$$

но это как раз противоречит второму определению. ▶



Кто увидит здесь фундаментальный провал рассуждения? Оказывается, несмотря на разрешимость системы неравенств (2.4) при любом $\delta_n > 0$, нельзя утверждать существование последовательности x_n . Не на что сослаться. Нужна специальная аксиома выбора: в любом семействе множеств в каждом множестве можно выбрать по элементу. Идиотизм вроде бы. Кажется, что ломимся в открытую дверь. Ясно, что можно. Ан нет. И эта аксиома выбора запускает феноменальный детектив, см. парадокс Банаха — Тарского.

Определение непрерывности с помощью x_n полезно на практике. Например, если существование предела установлено, то для его нахождения достаточно определить предел $f(x)$ для какой-нибудь одной подпоследовательности x_n .

Упражнение. При $x \rightarrow \infty$

$$\bullet \quad \frac{\ln x}{x} \rightarrow 0, \quad \frac{a^x}{x} \rightarrow \infty, \quad \left(1 + \frac{1}{x}\right)^x \rightarrow e.$$

• Если $f(x) \rightarrow A$, $g(x) \rightarrow B$ при $x \rightarrow a$ (или $x \rightarrow \infty$), то

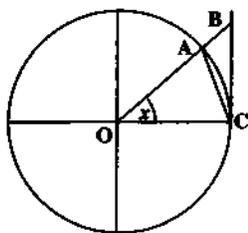
$$f(x) + g(x) \rightarrow A + B, \quad \gamma f(x) \rightarrow \gamma A, \quad f(x)g(x) \rightarrow AB$$

и $f(x)/g(x) \rightarrow A/B$ при условии $g(x), B \neq 0$.

Справедливы и другие аналоги утверждений о пределах числовых последовательностей. Теорема о трёх собачках, например. Существование предела для ограниченной и монотонной $f(x)$ тоже имеет место, но его формулировку приходится уточнять⁵. Однако в целом никаких особых трудностей здесь не возникает.

Плодящиеся определения пределов могут переполнить голову, если не смотреть в корень. Есть лишь одно понятие предела. Одна идея, одна схема. Остальное — вариации, каковые должны всплывать в голове сами по себе. Если не всплывают — тренируйтесь ещё, ходите кругами, но добейтесь результата. Что касается вариаций, то они определяются природой переменных (которые могут быть векторами, функциями и т. п.) и мерами близости $|x - a|$, $|f - A|$.

2.4.3 Важный пример. Легко видеть, что площадь $\triangle OAC$ меньше ($<$) площади сектора OAC , которая $<$ площади $\triangle OBC$, т. е.



$$\frac{1}{2}R^2 \sin x < \frac{1}{2}R^2 x < \frac{1}{2}R^2 \operatorname{tg} x,$$

откуда

$$\cos x < \frac{\sin x}{x} < 1,$$

что по теореме 2.2.1 о трёх собачках даёт

⁵ Необходимо ввести понятия пределов слева/справа или снизу/сверху, когда x приближается к a снизу $-x \uparrow a$, или сверху $-x \downarrow a$.



$$\boxed{\frac{\sin x}{x} \rightarrow 1 \text{ при } x \rightarrow 0} \quad (2.5)$$

Как говорил *Плутарх*, отвечать на вопрос можно тремя способами. (1) Сказать необходимое, (2) отвечать вежливо и (3) наговорить лишнего. В доказательствах мы стараемся, конечно, двигаться по первому варианту. Но здесь тоже есть три способа. Сказать необходимое можно

- промолчав,
- ограничившись сутью,
- выбрав максимально лаконичный путь.

Причём в последнем случае добиваться краткости нельзя за счёт механического урезания умозаключений. Действовать необходимо не только не теряя ясности, но даже приумножая её. Подобное требует определённых усилий, и если достигается успех, заслуживает подражания. Поэтому в литературе полезно обращать внимание на удачные доказательства и учиться на них эффективному математическому маневрированию. Изящное рассуждение сродни афоризму, как «хорошо отредактированному роману».

2.4.4 Функция $f(x)$ в случае $f(x) \rightarrow 0$ называется *бесконечно малой*, в случае $|f(x)| \rightarrow \infty$ — *бесконечно большой*.

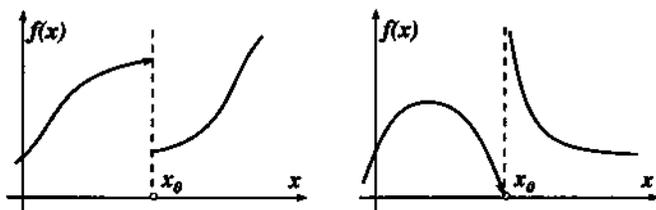
2.4.5 Если $f(x) \rightarrow 0$, $g(x) \rightarrow 0$, и $f(x)/g(x) \rightarrow 0$, то говорят, что f имеет более высокий порядок малости по сравнению с g , и пишут $f = o(g)$, читая « f есть о малое от g ».

Например, $x^2 = o(x)$ при $x \rightarrow 0$, $(x - 5)^{4/3} = o(x - 5)$ при $x \rightarrow 5$.

2.5 Непрерывные функции

2.5.1 *Определение.* Функция $f(x)$ называется *непрерывной* в точке x_0 , если $f(x) \rightarrow f(x_0)$ при $x \rightarrow x_0$. Функцию, непрерывную в любой точке $[a, b]$, называют *непрерывной на $[a, b]$* .

Дабы «чувствовать берега», сразу полезно взглянуть на разрывные функции. Разрывы могут выглядеть, например, так



Можно и пофантазировать в диапазоне от $y = \sin \frac{1}{x}$ до функции, которая в рациональных точках равна нулю, в иррациональных — единице.

2.5.2 Определение. Функция $f(x)$ непрерывна в точке x_0 и том же случае, если по любому $\varepsilon > 0$ можно указать такое δ , что

$$|f(x) - f(x_0)| < \varepsilon, \text{ если } |x - x_0| < \delta.$$

2.5.3 Теорема. Непрерывная на $[a, b]$ функция ограничена снизу и сверху⁶, и достигает минимального и максимального значения.

Упражнения

Плавать — только по книгам не научишься. Приходится лезть в воду. С математикой — такая же история.

1. Сконструировать определения: непрерывности слева и справа⁷; разрывов слева и справа; бесконечного предела, $f(x) \rightarrow \infty$.
2. Если $f(x)$ непрерывна на отрезке $[a, b]$, и $f(a) = A$, $f(b) = B$, то для любого C , лежащего между A и B , можно указать точку $c \in [a, b]$, в которой $f(c) = C$ (теорема Больцано — Коши)⁸.

$$3. \quad \lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \ln a \quad (a > 0).$$

⁶Это утверждение называют теоремой Вейерштрасса.

⁷Функции в точке.

⁸Подсказка: сначала полезно убедиться (бесконечным последовательным делением $[a, b]$ пополам) в существовании такого $c \in [a, b]$, что $f(c) = 0$, если A и B имеют разные знаки.

$$4. \quad \lim_{x \rightarrow 0} \frac{(1+x)^\lambda - 1}{x} = \lambda \quad (\lambda > 0).$$

$$5. \quad \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

$$6. \quad \lim_{x \rightarrow 0} x^{\sin x} = 1.$$

$$7. \quad \lim_{n \rightarrow \infty} \left(\cos \frac{x}{n} + \lambda \sin \frac{x}{n}\right)^n = e^{\lambda x}.$$

2.6 О теории вещественных чисел

Человеку удобнее думать о целых числах как о делениях на бесконечной прямой. Далее мысль течёт сама собой, сопоставляя точкам между делениями дроби. А когда выясняется, что этих p/q не хватает, возникает идея использования сходящихся последовательностей. Препятствия на этом пути с некоторыми усилиями преодолеваются, и вещественная прямая в самом деле может быть построена таким способом. Есть также порочный путь: «вещественные числа — это бесконечные десятичные дроби» — соблазняющий обманчивой простотой. Здесь не удаётся снять логические противоречия, но мало кто думает, что проблема так уж серьёзна. Мол, разве имеют значение мелкие нюансы?



Имеют. От мелочей зависит ВСЁ. Возьмём аксиому выбора: «в любом семействе Φ непустых множеств в каждом множестве $X \in \Phi$ можно выбрать по одному элементу». Словоблудие, казалось бы. Вещь очевидная, стоимость нулевая. Да и как оно может влиять на «грубую реальность»?

До Цермело так и думали, вернее, даже не задумывались, принимая «аксиому» за самоочевидный факт. Но будучи явно сформулирована, аксиома выбора вызвала ожесточенную критику из-за ра-

зительного контраста между своей банальностью и «невероятностью» следствий. Яркий пример — парадокс Банаха — Тарского:

Шар B в \mathbb{R}^3 допускает разбиение на конечное число непересекающихся множеств $B_1, \dots, B_k \subset \mathbb{R}^3$, из которых можно составить передвижением B_j как твёрдых тел (перенос плюс поворот) шар вдвое большего радиуса.

Так что очень отдалённая казуистика может революционным образом вмешиваться в будничную жизнь. Мелкие логические нестыковки и «пустячки» аukaются так далеко и так громко, что предосторожности выходят в математике на передний план.

Теория Дедекинда ликвидирует такие нестыковки и строит удобную игровую площадку (вещественную прямую) для матанализа. Стартовой площадкой служат рациональные числа с определёнными арифметическими операциями.

2.6.1 *Непустое множество A рациональных чисел $d(A)$ назовём сечением Дедекинда при выполнении двух условий:*

1. *В A нет наибольшего числа.*
2. *Если $\alpha \in A$, $\beta < \alpha$ и β — рациональное число, то $\beta \in A$.*



Называть сечением множество, разумеется, противоестественно. Но дело в том, что мы хотим поймать журавля в небе, имея в руках сивичу. Если бы речь шла о множестве A рациональных $x < 3$, то на роль сечения годилось бы число 3. Однако в случае « $x^2 < 5$ » указать сечением $\sqrt{5}$ нет возможности, поскольку игра начинается в отсутствие иррациональных чисел. Поэтому в роли сечения оказывается само множество A , что режет слух, но таковы стартовые условия.

Подготавливая почву для сечений стать числами, надо определить для них понятия больше, меньше, равенства, суммы и т. д. Это делается совсем легко, но скучно. Например, неравенству $d(A) < d(B)$ приходится сопоставить⁹ строгое включение: $A \subset B$, но $A \neq B$. Сумме

⁹Чтобы не нарушить уже имеющиеся неравенства для рациональных чисел.

$d(A) + d(B)$ — сечение множества $A + B$, состоящего из рациональных чисел $\alpha + \beta$, где $\alpha \in A$, $\beta \in B$. Чтобы такие определения имели смысл, надо проверить стандартные условия, которым они обязаны удовлетворять. Скажем, отношение неравенства должно быть транзитивно:

$$\alpha < \beta, \quad \beta < \gamma \Rightarrow \alpha < \gamma.$$

В данном случае это обеспечивается транзитивностью строгого включения для множеств: $A \subset B, B \subset C \Rightarrow A \subset C$. Так же легко проверяются обычные свойства сложения для сечений. И так далее. Короче, все это рутинная работа, которая заканчивается определением на сечениях обычных числовых операций. После этого термин «сечение» приравнивается термину «вещественное число». Рациональные сечения (множество элементов $x < \alpha$, где α рационально) оказываются рациональными числами. Все другие сечения называются — *иррациональными*.

Особо выделяются понятия *инфимума* и *супремума*. Пусть множество M ограничено снизу. Определим множество Γ нижних граней M , как множество таких рациональных γ , что $\gamma < t$ для любого рационального $t \in M$. Легко убедиться, что Γ является сечением.

2.6.2 Число $d(\Gamma)$ называется *точной нижней гранью* множества M и обозначается $\inf M$. Если M не ограничено снизу, то полагают $\inf M = -\infty$.

Аналогично определяется *точная верхняя грань* $\sup M$.

Остаётся главный вопрос «о полноте вещественной прямой». Рациональные числа не исчерпывали всех точек. Хватит ли для «сплошного заполнения» дедекиндовых сечений? Рациональным последовательностям a_n теперь есть «куда сходиться». Но *не потребуются ли новое пополнение для иррациональных a_n* ? Не появятся ли «новые числа», если сечения производить уже с помощью вещественных множеств A , удовлетворяющих тем же условиям 1, 2? Не появятся.

2.6.3 *Основная теорема Дедекнда. Любое сечение в области вещественных чисел является вещественным числом.*

◀ *Доказательство* совсем просто. Пусть сечение определяется множеством A вещественных чисел. Пусть A_r — множество всех рациональных чисел из A . *Вещественное число* $\sup A_r$ определяет сечение A как множество чисел $x < \sup A_r$. ▶

2.7 Надумана ли проблема и каковы блага

Лемма Вейерштрасса¹⁰ 2.2.2 получает теперь вместо наводящих соображений совершенно строгое обоснование. «У ограниченного множества $\{a_n\}$ «у Дедекунда» есть супремум $a = \sup\{a_n\} < \infty$. Из монотонного возрастания a_n теперь следует $a_n \rightarrow a$. Всё. ►

Лемма Вейерштрасса порождает цепную реакцию возникновения удобных инструментов в теории пределов.

2.7.1 Лемма о вложенных отрезках. *Пересечение бесконечного множества вложенных друг в друга отрезков,*

$$[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \quad \text{при любом } n = 1, 2, \dots,$$

*длина которых стремится к нулю, всегда не пусто*¹¹.

◀ По условию $a_1 \leq \dots \leq a_n \leq \dots \leq b_n \dots \leq \dots \leq b_1$. Поэтому последовательность a_n монотонна и ограничена, и потому сходится, $a_n \rightarrow a$. А поскольку $b_n = a_n + (b_n - a_n)$ и $b_n - a_n \rightarrow 0$, то и $b_n \rightarrow a$. Точка a , в силу $a_n \leq a \leq b_n$, принадлежит всем отрезкам. ►

Рассматривая последовательности, удобно пользоваться разными трюками перехода к *подпоследовательностям*.

• Пусть n_k произвольная расходящаяся последовательность целых чисел¹². Последовательность a_{n_k} называют *подпоследовательностью* последовательности a_n .

Если подпоследовательность a_{n_k} сходится, то её предел называют *пределной точкой* (или *точкой сгущения*) последовательности a_n .

2.7.2 Лемма Больцано — Вейерштрасса¹³. *В ограниченной последовательности a_n всегда можно выделить сходящуюся подпоследовательность.*

¹⁰ Если последовательность a_n монотонно возрастает и ограничена сверху, то она сходится, $a_n \rightarrow a < \infty$.

¹¹ Имеется точка, принадлежащая всем отрезкам.

¹² $n_k \rightarrow \infty$ при $k \rightarrow \infty$

¹³ Широко используется в рассуждениях и доказательствах.

◀ В силу ограниченности все a_n принадлежат некоторому отрезку $I_0 = [a, b]$. Разделим I_0 пополам и выберем ту его половину I_1 , которая содержит бесконечно много элементов a_n . Затем разделим I_1 пополам и выберем ту его половину I_2 , которая опять содержит бесконечное число элементов a_n . Продолжая процесс до бесконечности, получим бесконечную цепочку вложенных отрезков $I_0 \supset I_1 \supset \dots \supset I_k \supset \dots$, длины которых стремятся к нулю. В силу *леммы Вейерштрасса* все I_k имеют общую точку c . Тогда $a_{n_k} \rightarrow c$, поскольку все $a_{n_k} \in I_k$. ▶

Теперь упоминавшаяся ранее *теорема Вейерштрасса* 2.5.3 может быть обоснована строго.

2.5.3 Теорема Вейерштрасса. *Непрерывная на $[a, b]$ функция ограничена снизу и сверху.*

Утверждение интуитивно естественное¹⁴. Но попробуйте его доказать кустарными методами.

◀ Неограниченность $f(x)$, например сверху, означает, что для любой последовательности $M_n \rightarrow \infty$ можно указать такую последовательность $c_n \in [a, b]$, что $f(c_n) > M_n$. *Лемма Больцано — Вейерштрасса* гарантирует существование у c_n сходящейся подпоследовательности. Чтобы не усложнять обозначений, можно считать сходящейся саму последовательность $c_n \rightarrow c$. Тогда в силу непрерывности $f(c_n) \rightarrow f(c)$, что вступает в противоречие с $f(c_n) > M_n \rightarrow \infty$. ▶

В данном контексте стоит отметить ещё один результат.

2.7.3 Лемма Гейне — Бореля. *Из любого покрытия σ отрезка $[a, b]$ интервалами можно выбрать конечное подпокрытие.*

◀ Допустим противное. Разделим $[a, b]$ пополам и выберем ту половину, которая не покрывается конечным числом интервалов. Эту половину снова разделим пополам — и так далее. В результате получим цепочку вложенных отрезков $[a_k, b_k]$, длины которых стремятся к нулю, но каждый сегмент $[a_k, b_k]$ не покрывается конечным множеством интервалов из σ . Но тогда (*лемма о вложенных отрезках*) все

¹⁴В значительной мере потому, что интуиция считает функцию непрерывной, «если её график можно нарисовать, не отрывая карандаша от бумаги».

$[a_k, b_k]$ имеют общую точку c . Точка c принадлежит некоторому интервалу $[a^*, b^*] \in \sigma$, который, начиная с какого-то номера, содержит все последующие $[a_k, b_k]$, что рождает противоречие. ►

Леммы Гейне – Бореля и Бальцано – Вейерштрасса глубоко «эквивалентны», поскольку отражают одно и то же свойство компактности. Множество X в общем случае определяется как компактное, если из любого его покрытия открытыми множествами можно выбрать конечное подпокрытие. Так что *лемма Гейне – Бореля* устанавливает компактность любого отрезка $[a, b]$. Вообще говоря, на прямой все это большого смысла не имеет – компактность X означает ограниченность и замкнутость. Но в функциональных пространствах понятие компактности выдвигается на передний фронт и работает весьма эффективно.

2.8 Равномерная непрерывность

2.8.1 *Функция $f(x)$, непрерывная на некотором множестве X , называется равномерно непрерывной на X , если по любому $\varepsilon > 0$ можно указать такое δ , что $|f(x) - f(y)| < \varepsilon$, как только $|x - y| < \delta$, для любых $x, y \in X$.*

Функции $y = \ln x$, $y = 1/x$ на интервале $(0, 1)$ не являются равномерно непрерывными.

2.8.2 *Теорема Кантора. Функция $f(x)$, непрерывная на отрезке $[a, b]$, автоматически равномерно непрерывна на $[a, b]$.*

◀ В предположении противного для некоторого ε не найдется нужного δ . Это означает, что для любой последовательности положительных $\delta_n \rightarrow 0$ можно указать такие x_n, y_n , что $|x_n - y_n| < \delta_n$, но

$$|f(x_n) - f(y_n)| \geq \varepsilon$$

при любом $n = 1, 2, \dots$

По *лемме Бальцано – Вейерштрасса* из x_n можно выделить сходящуюся подпоследовательность. Для простоты, будем считать, что сходится сама последовательность x_n , т. е. $x_n \rightarrow c \in [a, b]$. Но тогда и $y_n \rightarrow c$ в силу $x_n - y_n \rightarrow 0$. В этом случае $|f(x_n) - f(y_n)| \geq \varepsilon$ противоречит непрерывности $f(x)$ в точке c . ►

2.9 Фундаментальные последовательности

Порок исходного определения предела заключается в опоре на знание самого предела. Имея дело с $a_n \rightarrow a$, мы часто не имеем представления о значении a , и тогда инструменты типа $|a_n - a| < \varepsilon$ теряют смысл. А если a_n не сходится (ни к чему), как это доказать? Перебирать всевозможные a ?

Снимает напряжение и даёт в руки эффективный инструмент понятие фундаментальной последовательности.

2.9.1 Последовательность a_n называется *фундаментальной*, или *последовательностью Коши*, если по любому $\varepsilon > 0$ можно указать такое N , что



$$|a_n - a_m| < \varepsilon$$

для любых $n, m > N$.

То есть у последовательности Коши члены с большими номерами не могут сильно отличаться друг от друга.

2.9.2 Критерий Коши. Последовательность a_n сходится в том и только том случае, когда она фундаментальна.

◀ Если a_n сходится, то она является последовательностью Коши. Это устанавливается «в одно касание». Действительно, в случае $a_n \rightarrow a$ все a_n при достаточно больших n оказываются в сколь угодно малой ε -окрестности точки a , и там уже a_n от a_m не могут сильно отличаться (не более, чем на 2ε).

Обратно. Пусть a_n — последовательность Коши. Покажем, что она имеет предел. Возьмём произвольную сходящуюся к нулю последовательность $\varepsilon_1, \dots, \varepsilon_k, \dots$, все $\varepsilon_k > 0$. И пусть N_1, \dots, N_k, \dots таковы, что

$$|a_n - a_m| < \varepsilon_k \quad \text{при} \quad n, m \geq N_k.$$

Тогда все a_n при $n \geq N_k$ принадлежат отрезкам

$$I_k = [a_{N_k} - \varepsilon_k, a_{N_k} + \varepsilon_k],$$

длины которых убывают до нуля при $k \rightarrow \infty$. Чтобы воспользоваться леммой о вложенных отрезках, перейдем к последовательности вложенных отрезков $J_k \subset I_k$, которые строятся по правилу

$$J_1 = I_1, \quad J_{k+1} = J_k \cap I_{k+1}.$$

Теперь «лемма» гарантирует существование общей точки a у всех J_k , которая и есть предел a_n , поскольку J_k стягиваются к a , и все $a_n \in J_k$ при $n \geq N_k$. ►

2.10 Числовые ряды

Особое внимание в анализе уделяется числовым последовательностям вида

$$A_n = a_1 + \dots + a_n. \quad (2.6)$$

Их называют *частичными суммами бесконечных рядов*

$$\sum_{n=1}^{\infty} a_n = a_1 + \dots + a_n + \dots \quad (2.7)$$

2.10.1 Конечный или бесконечный предел A частичной суммы (2.6) определяют как сумму ряда (2.7). Ряд, имеющий конечную (бесконечную) сумму, называют *сходящимся (расходящимся)*.

Примеры

- Сумма бесконечной геометрической прогрессии

$$1 + q + q^2 + \dots = \frac{1}{1 - q}, \quad |q| < 1.$$

- Вещественное число в десятичной записи

$$a_0, a_1 a_2 \dots = a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \dots$$

- Часто встречается *гармонический ряд*

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots = \infty. \quad (2.8)$$

Докажем (2.8). Установленное ранее неравенство $\left(1 + \frac{1}{n}\right)^n < e$ после логарифмирования даёт $n \ln\left(1 + \frac{1}{n}\right) < 1$, откуда

$$\frac{1}{n} > \ln\left(1 + \frac{1}{n}\right) = \ln(n+1) - \ln n.$$

Суммирование n первых таких неравенств приводит к

$$h_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} > \ln(n+1),$$

что влечёт за собой расходимость последовательности h_n .

Бесконечные ряды представляют собой эквивалентный язык для изучения последовательностей и пределов. Действительно, сходимость ряда означает сходимость *варианты*¹⁵ A_n . Обратно, сходимость любой последовательности b_n равносильна сходимости ряда



$$b_1 + (b_2 - b_1) + \dots + (b_n - b_{n-1}) + \dots$$

Отметим два простых факта.

2.10.2 *Сходящиеся ряды можно почленно складывать, вычитать и умножать на константу.* (?)

2.10.3 *Если ряд (2.6) сходится, то $a_n \rightarrow 0$.*

Интересно, что последнее утверждение на миг заставляет задуматься. Когда все $a_n > 0$, результат сразу очевиден. В общем случае возникает заминка, которая тривиально разрешается. Последовательности A_n и A_{n-1} сходятся к одному и тому же пределу. Поэтому

$$a_n = A_n - A_{n-1} \rightarrow 0. \quad \blacktriangleright$$



Многие результаты в теории числовых рядов представляют собой несложные переформулировки известных фактов из теории пределов. «Снять маску» обычно нетрудно.

¹⁵ *Варианта* — синоним *числовой последовательности*.

1. Когда все $a_n \geq 0$, ряд $a_1 + a_2 + \dots$ сходится, если его частичные суммы ограничены сверху, и расходится в противном случае.

2. Теорема сравнения положительных рядов

$$(A): a_1 + a_2 + \dots, \quad (B): b_1 + b_2 + \dots$$

Если $a_n \leq b_n$, начиная с некоторого n , либо $\frac{a_{n+1}}{a_n} \leq \frac{b_{n+1}}{b_n}$, и ряд (B) сходится, то и ряд (A) сходится. Если (A) расходится, то и (B) расходится. Ряд (B) называют мажорирующим рядом для (A).

Пример. $\frac{1}{n^2} < \frac{1}{n(n-1)}$ при $n \geq 2$. По ряд $\sum_{n=2}^{\infty} \frac{1}{n(n-1)}$ сходится,

$$\sum_{n=2}^{\infty} \frac{1}{n(n-1)} = \frac{1}{1} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots < 1.$$

Поэтому сходится и ряд $\sum_{n=2}^{\infty} \frac{1}{n^2}$, причём здесь неожиданно всплывает число π ,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6},$$

но это уже другая история.

3. Если $\frac{a_n}{b_n}$ имеет конечный строго положительный предел, то оба положительных ряда (A) и (B) сходятся или расходятся одновременно.

• В общем случае ряд (2.7) называют абсолютно сходящимся, если сходится ряд $|a_1| + |a_2| + \dots$ из абсолютных величин. Любой абсолютно сходящийся ряд сходится (?).

• Любой знакпеременный ряд

$$a_0 - a_1 + \dots + (-1)^n a_n + \dots \quad (\text{все } a_n > 0)$$

при условии монотонного стремления a_n к нулю — сходится. Поэтому, например, не сходящийся абсолютно, ряд $1 - \frac{1}{2} + \frac{1}{3} - \dots$, сходится.

◀ Частичные суммы

$$A_{2k} = (a_1 - a_2) + (a_3 - a_4) + \dots + (a_{2k-1} - a_{2k})$$

монотонно возрастают, поскольку все скобки положительны из-за монотонности убывания a_n и ограничены,

$$A_{2k} = a_1 - (a_2 - a_3) - \dots - (a_{2k-2} - a_{2k-1}) - a_{2k} < a_1.$$

Поэтому $A_{2k} \rightarrow A$. Но тогда и $A_{2k+1} = A_{2k} + a_{2k+1} \rightarrow A$. ►

Приведем два признака сходимости положительных рядов, накрывающих 90% практических ситуаций.

2.10.4 Признак Коши. Ряд (2.7) сходится, если

$$\sqrt[n]{a_n} \rightarrow a < 1,$$

и расходится, если $a > 1$.

◀ Для достаточно малого $\varepsilon > 0$, такого, что $a + \varepsilon < 1$, начиная с некоторого $n = N$, будет $a_n \leq (a + \varepsilon)^n$. Поэтому ряд (2) мажорируется сходящимся рядом

$$\sum_{n=1}^{\infty} (a + \varepsilon)^n. \quad \blacktriangleright$$

2.10.5 Признак Даламбера. Ряд (2.7) сходится, если

$$\frac{a_{n+1}}{a_n} \rightarrow a < 1,$$

а если $a > 1$ — расходится¹⁶.

Примеры

1. Ряд

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \dots \quad (2.9)$$

сходится при $s > 1$ и расходится при $s \leq 1$. Расходимость (2.9) при $s = 1$ уже была установлена. Отсюда тем более следует расходимость (2.9) при $s < 1$. Установим сходимость при $s > 1$. Очевидно,

$$\frac{1}{(2^k + 1)^s} + \frac{1}{(2^k + 2)^s} + \dots + \frac{1}{(2^{k+1})^s} < 2^k \frac{1}{(2^{k+1})^s} < \frac{1}{2^{k(s-1)}}.$$

¹⁶ Доказательство аналогично предыдущему.

Поэтому

$$\sum_{n=1}^{\infty} \frac{1}{n^s} < \sum_{k=0}^{\infty} \frac{1}{2^{k(s-1)}},$$

а мажорирующий ряд справа сходится, ибо представляет сумму бесконечно убывающей геометрической прогрессии со знаменателем $1/2^{s-1}$.

2. Ряд $\sum_n \ln\left(1 + \frac{x}{n}\right)$ при любом $x > 0$ расходится, поскольку

$$\ln\left(1 + \frac{x}{n}\right) : \frac{1}{n} = \ln\left[\left(1 + \frac{x}{n}\right)^{\frac{n-1}{x}}\right] \rightarrow 1.$$

При изучении пределов и числовых рядов довольно часто оказывается полезным следующий результат.

2.10.6 Теорема Штольца. Если последовательность y_n монотонно возрастает и $y_n \rightarrow +\infty$, то обе последовательности

$$\frac{x_n}{y_n} \quad \text{и} \quad \frac{x_n - x_{n-1}}{y_n - y_{n-1}}$$

имеют одинаковый предел (либо обе расходятся).

Инструмент заточен на неопределённости $\frac{x_n}{y_n}$ типа $\frac{\infty}{\infty}$. Теорема Штольца особенно хорошо работает в ситуациях типа следующей:

$$\text{если } a_n \rightarrow a, \quad \text{то } \frac{a_1 + \dots + a_n}{n} \rightarrow a.$$

Результат получается «сам собой», если положить

$$x_n = a_1 + \dots + a_n, \quad y_n = n.$$

Упражнения

$$1. \quad \lim_{n \rightarrow \infty} \frac{1 + \sqrt{2} + \sqrt[3]{3} + \dots + \sqrt[n]{n}}{n} = 1.$$

$$2. \quad \lim_{n \rightarrow \infty} \frac{1^k + 2^k + \dots + n^k}{n^{k+1}} = \frac{1}{k+1}.$$

$$3. \quad \lim_{n \rightarrow \infty} \left(\frac{\sqrt[n]{a} + \sqrt[n]{b}}{2} \right)^n = \sqrt{ab} \quad (a, b > 0).$$

4. В случае $a_n \rightarrow a > 0$ ряд $\sum_n (a_n x)^n$ сходится при $|x| < \frac{1}{a}$.

5. Ряд $\sum_n \frac{x^n}{n!}$ сходится при любом x (воспользуйтесь признаком Даламбера).

6. Ряд $\sum_n n^k x^n$ сходится при любых $|x| < 1$ и k .

Числовые ряды служат базой для последующего изучения функциональных рядов, которые образуют одно из магистральных направлений в анализе. Это кое-что объясняет в расстановке акцентов, наблюдаемой в учебниках. Например, особое внимание к бесконечным суммам $\sum_n a_n b_n$ может показаться странным, если не знать, что в дальнейшем подразумевается переход к изучению рядов $\sum_n a_n x^n$.

Достаточно естественными представляются вопросы влияния на сходимость рядов стандартных операций: перегруппировки членов (изменения порядка суммирования), умножения рядов. При этом выясняется, что между абсолютно сходящимися рядами и всеми остальными проходит мощный водораздел. Абсолютно сходящиеся ряды «беспроблемны». Они допускают любое изменение порядка суммирования. Их можно без предосторожностей перемножать,

$$\sum_n a_n = A, \quad \sum_n b_n = B \quad \rightarrow \quad \sum_{n,k} a_n b_k = AB,$$

независимо от порядка суммирования членов $a_n b_k$. В отсутствие абсолютной сходимости возможны любые сюрпризы.

2.10.7 Теорема Римана. *Не сходящийся абсолютно ряд всегда допускает изменение порядка суммирования, при котором его сумма оказывается равной любому наперёд заданному числу (конечному или бесконечному).*

Глава 3

Производная и дифференциал

*Существует достаточно света для тех,
кто хочет видеть,
и достаточно мрака для тех,
кто — не хочет.*
Паскаль

3.1 Производная

Производная $f'(x)$ — это *мгновенная скорость изменения* функции $f(x)$ в точке x . Если x — время¹, а $f(x)$ — пройденный путь, производная $f'(x)$ — это обычная мгновенная скорость в момент времени x . Если « x — время», производную принято обозначать точкой сверху, \dot{f} .

Подсознание имеет смысл приучать к интерпретации $f'(x)$ как скорости изменения $f(x)$ при изменении аргумента x , *независимо от природы x и $f(x)$* . Скажем, вполне естественно говорить о скорости $f'(x)$ изменения плотности $f(x)$ атмосферы по мере увеличения высоты x .

Теперь о формальном определении.

¹Коль скоро внутри возникает дискомфорт, когда время обозначается не буквой t , над этим стоит поработать. Подсознание не должно приклеивать суть к случайным обстоятельствам.

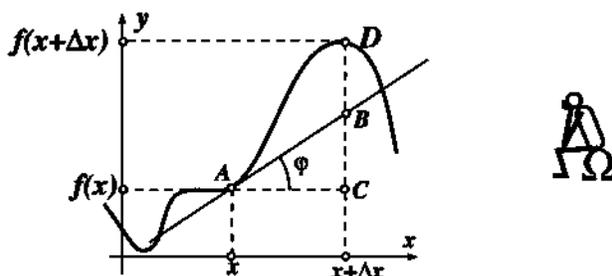
3.1.1 Производной $f'(x)$ в точке x называется предел



$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}, \quad (3.1)$$

если таковой существует. Операцию взятия производной называют *дифференцированием*. При существовании предела (3.1) функцию $f(x)$ называют *дифференцируемой* в точке² x .

На картинке ситуация выглядит следующим образом.



Касательная к графику $y = f(x)$ в точке A тем точнее приближает дугу AD , чем меньше Δx . Разумеется, это само по себе сказано не вполне точно. Во-вторых, это ниоткуда не следует. Так будет, когда всё хорошо, т. е. существует предел (3.1). Тогда при уменьшении Δx вертикальная прямая CD движется влево, и $BD/DC \rightarrow 0$, что и является эквивалентом существования предела (3.1), который в терминах геометрического рисунка оказывается равным $\operatorname{tg} \varphi$, т. е. *производная численно равна тангенсу угла наклона касательной к графику $f(x)$ в точке x* ,

$$f'(x) = \operatorname{tg} \varphi$$

Имеет смысл сразу осознать, что существование предела (3.1) — штука достаточно естественная и широко распространённая. Пусть $f(x) = x^2$, тогда в соответствии с (3.1)

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \rightarrow 0} (2x + \Delta x) = 2x.$$

²И говорят, что $f(x)$ дифференцируема на множестве $X \subset \mathbb{R}$, если она дифференцируема в любой точке $x \in X$.

Совсем просто получается $x' = 1$. (?) Почти так же легко вычисляются производные других элементарных функций. Но прежде чем этим заниматься, мы рассмотрим некоторые вспомогательные инструменты, облегчающие взятие производных конкретных функций.

3.2 Правила дифференцирования

1. Если известна производная $f'(x)$, то для $y = cf(x)$, c — константа, можно сразу написать $y' = cf'(x)$, ибо

$$\frac{cf(x + \Delta x) - cf(x)}{\Delta x} = c \frac{f(x + \Delta x) - f(x)}{\Delta x} \rightarrow cf'(x).$$

• $(\pi x^2)' = 2\pi x$, $(3x)' = 3$.

2. Производная суммы равна сумме производных



$$(f + g)' = f' + g'$$



Доказательство — в качестве упражнения.

• $(x^2 - 7x)' = 2x - 7$.

3. Производная произведения равна

$$(fg)' = f'g + fg' \tag{3.2}$$

◀ Пусть $y = f \cdot g$, а Δf , Δg — обозначают приращения функций, например, $\Delta f = f(x + \Delta x) - f(x)$. Тогда

$$y' = \lim_{\Delta x \rightarrow 0} \frac{(f + \Delta f)(g + \Delta g) - fg}{\Delta x} = \lim_{\Delta x \rightarrow 0} \left(\frac{\Delta f}{\Delta x} g + f \frac{\Delta g}{\Delta x} + \frac{\Delta f}{\Delta x} \Delta g \right),$$

что и даёт (3.2), поскольку

$$\frac{\Delta f}{\Delta x} \rightarrow f', \quad \frac{\Delta g}{\Delta x} \rightarrow g', \quad \frac{\Delta f}{\Delta x} \Delta g \rightarrow f' \Delta g \rightarrow 0. \quad \blacktriangleright$$

$$\bullet (x^3)' = (x^2 \cdot x)' = 2x \cdot x + x^2 \cdot 1 = 3x^2.$$

4. *Правило дифференцирования сложной функции (композиции функций):*

$$\boxed{y = f(g(x)) \quad \rightarrow \quad y' = f' \cdot g'} \quad (3.3)$$

Здесь f дифференцируется «по всему своему аргументу, т. е. по $g(x)$ ». Допустим, $y = (3x)^2$. Полагаем $g(x) = 3x$, $f(\cdot) = (\cdot)^2$. Дифференцирование квадрата (по $3x$) даёт $2 \cdot 3x$, $(3x)' = 3$, поэтому $y' = 18x$.

◀ Доказательство (3.3) занимает одну строчку³,

$$y' = \lim_{\Delta x \rightarrow 0} \frac{f(g + \Delta g) - f(g)}{\Delta g} \cdot \frac{\Delta g}{\Delta x} = f' \cdot g'. \quad \blacktriangleright$$

Понятно, что цепное правило дифференцирования сложной функции можно индуктивно продолжить

$$y = f(g(h(x))) \quad \Rightarrow \quad y' = f'_g \cdot g'_h \cdot h'_x, \quad (3.4)$$

и так далее. Нижний индекс в (3.4) подчёркивает (напоминает), по какому аргументу производится дифференцирование.

5. Производная частного

$$\boxed{\left(\frac{f}{g}\right)' = \frac{f' \cdot g - f \cdot g'}{g^2}}$$



легко устанавливается с помощью предыдущих правил. Частное можно считать произведением f и $1/g$, а производную $(1/g)'$ вычислять как производную сложной функции⁴.

³Напомним, $\Delta g = g(x + \Delta x) - g(x)$.

⁴Здесь мы, немного забегая вперёд, имеем в виду $\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$.

6. *Производная обратной функции.* Если для $y = f(x)$ определена обратная функция $x = f^{-1}(y)$, то⁵

$$x'_y = \frac{1}{y'_x}$$

что легко получается⁶ предельным переходом в очевидном равенстве $\Delta x / \Delta y = 1 / (\Delta y / \Delta x)$.

Человек, увы, мыслит в «обратном направлении» всегда хуже. Трижды набивший оскомину пример тюбика, из которого выдавить пасту легче, чем загнать её обратно, образно отражает положение дел. Именно в связи с трудностями манипулирования обратными функциями — использование правила $x'_y = 1/y'_x$ часто вызывает затруднения. Осваивать такие вещи лучше всего на примерах. Определим производную арксинуса, исходя из формулы для производной синуса, $(\sin x)' = \cos x$. Пусть $y = \sin x$, $x = \arcsin y$. Тогда

$$x'_y = \frac{1}{y'_x} = \frac{1}{\cos x} = \frac{1}{\sqrt{1 - \sin^2 x}} = \frac{1}{\sqrt{1 - y^2}},$$

что даёт необходимый результат. При желании буквы x, y теперь можно поменять местами, $(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}}$.

3.3 Дифференциалы

Определение производной (3.1) означает

$$\text{☺} \quad f'(x) = \frac{f(x + \Delta x) - f(x)}{\Delta x} + \xi(\Delta x), \quad (3.5)$$

где $\xi(\Delta x) \rightarrow 0$ при $\Delta x \rightarrow 0$. Равенство (3.5) запишем в виде⁷

$$\Delta f = f'(x) \cdot \Delta x + o(\Delta x), \quad (3.6)$$

⁵Ещё раз обратите внимание на использование нижнего индекса, который указывает, по какому аргументу производится дифференцирование. Это удобный приём, широко используемый в анализе.

⁶Мы избегаем «разжевывания», которое иногда создаёт иллюзию исчерпывающего объяснения, но всегда топит суть в трясине деталей.

⁷Здесь $\Delta f = f(x + \Delta x) - f(x)$.

где $o(\Delta x) = \xi(\Delta x) \Delta x$, и потому $o(\Delta x)$ не просто стремится к нулю при $\Delta x \rightarrow 0$, а гораздо быстрее, чем Δx . Точнее говоря,

$$\frac{o(\Delta x)}{\Delta x} \rightarrow 0 \quad \text{при} \quad \Delta x \rightarrow 0. \quad (3.7)$$

Условие (3.7) определяет *о-малое*, $o(\cdot)$, как стенографический трюк, удобный и широко используемый. Для усвоения полезно объяснить соседу, что такое

$$o(x), \quad o(x^3), \quad o(x^2 - 1), \quad o\left(\frac{1}{x}\right).$$

Два слова о понятии *порядка малости*, позволяющем в одно касание уходить от неуклюжих пояснений. Пусть $\alpha(x)$ и $\beta(x)$ — бесконечно малые величины, т. е. функции, убывающие до нуля при $x \rightarrow 0$, и

$$\frac{\alpha(x)}{\beta(x)} \rightarrow 0 \quad \text{при} \quad x \rightarrow 0.$$

Тогда говорят, что $\alpha(x)$ имеет *более высокий порядок малости* чем $\beta(x)$. На этом языке $o(\Delta x)$ — величина более высокого порядка малости чем Δx .

Перейдём теперь к главной теме параграфа.

3.3.1 *Линейная часть приращения Δy , равная $f'(x) \cdot \Delta x$ в представлении (3.6), называется дифференциалом функции $y = f(x)$ в точке x , и обозначается $dy = df$.*

Приращение Δx в этом контексте обозначают как dx , и вместо (3.6) в линейном приближении имеем

$$df = f'(x) \cdot dx. \quad (3.8)$$



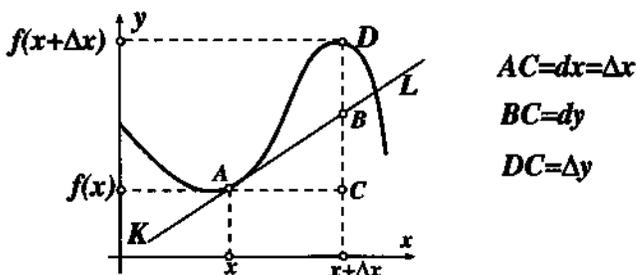
Поскольку рассмотрению дифференциалов сопутствует обычно большая неразбериха, остановимся на некоторых реверансах, хотя это и сопряжено с издержками типа «объяснения анекдота». Часто определяют df как *линейную функцию*, что у новичков уводит почву из под ног. Это функция в том же самом смысле, что и в случае $y = 2x$, переменная y — функция, но при данном значении аргумента y — это уже число. Точно также dy при заданном dx — это число. И потому с дифференциалами можно обращаться как с числами, делить одно на другое, сокращать. В частности, деление (3.8) на dx даёт

$$f'(x) = \frac{df}{dx}, \quad (3.9)$$

что служат ещё одним общепринятым обозначением производной.

Кроме того, дифференциалы — это отнюдь не «маленькие дельта», как иногда думают. Они могут принимать любые значения.

Ситуацию (3.8) адекватно описывает следующее представление. Функция $y = f(x)$ в окрестности точки x аппроксимируется линейным приближением $s = f'(x)r$, где r и s обозначают линейные приращения вдоль осей x , y . Обозначая r через dx , s — через dy , имеем (3.8). На рисунке



линейная аппроксимация $y = f(x)$ представляет собой прямую KL , касающуюся графика $f(x)$ в точке A и описываемую линейной функцией $df = f'(x) \cdot dx$ в координатах $\{dx, dy\}$ с началом отсчёта в точке A .

Всё это не имело бы особого смысла, если бы речь шла только о ситуациях типа (3.6), (3.8). Выигрыш от рассмотрения диффе-

ренциалов невелик, а головной боли много⁸. Выгоды, конечно, есть. Например, при работе с приращениями функций приходится использовать приближённые равенства $\Delta y \approx f'(x) \Delta x$, таская из строчки в строчку хвосты нелинейных добавок $o(\cdot)$. Понятие дифференциала освобождает от этих неудобств, позволяя писать для линейных частей приращений абсолютно строгие равенства. Это особенно удобно в более сложных ситуациях. Например, для $y = f(x)g(x)$ имеем⁹

$$\boxed{dy = df \cdot g + f \cdot dg},$$

что сразу получается умножением формулы производной произведения двух функций на dx . Одна лишь возможность обращаться с df/dx , как с обыкновенной дробью, порождает массу удобств. Например, $\frac{df}{dx} \frac{dx}{dt} = \frac{df}{dt}$.

3.3.2 Инвариантность формы дифференциала

Допустим, в $y = f(x)$ производится замена $x = u(\tau)$, после которой x становится зависимой переменной. И хотя теперь уже $\Delta x \neq dx$ — равенство $dy = f'(x)dx$ сохраняется, в чем легко убедиться используя формулу дифференцирования сложной функции:

$$dy = y'_\tau d\tau = y'_x x'_\tau d\tau = y'_x dx.$$

Это свойство называют *инвариантностью формы дифференциала*.

Подобные «пустячки» пускают анализ в другое русло. Дифференциалы шаг за шагом обретают другой облик. Становится ясно, что линейные аппроксимации функций представляют собой удобную и продуктивную *категорию мышления*. Оказывается, их (аппроксимации — дифференциалы) в своей нише можно полезным образом комбинировать, оставляя за бортом кандалы нелинейных добавок.

⁸Поначалу действительно не ясно, зачем вводятся дифференциалы. Тот небольшой выигрыш, который они дают, на первом курсе выглядит неубедительно, ибо пока ничего не известно о «территориях», где всё это может хорошо «выстрелять». Но если присмотреться, то это хорошо «стреляет» и в самых простых ситуациях.

⁹Не говоря уж о $d(f + g) = df + dg$ или $d\zeta f = \zeta df$.

Формула для производной dy/dx , допускающая сокращения как в обыкновенных дробях (но не буквы $d!$), часто оказывается предпочтительнее. Например, в случае параметрического задания кривой,

$$x = \varphi(\tau), \quad y = \psi(\tau),$$

производную y'_x с помощью дифференциалов можно вычислить, не восстанавливая зависимости $y(x)$:

$$y'_x = \frac{dy}{dx} = \frac{y'_\tau d\tau}{x'_\tau d\tau} = \frac{\psi'(\tau)}{\varphi'(\tau)}$$



3.3.3 Дифференциалы высших порядков

Вторым дифференциалом d^2y функции $y = f(x)$ называется (первый) дифференциал функции $dy = f'(x)dx$ (как функции x , но не dx), т. е.

$$d^2y = d(f'(x))dx = (f''(x)dx)dx = f''(x)dx^2.$$

Дифференциалы более высокого порядка определяются индуктивно¹⁰, $d^n y = f^{(n)}(x)dx^n$.

В отличие от дифференциалов первого порядка, для которых всегда $dy = f'(x)dx$ (инвариантность формы), приведённые формулы для дифференциалов d^n ($n > 1$) справедливы лишь в предположении независимости аргумента x . Если $x = \varphi(t)$, то

$$\begin{aligned} d^2y &= [f(\varphi(t))]''_{t^2} dt^2 = [f'(\cdot)\varphi'(\cdot)]' dt^2 = \\ &= (f'' \cdot (\varphi')^2 + f' \cdot \varphi'') dt^2 = f'' dx^2 + f' d^2x. \end{aligned}$$

3.4 Производные элементарных функций

• Производная степенной функции

В случае x^n при целом $n \geq 1$ из формулы *бинома Ньютона*,

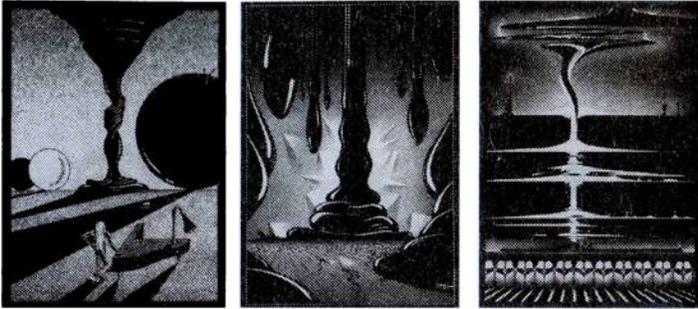
$$(x + \Delta x)^n = x^n + nx^{n-1} \cdot \Delta x + C_n^2 x^{n-2} \cdot \Delta x^2 + \dots,$$

¹⁰ Вторая производная $f''(x)$ получается дифференцированием $f'(x)$. Дифференцируя $f(x)$ n раз, будем иметь n -ю производную $f^{(n)}(x)$.

следует

$$(x + \Delta x)^n - x^n = nx^{n-1} \Delta x + o(\Delta x),$$

откуда легко вытекает $(x^n)' = nx^{n-1}$.



Рисунки А. Фоменко

В общем случае степенной функции $y = x^\lambda$ с произвольным $\lambda \neq 0$ возни чуть больше. Сначала установим вспомогательный факт

$$\blacktriangleleft \frac{(1+t)^\lambda - 1}{t} \rightarrow \lambda \quad \text{при } t \rightarrow 0.$$

Вводя новую переменную $s = (1+t)^\lambda - 1$ (очевидно, $s \rightarrow 0$ при $t \rightarrow 0$) и логарифмируя равенство $(1+t)^\lambda = 1+s$, имеем

$$\lambda \ln(1+t) = \ln(1+s),$$

откуда

$$\frac{(1+t)^\lambda - 1}{t} = \frac{s}{t} = \lambda \frac{s}{\ln(1+s)} \frac{\ln(1+t)}{t} \rightarrow \lambda,$$

в силу того, что (при $t \rightarrow 0$)

$$\frac{\ln(1+t)}{t} = \ln(1+t)^{\frac{1}{t}} \cdot \ln e = 1.$$

Теперь легко получаем нужный результат

$$\begin{aligned}
 y' &= \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^\lambda - x^\lambda}{\Delta x} = \\
 &= \lim_{\Delta x \rightarrow 0} \frac{\left(1 + \frac{\Delta x}{x}\right)^\lambda - 1}{\frac{\Delta x}{x}} x^{\lambda-1} = \lambda x^{\lambda-1}. \blacktriangleright
 \end{aligned}$$

Таким образом $\boxed{(x^\lambda)' = \lambda x^{\lambda-1}}$. В частности,

$$\boxed{\text{const}' = 0, \quad x' = 1, \quad (\sqrt{x})' = \frac{1}{2\sqrt{x}}}$$

• Производная показательной функции

◀ Для $y = a^x$ имеем

$$y' = \lim_{\Delta x \rightarrow 0} \frac{a^{x+\Delta x} - a^x}{\Delta x} = \lim_{\Delta x \rightarrow 0} a^x \frac{a^{\Delta x} - 1}{\Delta x} = a^x \ln a,$$

поскольку

$$\frac{a^\tau - 1}{\tau} \rightarrow \ln a \quad \text{при } \tau \rightarrow 0,$$

что следует из

$$\frac{a^\tau - 1}{\tau} = \frac{\sigma}{\log_a(1 + \sigma)} = \frac{1}{\log_a(1 + \sigma)^{1/\sigma}} \rightarrow \frac{1}{\log_a e} = \ln a,$$

где $a^\tau - 1 = \sigma$. ▶

Таким образом $\boxed{(a^x)' = a^x \ln a}$. В частности,



$$\boxed{(e^x)' = e^x}$$

(3.10)

Особая роль числа e при дифференцировании показательной функции — есть *та самая причина*, которая ставит e в ряд важнейших констант. Принципиальная роль показательной

функции в «устройстве Вселенной» — это уже другой вопрос. Из приведённых формул ясно, например, что дифференциальному уравнению $y' = ky$ удовлетворяет функция $y = e^{kx}$. Аналогичным образом e появляется при решении любых линейных дифференциальных уравнений, которыми описывается большинство прикладных задач в физике, биологии, экономике и других областях. Функцию e^x называют также *экспонентой*. Вместо e^x иногда используют обозначение $\boxed{\text{exp } x}$.

• Производная логарифмической функции

Функции $x = a^y$ и $y = \log_a x$ взаимнообратны. Поэтому производная логарифмической функции определяется формулой *производной обратной функции*.

$$y'_x = \frac{1}{x'_y} = \frac{1}{a^y \ln a} = \frac{1}{x \ln a}.$$

Таким образом,

$$\boxed{(\log_a x)' = \frac{1}{x \ln a}, \quad (\ln x)' = \frac{1}{x}}.$$



• Производные тригонометрических функций

В случае $y = \sin x$ имеем¹¹.

$$y' = \lim_{\Delta x \rightarrow 0} \frac{\sin(x + \Delta x) - \sin x}{\Delta x} = \lim_{\Delta x \rightarrow 0} \cos \left(x + \frac{\Delta x}{2} \right) \frac{\sin(\Delta x/2)}{(\Delta x/2)} = \cos x,$$

с учётом установленного ранее предела (2.5),

$$\frac{\sin \tau}{\tau} \rightarrow 1 \quad \text{при} \quad \tau \rightarrow 0.$$

Производная косинуса вычисляется аналогично. Можно также воспользоваться формулой производной композиции функций либо для представления

$$\cos x = \sqrt{1 - \sin^2 x}, \quad (3.11)$$

¹¹Пользуясь формулой разности синусов.

либо, что целесообразнее, для

$$\cos x = \sin \left(\frac{\pi}{2} - x \right). \quad (3.12)$$

Тут, как говорится, все дороги ведут в Рим. Но в случае (3.11) необходимы предосторожности и лишние усилия из-за «плохо выбранной дороги». Манипуляция

$$(\cos x)' = \frac{1}{2\sqrt{1 - \sin^2 x}} (-2 \sin x \cos x) = -\frac{\sin x \cos x}{\cos x} = -\sin x$$

представляет собой лишь канву. Корень — функция неоднозначная, сокращать на $\cos x$, который время от времени обращается в ноль, тоже не с руки. Все эти препятствия можно обойти, но зачем лишняя работа. При опоре на (3.12) производная косинуса вычисляется без проблем. В итоге

$$\boxed{(\sin x)' = \cos x, \quad (\cos x)' = -\sin x.}$$

Производные тангенса и котангенса вычисляются по формуле производной отношения двух функций.

$$\boxed{(\operatorname{tg} x)' = \frac{1}{\cos^2 x}, \quad (\operatorname{ctg} x)' = -\frac{1}{\sin^2 x}.}$$

Наконец, производные обратных тригонометрических функций вычисляются по формуле производной для обратной функции, что уже было продемонстрировано ранее на примере арксинуса.

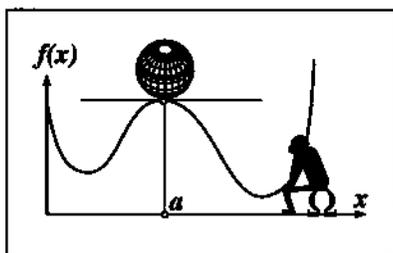
$$\boxed{(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}}, \quad (\arccos x)' = -\frac{1}{\sqrt{1 - x^2}},}$$

$$\boxed{(\operatorname{arctg} x)' = \frac{1}{1 + x^2}, \quad (\operatorname{arcctg} x)' = -\frac{1}{1 + x^2}.}$$

3.5 Тропа на вершину Тейлора

О полезности изучения закоулков Вселенной с дифференциальной точки зрения свидетельствуют многие факты. Один из них

3.5.1 Теорема Ферма. Пусть $f(x)$ в точке $x = a$ дифференцируема и принимает локально максимальное значение, т. е. $f(a) \geq f(x)$ для всех x из достаточно малой окрестности точки a . Тогда $f'(a) = 0$.



◀ В предположении противного, $f'(a) \neq 0$, например $f'(a) > 0$, линейная (самая большая при малом Δx) часть приращения

$$f'(a) \Delta x > 0 \quad \text{при} \quad \Delta x > 0,$$

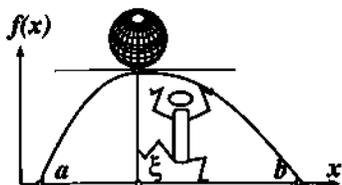
т. е. $f(a + \Delta x) > f(a)$ при достаточно малых $\Delta x > 0$, что противоречит наличию локального максимума в a . ▶

На ситуацию полезно взглянуть и с других позиций. При движении точки в положении максимального удаления точка останавливается¹² - скорость обнуляется. Очевидно также, что касательная к графику в точке локального максимума — горизонтальна, т. е. $\text{tg } \varphi = f'(a) = 0$.

Обратное, разумеется, неверно. У x^3 производная ($3x^2$) в нуле — нуль, но максимума нет. В нуле у x^3 точка перегиба.

¹²Чтобы двинуться обратно.

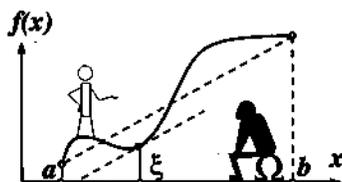
3.5.2 Теорема Ролля. Пусть $f(x)$ дифференцируема на $[a, b]$ и $f(a) = f(b)$. Тогда есть точка $\xi \in [a, b]$, в которой $f'(\xi) = 0$.



◀ Из $f(a) = f(b)$ вытекает, что $f(x)$ на $[a, b]$ имеет или минимум, или максимум. Далее решает ссылка на предыдущую теорему. ▶

3.5.3 Теорема Лагранжа. Пусть $f(x)$ дифференцируема на $[a, b]$. Тогда существует точка $\xi \in [a, b]$, в которой

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$



Последнее равенство чаще записывают в виде



$$f(b) - f(a) = f'(\xi)(b - a),$$

подчеркивая способ выражения приращения Δf с помощью умножения $\Delta x = a - b$ на «среднюю скорость роста» $f'(\xi)$.

◀ Для доказательства введем вспомогательную функцию

$$\omega(x) = f(x) - kx,$$

которая при $k = [f(b) - f(a)] / (b - a)$ удовлетворяет условию $\omega(a) = \omega(b)$. Доказательство завершает применение к $\omega(x)$ теоремы Ролля. Для некоторого $\xi \in [a, b]$ будет $\omega'(\xi) = f'(\xi) - k = 0$. ▶

3.5.4 Теорема Коши. Пусть $f(x)$ и $g(x)$ дифференцируемы на $[a, b]$ и $g'(x) \neq 0$ в промежутке $[a, b]$. Тогда существует точка $\xi \in [a, b]$, в которой

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

◀ Доказательство снова получается сведением к *теореме Ролля* с помощью функции $\omega(x) = f(x) - kg(x)$, которая при

$$k = [f(b) - f(a)]/[g(b) - g(a)]$$

удовлетворяет условию $\omega(a) = \omega(b)$. Точка $\xi \in [a, b]$, в которой $\omega'(\xi) = f'(\xi) - kg'(\xi) = 0$, даёт «то, что нужно». ▶

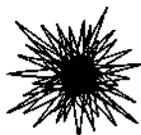
Теорема Коши допускает следующее усиление.

3.5.5 Теорема Коши-плюс. Пусть функции $f(x)$ и $g(x)$ дифференцируемы $n + 1$ раз в некоторой окрестности точки a и обращаются в этой точке в нуль вместе со своими n производными, причём для $x \neq a$ из этой окрестности

$$g(x) \neq 0, \quad g'(x) \neq 0, \quad \dots, \quad g^{(n+1)}(x) \neq 0.$$

Тогда существует точка ξ в окрестности a , в которой

$$\frac{f(x)}{g(x)} = \frac{f^{(n+1)}(\xi)}{g^{(n+1)}(\xi)}.$$



◀ Если $f(a) = g(a) = 0$, то теорема 3.5.4 в применении к промежутку $[a, x]$ (или $[x, a]$, если $x < a$) гарантирует существование такого $\xi \in [a, x]$, что

$$\frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)}.$$

Применим теперь этот факт к $\frac{f'(\xi)}{g'(\xi)}$, и так — n раз. ▶

3.5.6 Теорема Дарбу. Если $f(x)$ дифференцируема на $[a, b]$, то $f'(x)$ принимает все промежуточные значения между $f'(a)$ и $f'(b)$. (?)

Формально к *теореме Коши-плюс* надо добавить замечание о **производных высших порядков**, поскольку до сих пор о повторном дифференцировании речь не заходила.

• Дифференцируя $f'(x)$, получаем вторую производную $f''(x)$ функции $f(x)$. Дифференцируя $f(x)$ n раз, будем иметь n -ю производную $f^{(n)}(x)$ функции $f(x)$. Для обозначения второй (n -й) производной используется также запись $\frac{d^2 f}{dx^2}$ ($\frac{d^n f}{dx^n}$). Когда аргумент время — популярны точки сверху, \ddot{z} вместо z'' . Обычно $f(x)$ предполагается столько раз дифференцируемой, сколько требуется по контексту.

3.6 Разложение Тейлора

Кульминационным моментом дифференциального исчисления обычно представляется следующий факт.

3.6.1 Формула Тейлора. Пусть функция $f(x)$ дифференцируема $n + 1$ раз в некоторой окрестности точки a . Тогда для x достаточно близких к a справедлива формула

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \dots + \frac{f^n(a)}{n!}(x-a)^n + o((x-a)^n)$$

◀ Легко видеть что аппроксимирующий полином

$$P_n(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \dots + \frac{f^n(a)}{n!}(x-a)^n$$

имеет в a те же производные (до n -й включительно), что и $f(x)$, поэтому функции¹³

$$\varphi(x) = f(x) - P_n(x) \quad \text{и} \quad \psi(x) = (x-a)^{n+1}$$

удовлетворяют условиям *теоремы Коши-плюс*. Следовательно,

$$\frac{\varphi(x)}{\psi(x)} = \frac{\varphi^{n+1}(\xi)}{\psi^{n+1}(\xi)}$$

в некоторой точке ξ в окрестности a .

Очевидно, $\varphi^{n+1}(\xi) = f^{n+1}(\xi)$, $\psi^{n+1}(\xi) = (n+1)!$. Поэтому

$$\frac{f(x) - P_n(x)}{(x-a)^{n+1}} = \frac{f^{n+1}(\xi)}{(n+1)!},$$

что и требуется, поскольку

$$\frac{f^{n+1}(\xi)}{(n+1)!}(x-a)^{n+1} = o((x-a)^n). \quad \blacktriangleright$$

Если функция $f(x)$ бесконечное число раз дифференцируема, скажем, в нуле, то велик соблазн представить её в виде бесконечного ряда Тейлора

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \dots + \frac{f^n(0)}{n!}x^n + \dots \quad (3.13)$$

¹³Имеющие в a нулевые производные до n -й включительно.

Вопрос в том, имеет ли смысл равенство (3.13). Оценки малости «хвоста» ряда Тейлора (остаточного члена) $f(x) - P_n(x)$ ничего не дают. Качество приближения $f(x)$ полиномом $P_n(x)$ улучшается с ростом n , но окрестность, где это происходит, может уменьшаться до нуля. Поэтому для знака равенства в (3.13) нет оснований. Ряд Тейлора может расходиться или сходиться к другой функции.

Например, для функции $f(x) = e^{-1/x^2}$, доопределённой в нуле по непрерывности ($f(0) = 0$), все производные $f^{(n)}(0) = 0$. Поэтому её ряд Тейлора (3.13) сходится к (равен) тождественному нулю.

Тем не менее для многих функций представление (3.13) справедливо, причём не в малой окрестности, а на довольно широких областях и даже на всей числовой прямой. Например, представления¹⁴

$$\begin{aligned} e^x &= 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots, \\ \sin x &= x - \frac{x^3}{3!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots, \\ \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + \dots \end{aligned}$$

справедливы при любом $x \in (-\infty, +\infty)$.

Так или иначе, но в теории функций комплексного переменного (глава 10) устанавливается, что ряды Тейлора, с которыми приходится иметь дело на практике, сходятся в обширных областях.

Вот ещё несколько рядов

$$\begin{aligned} \operatorname{arctg} x &= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots, \\ \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^5}{5} + \dots, \quad (|x| < 1), \\ \frac{1}{1-x} &= 1 + x + x^2 + x^3 + \dots, \quad (|x| < 1), \\ (1+x)^\alpha &= 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \dots, \quad (|x| < 1). \end{aligned}$$

¹⁴Именно глядя на эти ряды, как гласит легенда, Эйлер открыл свою знаменитую формулу

$$e^{i\varphi} = \cos \varphi + i \sin \varphi$$

Ради справедливости надо подчеркнуть, что тейлоровская тематика ориентирована на изучение особенностей поведения функции в точке — а там ограничение ракурса малыми окрестностями естественно. Если бы задача состояла в аппроксимации функций на больших промежутках, то можно было бы вообще заходить с другого конца.

3.6.2 Теорема Вейерштрасса (аппроксимационная). Для непрерывной на $[a, b]$ функции $f(x)$ и любого $\varepsilon > 0$ всегда можно указать такой полином $P_n(x)$, что

$$|f(x) - P_n(x)| < \varepsilon \quad \text{для любого } x \in [a, b].$$

Это знаменитый результат, но уже из другой области и для других целей. Равномерную аппроксимацию $f(x)$ на $[0, 1]$ (при $n \rightarrow \infty$) обеспечивают, например, *полиномы Бернштейна*:

$$P_n(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}. \quad \bullet \quad (3.14)$$

3.7 Контрпримеры и парадоксы

Об ощущении берегов. Когда плывёшь по течению в окружении теорем, необходимо чувствовать берега. То есть иметь представление, что там, за пределами оговоренных условий. Куда можно угодить, делая шаг в сторону. Поэтому примеры и контрпримеры не менее важны чем теоремы.

Ещё важнее понимать, обеспечивают ли «оговоренные условия» безопасность «рейса». Исключены ли аномалии. Не могут ли хлынуть в пробойну определений какие-нибудь несуразности. Дело в том что некоторые предположения теорем, находясь в тени, кажутся несущественными, и выводы мерещатся справедливыми без них. Тогда и возникает потребность в контрпримерах.

Ловушка непрерывности. Всякое определение минирует математику. Казалось бы, что может быть проще и естественнее понятия непрерывной функции. Но в расщелину соответствующего определения просачиваются многочисленные уродцы.

Первые признаки того, что непрерывность вмещает в себя всякую чертовщину, возникают довольно просто. График везде непрерывной¹⁵ функции $y = x \sin \frac{1}{x}$ невозможно нарисовать, а график непрерывной в нуле функции

$$f(x) = \begin{cases} 0, & \text{если } x \text{ рационально,} \\ x, & \text{если } x \text{ иррационально} \end{cases}$$

всюду дыряв. Ещё экзотичнее *функция Римана*



$$f(x) = \begin{cases} \frac{1}{n}, & \text{если } x = \frac{m}{n}, \text{ дробь } \frac{m}{n} \text{ несократима,} \\ 0, & \text{если } x \text{ иррационально,} \end{cases} \quad (3.15)$$

В подобное загода трудно поверить. Ибо можно ли заранее образовать *функцию непрерывную в иррациональных и разрывную в рациональных точках*¹⁶.

◀ Известно, что иррациональное x_0 невозможно хорошо приблизить рациональной дробью $\frac{m}{n}$ с ограниченным знаменателем¹⁷. Поэтому чем меньше окрестность x_0 , тем больше знаменатели входящих туда дробей $\frac{m}{n}$ и тем меньше значения $f\left(\frac{m}{n}\right) = \frac{1}{n}$. В итоге $f(x) \rightarrow f(x_0) = 0$ при $x \rightarrow x_0$. ▶

Далее можно вспомнить *кривую Пеано* — непрерывный образ отрезка $(x(t), y(t))$, $t \in [0, 1]$, заполняющий весь квадрат [2, т. 12]. Конечно, серьёзного удара по непрерывности перечисленные примеры не наносят. Да, ёмкость понятия оказывается больше, чем хотелось бы, однако неприятности терпимы. Не ка-

¹⁵Здесь и далее подразумевается непрерывное доопределение функций с «выколотыми точками».

¹⁶Тем более что, наоборот, «*функцию непрерывную в рациональных и разрывную в иррациональных точках*» построить невозможно.

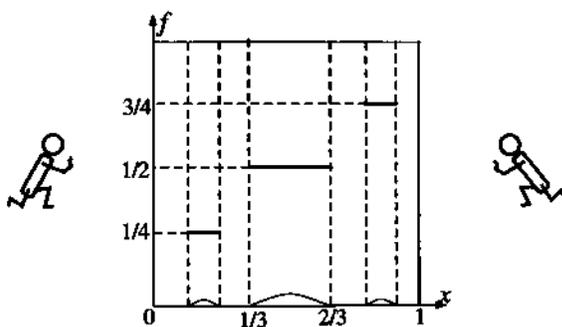
¹⁷См. например, Кац М., Улам С. Математика и логика: Ретроспектива и перспективы. «Мир».

гастрофа, в конце концов. Но лиха беда начало.



• **Канторова лестница** бьёт по мозгам уже сильнее. Строится она на базе канторова множества C , получаемого выбрасыванием третей из сегмента $[0, 1]$. Сначала $[0, 1]$ делится на три равные части, и средняя (интервал) удаляется. С каждой из оставшихся частей повторяется аналогичная операция — и так до бесконечности. В пределе от $[0, 1]$ остается как раз множество C . При этом длина выброшенных третей равна $\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots = 1$, т. е. «вся длина» $[0, 1]$ выбрасывается. Но оставшееся множество C оказывается равномощно континууму, что и производит главный сюрприз в теории множеств.

Теперь далее. Пусть функция $f: [0, 1] \rightarrow [0, 1]$ на замыкании каждого выбрасываемого (в процессе построения канторова множества) интервала — принимает значение, равное середине этого интервала. Вот так выглядит график $f(x)$ после двух шагов.



Доопределение $f(x)$ в точках второго рода по непрерывности¹⁸ даёт непрерывную монотонную функцию $f(x)$, производная которой

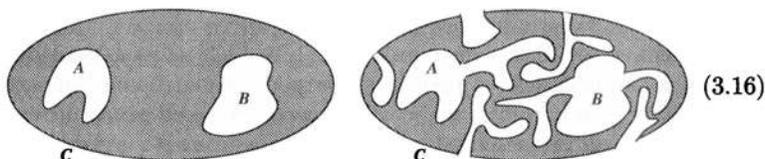
¹⁸Концы выброшенных интервалов в канторовом множестве C называют точками первого рода, остальные — второго рода. В точках второго рода \tilde{x} полагается $f(\tilde{x}) = \lim f(x_k)$, где $x_k \rightarrow \tilde{x}$ и все x_k — точки первого рода.

почти всюду равна нулю, тем не менее $f(0) = 0$, $f(1) = 1$. Другими словами, $f(x)$ почти нигде на $[0, 1]$ не растет, успевая ощутимо вырасти на множестве нулевой меры (несмотря на непрерывность).

• **Знаменитый пример Брауэра.** Этот удар уже не по мозгам, а едва ли не по устройству Вселенной.

Зададимся дурацким на вид вопросом. Могут ли на плоскости три области иметь общую границу? Не общий участок — а одну и ту же границу. Похоже на бред, но ответ — положительный.

Идея конструкции довольно проста. Пусть в море C есть остров, на острове два озера — A и B . На сухопутной части острова выделим ε -сеть¹⁹ S_ε . Затем от каждого озера и от моря к каждой точке S_ε пророем канал, не доводя его до этой точки на расстояние $\frac{\varepsilon}{2}$. В духе рисунка (3.16) справа.



На оставшуюся часть суши поместим $\frac{\varepsilon}{2}$ -сеть, и к точкам $S_{\varepsilon/2}$ пророем каналы, не доходящие до точек сети на $\frac{\varepsilon}{4}$. Потом накроем сушу $\frac{\varepsilon}{4}$ -сетью и так далее. Понятно, что в пределе области A, B, C разрастутся до областей $A^\infty, B^\infty, C^\infty$ с общей границей Γ . Граница Γ — это все, что останется от суши. А если изначально взять остров с миллионом озер — получится *пример миллиона областей с общей границей*.

Такую галиматью возникает соблазн объявить фантазмагорией ума. Дескать, чего не бывает в больном воображении. Тем более удивительно, что картина нескольких областей с общей границей встречается не только в фантазиях, но и на практике. Итерационная процедура

$$z_{k+1} = z_k - \frac{z_k^3 - 1}{3z_k^2} \quad (3.17)$$

¹⁹Эпсилон-сетью множества X называют такое его подмножество S_ε , что для любого $x \in X$ можно указать $s \in S_\varepsilon$, удалённое от x не более чем на ε .

на комплексной плоскости вычисляет корень кубический из единицы, каковых имеется три,

$$1, \quad \frac{-1 + i\sqrt{3}}{2}, \quad \frac{-1 - i\sqrt{3}}{2}, \quad (3.18)$$

и есть, соответственно, три области притяжения A , B , C . Процесс (3.17) сходится к одному из корней (3.18) в зависимости от того, какой области принадлежит z_0 . Пусть Γ_A , Γ_B , Γ_C обозначают границы областей A , B , C . Невероятно, но факт:



$$\Gamma_A = \Gamma_B = \Gamma_C, \quad (!)$$

т. е. области притяжения имеют одну и ту же границу²⁰.

Дифференциальная западня. Дифференцируемая функция обладает свойством Коши: по теореме Дарбу её производная принимает все промежуточные значения. Это подталкивает к выводу, что в случае дифференцируемости $f(x)$ на $[a, b]$ производная $f'(x)$ обязана быть непрерывной функцией, поскольку, мол, свойство Коши представляется эквивалентом непрерывности. *Неправильно то и другое.* Аргументы:

- Производная везде дифференцируемой функции

$$f(x) = x^2 \sin \frac{1}{x}, \quad (3.19)$$

равная нулю при $x = 0$ и

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x} \quad \text{при } x \neq 0$$

разрывна в точке $x = 0$, но обладает свойством Коши²¹.

²⁰См. п. 8.5 в книге Кроновер Р. М. Фракталы и хаос в динамических системах. Основы теории. М.: Постмаркет, 2000.

²¹Напомним, функции мы доопределяем по непрерывности. В данном случае $f(0) = 0$, поскольку $f(x) \rightarrow 0$ при $x \rightarrow 0$.

Пример (3.19) задаёт эталон возможных неприятностей в дифференциальном исчислении. При выборе в качестве мишени точки $x = 0$ ядром замысла обычно является произведение двух функций, одна из которых подходящим образом обнуляется в нуле, а другая, типа $\sin \frac{1}{x}$, быстро колеблется, «ускоряясь» по мере приближения к $x = 0$. Если частоты колебаний не хватает, то вместо $\sin \frac{1}{x}$ берется что-нибудь вроде $\sin(1/x^k)$. Возникающий ассортимент довольно широк.

- Везде дифференцируемая функция

$$f(x) = x^2 \sin \frac{1}{x^2}$$

имеет неограниченно разрывную в нуле производную, $f'(0) = 0$ и

$$f'(x) = 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} \quad \text{при } x \neq 0.$$

- Функция

$$f(x) = x + x^2 \sin \frac{1}{x^2}$$

имеет производную $f'(0) = 1$, но не монотонна в окрестности нуля.

- Функция

$$f(x) = 2x^2 + x^2 \sin \frac{1}{x^2}$$

имеет в нуле строгий минимум и равную нулю производную, но её производная в сколь угодно малой окрестности нуля принимает как положительные, так и отрицательные значения, сколь угодно большие по модулю.

Такие примеры способствуют поддержанию бдительности, показывая не только чего надо опасаться, но и при каких обстоятельствах. Эпизодически поговаривают о чрезмерной либеральности определения производной, которое считает дифференцируемыми в нуле функции типа $x^2 \sin(1/x)$. Дескать, скверная природа таких функций портит столбовую дорогу матанализа компрометирующими закоулками. Однако проблема дифференцируемости абсолютного ответа не имеет и иметь не может. Как и непротиворечивость гражданского или уголовного кодекса.

Глава 4

Интеграл

*A not getting what you want
is sometimes a stroke of luck.*
Dalai Lamas

4.1 Первообразная

4.1.1 *Операция, обратная дифференцированию, называется интегрированием. Функцию $F(x)$, такую что*

$$F'(x) = f(x),$$

или эквивалентно $dF(x) = f(x)dx$, называют первообразной функции $f(x)$, или интегралом от $f(x)$.

Если $F(x)$ — первообразная $f(x)$, то и $F(x) + C$ — первообразная $f(x)$, поскольку производная константы C равна нулю.

4.1.2 *Совокупность всех первообразных $f(x)$ называют неопределённым интегралом и обозначают*

$$\int f(x)dx. \tag{4.1}$$

Под знаком интеграла \int в (4.1) стоит дифференциал первообразной, $f(x)dx = dF(x)$. Поскольку, например, $d \sin x = (\sin x)' dx = \cos x dx$, то

$$\int \cos x dx = \sin x. \quad \Omega$$

Точнее говоря, $\int \cos x dx = \sin x + C$, но константу мы будем опускать.

Таблица производных элементарных функций легко трансформируется в таблицу неопределённых интегралов.

$$\int e^x dx = e^x, \quad \int a^x dx = a^x / \ln a, \quad a > 0, a \neq 1.$$

$$\int x^\lambda dx = \frac{1}{\lambda+1} x^{\lambda+1}, \quad \lambda \neq -1.$$

$$\int \frac{dx}{x} = \ln |x|, \quad \int \frac{dx}{\sqrt{1-x^2}} = \arcsin x, \quad \int \frac{dx}{1+x^2} = \arctg x.$$

Не будем далее ломиться в открытую дверь. Некоторое количество формул, конечно, полезно иметь в запасе, дабы не ходить к соседу за справкой в процессе решения задач.

Что касается потенциальной полезности неопределённого интеграла, то первое, что приходит в голову, это задачи, в которых даны соотношения скоростей, и неизвестны — пройденные пути. Не обязательно о километрах речь. Скажем, численность популяции $x(t)$ растёт со скоростью $\dot{x} = kx$. Конечно, интеграл сразу бы решал вопрос в ситуации

$$\dot{x} = \varphi(t) \Rightarrow x(t) = \int \varphi(t) dt.$$

Но и в случае $\dot{x} = kx$ интегрирование спасает¹:

$$\begin{aligned} \frac{dx}{dt} = kx &\Rightarrow \frac{dx}{x} = k dt \Rightarrow \int \frac{dx}{x} = \int k dt \Rightarrow \\ &\Rightarrow \ln x = kt + C \Rightarrow x(t) = x(0)e^{kt}, \end{aligned}$$

¹Подобные трюки работают также в более сложных ситуациях.

где константа интегрирования² C выбрана из условия $e^C = x(0)$.

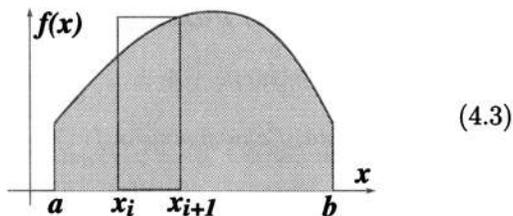
Простейшие свойства первообразной легко следуют из свойств производных.

- $\int [f(x) \pm g(x)] dx = \int f(x) dx \pm \int g(x) dx$ (?)
- $\int af(x) dx = a \int f(x) dx$ (?)
- $\int dx = x, \int df(x) = f(x)$ (?)

4.2 Определённый интеграл

Интегрирование возникло и какое-то время развивалось независимо от дифференцирования. Источником послужила задача определения площади под графиком функции $y = f(x)$. В основу решения была положена естественная аппроксимационная идея. Отрезок $[a, b]$ разбивается на n сегментов $[x_i, x_{i+1}]$ длины Δx_i , где $x_0 = a, x_n = b$, рис. (4.3). На каждом i -м отрезке выбирается произвольная точка ξ_i , и рассматривается сумма

$$\sigma = \sum_{i=0}^{n-1} f(\xi_i) \Delta x_i. \quad (4.2)$$



- Предел суммы (4.2), если таковой существует и не зависит от разбиения $[a, b]$ и выбора точек ξ_i , при стремлении к нулю максимальной длины Δx_i — называют определённым интегралом $f(x)$ от a до b и обозначают $\int_a^b f(x) dx$.

²О наличии которой надо вовремя вспоминать.

Конечно, в таком определении слишком много «если». Кто будет производить разбор полётов в каждом конкретном случае? Поэтому идут обычно по несколько иному пути. На каждом промежутке $[x_i, x_{i+1}]$ берутся точная нижняя m_i и точная верхняя M_i границы $f(x)$, и вводятся в рассмотрение две суммы Дарбу (нижняя и верхняя)

$$s = \sum_{i=0}^{n-1} m_i \Delta x_i, \quad S = \sum_{i=0}^{n-1} M_i \Delta x_i.$$

В силу $m_i \leq f(\xi_i) \leq M_i$ имеем $s \leq \sigma \leq S$. Поэтому для существования предела σ достаточно, чтобы суммы Дарбу сходились к одному пределу, т. е.

$$S - s \rightarrow 0.$$

Общий предел s и S , если таковой существует, и называют *определённым интегралом*. Оба определения эквивалентны, но второе легче проверять. При этом в случае существования интеграла $\int_a^b f(x)dx$ в указанном смысле функция $f(x)$ называется *интегрируемой по Риману*, а сам интеграл — *интегралом Римана*.

4.2.1 Теорема. Если функция $f(x)$ непрерывна на $[a, b]$, то интеграл $\int_a^b f(x)dx$ существует.

◀ Доказательство обеспечивает теорема Кантора 2.8.2, гарантирующая равномерную непрерывность $f(x)$. Поэтому по любому $\varepsilon > 0$ можно указать такое $\delta > 0$, что $\Delta x_i < \delta \Rightarrow M_i - m_i < \varepsilon$ для всех i , откуда

$$\sum_{i=0}^{n-1} (M_i - m_i) \Delta x_i < \varepsilon \sum_{i=0}^{n-1} \Delta x_i = \varepsilon(b - a),$$

что даёт $S - s \rightarrow 0$. ▶

В случае кусочной непрерывности $f(x)$ (конечного числа разрывов) результат остается в силе, поскольку³

$$\int_a^b = \int_a^{c_1} + \dots + \int_{c_i}^{c_{i+1}} + \dots + \int_{c_n}^b,$$

где c_i — точки разрыва.

4.3 Взаимосвязь интегралов

Если рассмотреть интеграл

$$\Phi(x) = \int_a^x f(t) dt$$

как функцию верхнего предела, то

$$\Delta \Phi = \Phi(x + \Delta x) - \Phi(x) = \int_x^{x+\Delta x} f(t) dt = f(\xi) \Delta x + o(\Delta x),$$

что приводит к $\Phi'(x) = f(x)$, т. е. Φ — первообразная.

◀ Вот более аккуратное рассуждение. Обозначим через ξ_{min} и ξ_{max} наименьшее и наибольшее значение функции $f(x)$ на отрезке $[x, x + \Delta x]$. Очевидно, площадь $\Delta \Phi$ заключена между площадями прямоугольников $\xi_{min} \Delta x$ и $\xi_{max} \Delta x$, поэтому

$$\xi_{min} \leq \frac{\Delta \Phi}{\Delta x} \leq \xi_{max}.$$

При $\Delta x \rightarrow 0$ $\xi_{min} \rightarrow f(x)$, $\xi_{max} \rightarrow f(x)$, в силу непрерывности $f(x)$. Тогда по теореме «о трёх собачках» $\frac{\Delta \Phi}{\Delta x} \rightarrow f(x)$, т. е. $\Phi'(x) = f(x)$. ▶

³ Аддитивность интеграла $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$ вытекает из определения.

Таким образом, *площадь* $\Phi(x)$ *под графиком* $y = f(x)$ — это *первообразная функции* $f(x)$. Изменение точки отсчёта a добавляет к $\Phi(x)$ некоторую константу. От любой другой первообразной $F(x)$ функция $\Phi(x)$ отличается на постоянную величину, поэтому всегда⁴

$$\Phi(b) - \Phi(a) = F(b) - F(a) = \int_a^b f(x)dx,$$

что численно равно площади под графиком $y = f(x)$ на промежутке $[a, b]$. Если $f(x)$ на $[a, b]$ меняет знак, то из построения ясно, что площади фигур между графиком и осью x -ов засчитываются со знаком «плюс» там, где $f(x) > 0$, и со знаком «минус» там, где $f(x) < 0$.

Дифференцируя $\Phi(\varphi(x))$ по формуле дифференцирования сложной функции, имеем

$$\frac{d}{dx} \int_a^{\varphi(x)} f(t)dt = f(\varphi(x))\varphi'(x).$$



При манипулировании определёнными интегралами нередко требуются оценки в виде неравенств. Вот два очень простых и часто используемых факта.

- Если $m \leq f(x) \leq M$ на $[a, b]$, то

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a). \quad (?)$$

⁴Заметим, что разность $F(b) - F(a)$ часто записывают в виде $F(x) \Big|_a^b$.

• Пусть на $[a, b]$ функция $g(x)$ знакопостоянна и $m \leq f(x) \leq M$.

Тогда

$$\int_a^b f(x)g(x)dx = \mu \int_a^b g(x)dx$$

при некотором $\mu \in [m, M]$. (?)

$$\bullet \int_a^b f(x)dx = - \int_b^a f(x)dx \quad (?)$$

$$\bullet \int_a^b kf(x)dx = k \int_a^b f(x)dx \quad (?)$$

$$\bullet \int_a^b [f(x) \pm g(x)]dx = \int_a^b f(x)dx \pm \int_a^b g(x)dx \quad (?)$$

$$\bullet \int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \quad (?)$$

$$\bullet \int_a^b f(x)dx \geq 0, \text{ если } f(x) \geq 0 \quad (?)$$

4.4 Техника интегрирования

*Whether you think you can or
whether you think you can't, you're right!*

Henry Ford

Замена переменной. Ретроспективно оглядываясь на дифференцирование, легко выделить в ремесле взятия производных две составляющих. Во-первых, нужно знать или иметь перед глазами таблицу производных простейших функций. Во-вторых, нужно уметь комбинировать элементы таблицы, опираясь на *свойства производных*: производная суммы, произведения, частного, композиции функций.



Со взятием неопределённых интегралов — такая же история. «Таблица», собственно, та же самая, разве что вывернутая наизнанку. Да и «приёмы комбинирования» те же, но «киноплёнка» крутится в обратном направлении. Впечатления, соответственно, другие. Пересматривать приходится заново.

• Интегрирование, как и дифференцирование, — линейная операция, т. е. если $F(x)$ — первообразная функции $f(x)$, а $G(x)$ — первообразная $g(x)$, то

$$\int [\alpha f(x) + \beta g(x)] dx = \alpha F(x) + \beta G(x)$$

при любых $\alpha, \beta \in \mathbb{R}$.

• Широко распространены переходы к интегрированию по другой переменной, в основе которых лежит формула

$$\int f(g(x))g'(x)dx = F(g(x)), \quad (4.4)$$

где $F(g)$ — первообразная $f(g)$.

Равенство (4.4) представляет собой опять-таки «вывернутое наизнанку» правило дифференцирования сложной функции. Дифференцируя в (4.4) $F(g(x))$, получаем

$$F'_g(g(x))g'(x) = f(g(x))g'(x),$$

т. е. как раз подынтегральное выражение.

Фокус практического использования формулы (4.4) для $\int \varphi(x)dx$ заключается в изобретении представления

$$\int \varphi(x)dx = \int f(g(x))g'(x)dx$$

с функцией f , первообразная которой известна или легко вычисляется. Часто это получается само собой.

• Например,

$$\int \sin^3 x \cos x dx = \int \sin^3 x d \sin x = \frac{1}{4} \sin^4 x$$

или

$$\Phi(x) = \int \frac{2x dx}{1+x^4} = \int \frac{dx^2}{1+x^4} = \operatorname{arctg} x^2.$$

За кадром-то здесь фигурирует замена $x^2 = t$, далее берётся интеграл $\int \frac{dt}{1+t^2}$, после чего исполняется возврат к переменной x . Понятно, что сию кухню проще держать в голове, не оставляя следов.

• Замена, конечно, не всегда сама напрашивается. Иногда приходится повозиться. Не так легко, например, додуматься, что проблему взятия интеграла $\int \frac{dx}{\sqrt{1+x^2}}$ решает замена $t = \sqrt{1+x^2} + x$. Возведение $t - x = \sqrt{1+x^2}$ в квадрат даёт

$$x = \frac{t^2 - 1}{2t}, \quad \sqrt{1+x^2} = \frac{t^2 + 1}{2t}, \quad dx = \frac{t^2 + 1}{2t^2} dt,$$

откуда

$$\int \frac{dx}{\sqrt{1+x^2}} = \int \frac{dt}{t} = \ln |t| = \ln |\sqrt{1+x^2} + x|.$$

Последний пример демонстрирует, что трюк (1) может быть лишь элементом более сложных манипуляций.

• Если $F(x)$ — первообразная $f(x)$, то

$$\int f(ax+b) dx = \frac{1}{a} \int f(ax+b) d(ax+b) = \frac{1}{a} F(ax+b).$$

Интегрирование по частям. Обращение правила дифференцирования произведения приводит к формуле *интегрирования по частям*

$$\int u dv = uv - \int v du, \quad (4.5)$$

проверяемой дифференцированием с учётом

$$(uv)' = u'v + uv'.$$

Диапазон применения (4.5) весьма широк. Осваивать инструмент надо на примерах

- $\int \ln x dx = x \ln x - \int x d \ln x = x \ln x - \int dx = x \ln x - x,$
- $\int x \cos x dx = \int x d \sin x = x \sin x - \int \sin x dx = x \sin x + \cos x.$

Возможны и более крутые выражи

- $\int e^x \cos x dx = \int e^x d \sin x = e^x \sin x - \int e^x \sin x dx =$
 $= e^x \sin x + \int e^x d \cos x = e^x \sin x + e^x \cos x - \int e^x \cos x dx,$

откуда $2 \int e^x \cos x dx = e^x (\sin x + \cos x).$

В том же духе:

- $\int \sqrt{\alpha^2 - x^2} dx = x \sqrt{\alpha^2 - x^2} - \int \frac{\alpha^2 - x^2 - \alpha^2}{\sqrt{\alpha^2 - x^2}} dx =$
 $= x \sqrt{\alpha^2 - x^2} - \int \sqrt{\alpha^2 - x^2} dx + \alpha^2 \int \frac{dx}{\sqrt{\alpha^2 - x^2}},$

откуда

$$\int \sqrt{\alpha^2 - x^2} dx = \frac{x}{2} \sqrt{\alpha^2 - x^2} + \frac{\alpha^2}{2} \arcsin \frac{x}{\alpha}.$$

Всё это к существу интегрирования имеет весьма отдалённое отношение. В основном, это касается жонглирования и комбинирования простых фактов, что широко распространено во всех сферах деятельности. Часть населения устремляется к задачам олимпиадного уровня, но и в рядовых ситуациях необходима определённая ловкость.

- $\int \frac{dx}{x^2 - a^2} = \int \frac{1}{2a} \left(\frac{1}{x - a} - \frac{1}{x + a} \right) dx =$
 $= \frac{1}{2a} (\ln |x - a| - \ln |x + a|) = \frac{1}{2a} \ln \left| \frac{x - a}{x + a} \right|.$

Следующий пример несколько сложнее.

$$\begin{aligned} \bullet \int \frac{dx}{\sin x} &= \int \frac{d(x/2)}{\sin(x/2) \cos(x/2)} = \int \frac{\cos(x/2) \frac{d(x/2)}{\cos^2(x/2)}}{\sin(x/2)} = \\ &= \int \frac{d \operatorname{tg}(x/2)}{\operatorname{tg}(x/2)} = \ln \left| \operatorname{tg} \frac{x}{2} \right|. \end{aligned}$$

Не следует забывать, что некоторые интегралы не выражаются через элементарные функции⁵, например,

$$\int \frac{\sin x}{x} dx, \quad \int \frac{dx}{\ln x}.$$

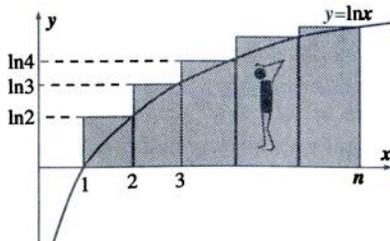
4.5 Прикладные задачи

Длины, площади

- Площадь под синусоидой

$$\int_0^{\pi} \sin x dx = -\cos x \Big|_0^{\pi} = -\cos \pi + \cos 0 = 2.$$

- **Формула Стирлинга.** Сумма $\ln n! = \ln 1 + \dots + \ln n$ равна сумме площадей заштрихованных прямоугольников,



⁵Тогда как дифференцирование не выводит из области элементарных функций.

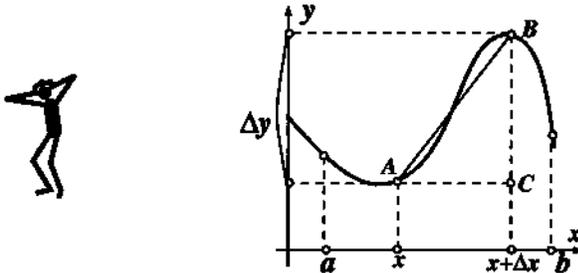
и может быть приближённо заменена площадью под кривой $\ln x$. Поэтому

$$\ln n! \simeq \int_1^n \ln x \, dx = x \ln x - x \Big|_1^n \simeq \ln \left(\frac{n}{e} \right)^n,$$

т. е. $n! \simeq (n/e)^n$. Прибавление площадей выступающих над кривой $\ln x$ треугольников приводит к более точному результату

$$n! \simeq \sqrt{n} \left(\frac{n}{e} \right)^n.$$

• *Длина дуги.* Пусть Δs обозначает длину дуги кривой $y = f(x)$ между точками A и B .



Из прямоугольного треугольника ABC

$$\Delta s \simeq \sqrt{\Delta x^2 + \Delta y^2} = \sqrt{1 + (\Delta y^2 / \Delta x^2)} \Delta x.$$

Поэтому $ds = \sqrt{1 + y'^2} dx$. Интегрирование дифференциала ds от a до b даёт формулу вычисления длины дуги на любом конечном участке.

Тестовый пример — длина s_R окружности $x^2 + y^2 = R^2$. На верхней полуокружности

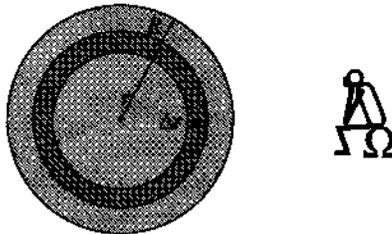
$$y = \sqrt{R^2 - x^2}, \quad y' = -\frac{x}{\sqrt{R^2 - x^2}}.$$

В результате

$$\frac{cR}{2} = \int_{-R}^R \sqrt{1 + \frac{x^2}{R^2 - x^2}} dx = R \int_{-R}^R \frac{dx}{\sqrt{R^2 - x^2}} = R \arcsin \frac{x}{R} \Big|_{-R}^R = \pi R.$$

Что касается площадей фигур, то они обычно строятся на базе покрытия фигур мелкими квадратиками с последующим предельным переходом, что приводит к *двойным интегралам*, см. далее. Метод достаточно универсален, но в конкретных случаях существуют более экономные приёмы, использующие специфику задачи, например симметрию. Скажем, разбиение на квадратики при вычислении площади круга на фоне известной формулы длины окружности — выглядело бы насмешкой над здравым смыслом. Действовать тут надо более рационально.

• Площадь круга радиуса R можно вычислять как сумму площадей тонких колец радиуса r и толщины Δr .



В свою очередь площадь кольца в первом приближении равна $2\pi r \cdot \Delta r$. Это определяет дифференциал площади круга $dS = 2\pi r dr$ при разбиении на кольцевые слои, что в итоге позволяет вычислить площадь круга

$$S_R = \int_0^R 2\pi r dr = \pi R^2.$$

Пример рассматривается не в качестве изобретения велосипеда, а как образец для подражания, демонстрирующий стереотип использования симметрии задачи.

• Найдём площадь поверхности вращения⁶ при вращении кривой $y = f(x)$ вокруг оси x . Малый элемент длины дуги ds вращается вокруг оси x на расстоянии $y = f(x)$. Поэтому площадь, описываемая элементом ds при полном обороте, равна

$$dS = 2\pi y ds = 2\pi y \sqrt{1 + y'^2} dx,$$

что после интегрирования в тех или иных пределах даёт желаемый «криволинейный» результат. Вращение полуокружности, например, порождает сферу. Площадь вычисляется по найденному рецепту:

$$S = 2\pi \int_{-R}^R y \sqrt{1 + y'^2} dx = 2\pi \int_{-R}^R \sqrt{R^2 - x^2} \sqrt{1 + \frac{x^2}{R^2 - x^2}} dx = 4\pi R^2.$$

Вычисление моментов инерции.

При вращении механического тела вокруг некоторой фиксированной оси с переменной угловой скоростью $\omega(t) = \dot{\varphi}(t)$ каждая частица массы Δm_i , находящаяся от оси на расстоянии r_i , движется со скоростью $v_i = r_i \omega$ и ускорением

$$\dot{v}_i = r_i \dot{\omega} = r_i \ddot{\varphi}.$$

Умножая закон движения каждой частицы $\Delta m_i \dot{v}_i = F_i$ на r_i и суммируя по i , после перехода к пределу при $\Delta m_i \rightarrow 0$ получим

$$J \ddot{\varphi} = M,$$

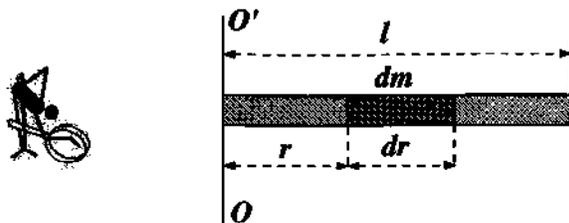
где M — результирующий момент⁷ действующих сил, J — момент инерции тела, представляющий собой интеграл $\int_V r^2 dm$, в котором суммирование производится по объёму тела.

Найдем моменты инерции некоторых стандартных тел.

⁶Мы пока делаем вид, что все просто, хотя здесь есть подводные камни, см. «парадокс Шварца».

⁷Обратим внимание, что момент сил, также как угловая скорость, — это вектор, см. «Векторный анализ».

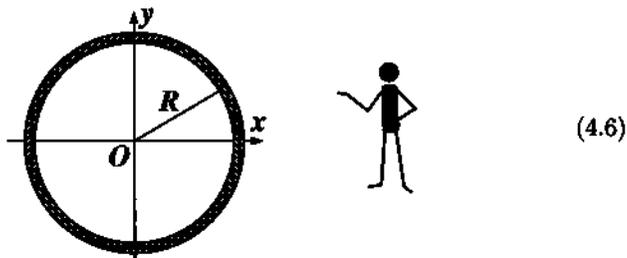
• *Однородный стержень длины l . Ось вращения OO' проходит перпендикулярно стержню через его конец.*



Элемент длиной dr имеет массу $dm = m \frac{dr}{l}$. Поэтому

$$J = \int_0^l r^2 dm = \int_0^l \frac{m}{l} r^2 dr = \frac{ml^2}{3}.$$

• *Обруч при условии, что ось вращения проходит через центр и лежит в плоскости обруча.*



Из соображений симметрии ясно, что моменты инерции относительно перпендикулярных осей x, y равны, $J_x = J_y$, т. е.

$$J_x = \int x^2 dm, \quad J_y = \int y^2 dm,$$

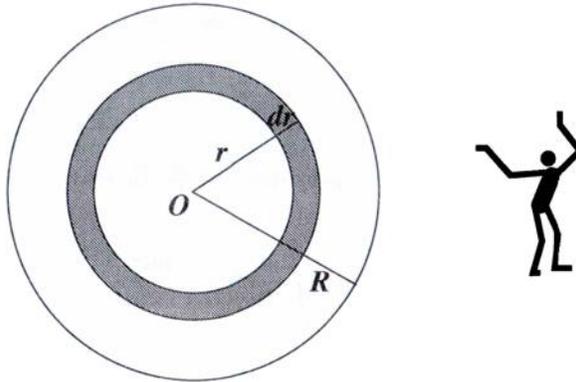
откуда

$$2J = J_x + J_y = \int (x_i^2 + y_i^2) dm = \int R^2 dm = mR^2,$$

что даёт

$$J = \frac{mR^2}{2}.$$

• **Однородный диск** радиуса R . Ось проходит через центр перпендикулярно плоскости диска.



Выделим из диска обруч радиуса r и ширины dr . Его масса равна

$$dm = m \frac{dS}{S} = m \frac{2rdr}{R^2}, \quad S = \pi R^2.$$

Поэтому

$$J = \frac{2m}{R^2} \int_0^R r^3 dr = \frac{mR^2}{2}.$$

• **Однородный диск** радиуса R относительно диаметра (рисунок тот же, но теперь ось лежит в плоскости диска). Момент инерции выделенного обруча (4.6) относительно диаметра равен $dJ = dm r^2 / 2$. Отсюда

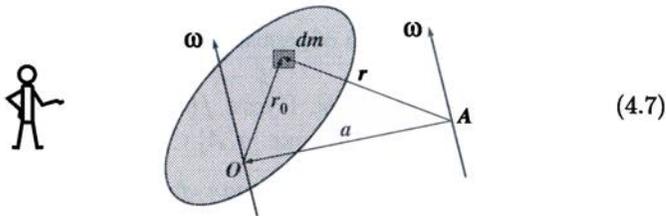
$$J = \frac{1}{2} \int_0^R r^2 dm = \frac{m}{R^2} \int_0^R r^3 dr = \frac{mR^2}{4}.$$

• Образовавшуюся небольшую коллекцию моментов инерции имеет смысл пополнить следующим результатом, который значительно расширяет ассортимент.

4.5.1 Теорема Штейнера. Если момент инерции J_0 относительно оси, проходящей через центр тяжести тела, известен, то момент инерции J_A относительно параллельной оси, отстоящей от первой на расстоянии a , определяется формулой

$$J_A = J_0 + ma^2.$$

◀ Ситуацию иллюстрирует рисунок

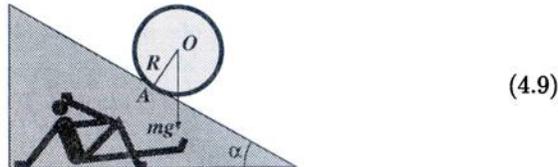


Подстановка в $J = \int r^2 dm$ значения⁸

$$r = r_0 + a \Rightarrow r^2 = r_0^2 + a^2 + 2r_0 \cdot a, \quad (4.8)$$

с учётом $\int r_0 dm = 0$ даёт нужный результат. ▶

• С каким ускорением полый цилиндр радиуса R скатывается с наклонной плоскости, рис. (4.9)?



Уравнение моментов относительно мгновенной оси вращения в данном случае имеет вид

$$J_A \frac{d\omega}{dt} = mgR \sin \alpha.$$

⁸Соотношения в (4.8) векторные, $r_0 \cdot a$ — скалярное произведение.

В силу $v = \omega R$ и $J_A = 2mR^2$, ускорение цилиндра получается равным

$$\dot{v} = \frac{1}{2}g \sin \alpha.$$

4.6 Несобственные интегралы

Бесконечный промежуток. Понятие определённого интеграла

$\int_a^b f(x)dx$ естественным образом обобщается на случай неограниченного промежутка $[a, b]$. Пусть $f(x)$ определена на $[a, \infty)$ и интегрируема на любом отрезке $[a, A] \subset [a, \infty)$, тогда:

4.6.1 Несобственный интеграл функции $f(x)$ от a до ∞ определяется как предел (конечный или бесконечный)

$$\int_a^{\infty} f(x)dx = \lim_{A \rightarrow \infty} \int_a^A f(x)dx \quad (4.10)$$

Если предел конечен, то говорят, что интеграл (1) *сходится*. Если бесконечен, то — *расходится*.

Аналогично определяется интеграл

$$\int_{-\infty}^a f(x)dx = \lim_{A \rightarrow -\infty} \int_A^a f(x)dx. \quad (4.11)$$

Наконец, если существуют оба интеграла (4.10) и (4.11), то

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^a f(x)dx + \int_a^{\infty} f(x)dx.$$

Примеры

• При $p \neq 1$

$$\int_1^A \frac{dx}{x^p} = \frac{x^{1-p}}{1-p} \Big|_1^A = \frac{1}{1-p} (A^{1-p} - 1),$$

в случае $p = 1$

$$\int_1^A \frac{dx}{x} = \ln x \Big|_1^A = \ln A.$$

В результате при $A \rightarrow \infty$ имеем

$$\int_1^{\infty} \frac{dx}{x^p} = \infty, \quad \text{если } 0 < p \leq 1$$

и

$$\int_1^{\infty} \frac{dx}{x^p} = \frac{1}{p-1}, \quad \text{если } p > 1$$

$$\bullet \int_0^{\infty} \frac{dx}{1+x^2} = \operatorname{arctg} x \Big|_0^{\infty} = \frac{\pi}{2}.$$

• По закону гравитационного притяжения вес mg тела массы m равен⁹

$$F = \gamma \frac{Mm}{R_0^2}, \quad (4.12)$$

где M — масса Земли, R_0 — радиус Земли.

Если потенциальную энергию бесконечно удалённого от Земли тела считать нулевой, $u(\infty) = 0$, то, интегрируя работу FdR по удалению тела в бесконечность, получаем

$$u(R) = \int_{\infty}^R F dR = -\gamma \frac{Mm}{R} \Big|_{\infty}^R = -\gamma \frac{Mm}{R}.$$

Вторая космическая скорость v_{∞} (позволяющая улететь в бесконечность) определяется из условия

$$u(\infty) - u(R_0) = \gamma \frac{Mm}{R_0} = \frac{mv_{\infty}^2}{2} \Rightarrow v_{\infty} = \sqrt{2\gamma \frac{M}{R_0}}.$$

⁹Хотя закон тяготения Ньютона декларирует (4.12) для точечных масс, он справедлив и в случае притяжения однородных шаров.

- Работа газа при расширении от объёма V_0 до V_1 равна

$$A = \int_{V_0}^{V_1} p dV.$$

Допустим, газ расширяется адиабатически (без теплообмена с окружающей средой) до бесконечности. Какая работа будет произведена?

При адиабатическом расширении взаимосвязь давления с объёмом определяется законом Пуассона

$$pV^k = \text{const}, \quad k = c_p/c_v \quad \text{т. е.} \quad pV^k = p_0V_0^k$$

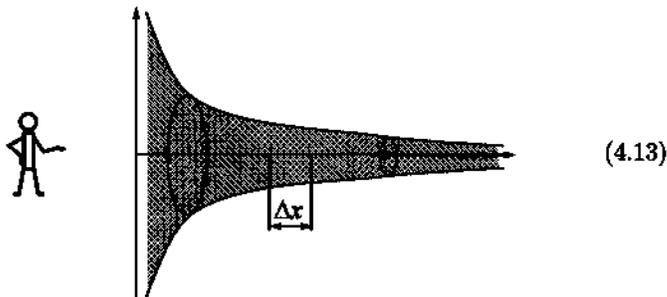
Следовательно,

$$A_{\text{max}} = \int_{V_0}^{\infty} \frac{p_0 V_0^k}{V^k} dV = \frac{p_0 V_0^k}{1-k} V^{1-k} \Big|_{V_0}^{\infty} = \frac{p_0 V_0}{k-1}.$$

- При вращении графика кривой $y = \frac{1}{x}$ в диапазоне $(1, \infty)$ вокруг оси x образуется тело вращения (4.13), объём которого равен

$$V = \int_1^{\infty} \frac{\pi dx}{x^2} = \pi,$$

что получается как результат предельного суммирования объёмов колец радиуса $y = \frac{1}{x}$ и толщины Δx ($\Delta V = \pi y^2 \Delta x$).

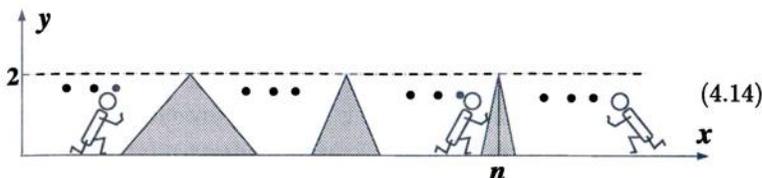


Интересно, что объём тела ограничен, но вертикальное плоское сечение, проходящее через ось x , имеет бесконечную площадь (!)

$$S = 2 \int_1^{\infty} \frac{dx}{x} = \infty.$$

Факт лежит в основе *парадокса маляра*.

• Часто думают, что для сходимости интеграла $\int f(x)dx$ от положительной функции $f(x)$ требуется $f(x) \rightarrow 0$ при $x \rightarrow \infty$. Это, правда, больше психологический вопрос, чем математический.



На самом деле, если $f(x)$ равна нулю вне треугольных пиков (высотой $f(n) = 2$ и шириной основания $1/n^2$), рис. (4.14), то¹⁰

$$\int f(x)dx = \sum \frac{1}{n^2} < \infty. \quad (4.15)$$

Понятно, что, не нарушая сходимости, высота пиков может расти до бесконечности при подходящем убывании ширины.

Многие результаты о сходимости несобственных интегралов перекликаются с признаками сходимости числовых рядов. Особую роль играют *абсолютно сходящиеся интегралы* $\int f(x)dx$, сходящиеся вместе с $\int |f(x)|dx$. Понятна в этом контексте роль интегралов положительных функций.

• Если $0 \leq f(x) \leq g(x)$ при достаточно больших x , то *сходимость*

$$G = \int g(x)dx \text{ влечёт за собой } \text{сходимость } F = \int f(x)dx,$$

а *расходимость* F — *расходимость* G .

¹⁰Запись (4.15) интегралов и сумм без указания нижних пределов интегрирования/суммирования используется обычно для обозначения факта сходимости или расходимости.

- Если существует конечный предел

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = M > 0,$$

то оба интеграла F и G сходятся или расходятся одновременно. (?)

Интегралы от неограниченных функций.

4.6.2 Пусть теперь функция $f(x)$ определена на $[a, b)$, интегрируема на любом отрезке $[a, c] \subset [a, b)$ и уходит в бесконечность при $x \rightarrow b - 0$, т. е. при x стремящемся к b слева. Тогда несобственный интеграл $\int_a^b f(x) dx$ определяется как предел (конечный или бесконечный)

$$\int_a^b f(x) dx = \lim_{c \rightarrow b-0} \int_a^c f(x) dx \quad (4.16)$$

Как и прежде, если предел конечен, то говорят, что интеграл (4.16) *сходится*. Если бесконечен, то — *расходится*.

В указанных условиях точку b называют *особой*. Из контекста ясно, что представляет собой несобственный интеграл, если особой является точка a , обе точки a и b , либо некоторая точка внутри отрезка $[a, b]$.

Примеры.

- При $p \neq 1$

$$\int_c^1 \frac{dx}{x^p} = \frac{x^{1-p}}{1-p} \Big|_c^1 = \frac{1}{1-p} (1 - c^{1-p}),$$

в случае $p = 1$

$$\int_c^1 \frac{dx}{x} = \ln x \Big|_c^1 = -\ln c.$$

В результате при $c \rightarrow 0 + 0$ имеем

$$\boxed{\int_0^1 \frac{dx}{x^p} = \frac{1}{p-1}, \quad \text{если } 0 < p < 1} \quad \text{и} \quad \boxed{\int_0^1 \frac{dx}{x^p} = \infty, \quad \text{если } p \geq 1}$$

- $$\int_0^1 \ln x \, dx = x \ln x - x \Big|_0^1 = -1.$$
- $$\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \arcsin 1 - \arcsin 0 = \frac{\pi}{2}.$$
- $$\int_0^{\pi/2} \operatorname{tg} x \, dx = - \int_0^{\pi/2} \frac{d \cos x}{\cos x} = - \ln \cos x \Big|_0^{\pi/2} = \infty.$$

Глубоко ли надо копать, изучая интегралы? Копать надо, сколько положено — скажут в деканате. Там учебная программа. Возникшая, разумеется, вследствие заботы о нуждах Галактики. Но многое в той программе включено/выключено по причине неуправляемого хаоса. Один промахнулся, другой поддержал, у третьего не было времени отфильтровать и причесать.

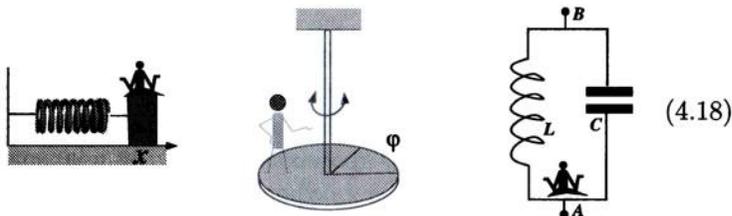
Но где провести черту в идеальной ситуации? В обстановке без нависающей дамокловым мечом системы образования, решившей всё за нас. Где остановиться, например, в изучении несобственных интегралов? Ограничиться сказанным выше или обратиться к *Физтенгольцу*? Понятно, универсального ответа быть не может. Кому-то и сказанного многовато, а другой чувствует «это моё!» — и отдаётся хронической тяге, забывая даже о продолжении рода. Тем не менее в ситуации «общего положения» есть всё-таки разумный ориентир. При изучении предмета планку имеет смысл поднимать на такую высоту, начиная с которой можно легко двигаться дальше. Или лучше сказать по-другому. Необходимо осваивать критическую массу фактов, способную далее расти самостоятельно в благоприятных условиях.

4.7 Дифференциальные уравнения

Дифференциальные описания. По мере изучения анализа и физики человек начинает понимать, что дифференциальное исчисление — это ДНК Вселенной. Возникает даже недоумение. Откуда у дифференциальных описаний всепроникающая сила? А секрет-то весьма прост. Процессы, разворачивающиеся во времени, не укладываются в голове одновременно, фотографически. О движении кометы или остывании утюга человеку, как и Создателю, легче судить «здесь и сейчас», а не как о траектории $x(t)$ целиком, на всей оси времени. Поэтому закон Ньютона¹¹,

$$m\ddot{x} = F, \quad \text{равносильно} \quad m\dot{v} = F, \quad (4.17)$$

описывает динамический процесс $x(t)$, или $v(t)$, в каждый данный момент, в каждой данной точке, озадачивая нас по поводу траекторий в целом.



Разумеется, в (4.17) необходимо указать, как сила F зависит от текущего положения точки $x(t)$ и/или скорости $\dot{x}(t) = v(t)$. В частности, «маятник на пружине» (4.18) движется в соответствии с уравнением $m\ddot{x} + kx = 0$, где k — коэффициент упругости пружины, m — масса грузика. Крутильные колебания описывает такое же уравнение

$$I\ddot{\varphi} + J\varphi = 0,$$

I — момент инерции диска, $J\varphi$ — крутящий момент стержня.

¹¹Равно как и другие физические законы.

В компанию (4.18) попадает и **колебательный контур**, представляющий собой электрический маятник. Уравнение колебаний снова имеет тот же вид,

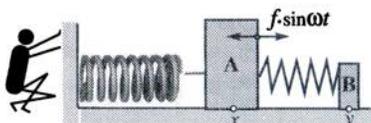
$$L\ddot{q} + \frac{q}{C} = 0,$$

названия коэффициентов другие: q — электрический заряд, L — индуктивность, C — емкость. Место «возвращающей силы» занимает электрическое напряжение q/C .

При наличии вязкого трения уравнение $m\ddot{x} + kx = 0$ усложняется, приобретая вид

$$m\ddot{x} + \beta\dot{x} + kx = 0. \quad (4.19)$$

Далее обстоятельства и факторы растут как снежный ком. Появляются задачи с несколькими взаимодействующими телами, и там возникают уже системы дифференциальных уравнений. Прикиньте хотя бы описание системы из двух грузиков:



Так что спрос на решение *дифференциальных уравнений* широк и многообразен¹², а соответствующая математическая дисциплина изящна и глубока. Но первые шаги (самые первые) имеет смысл сделать в рамках матанализа.

Простейшие уравнения. Решение дифференциального уравнения

$$\dot{x} = f(t) \quad (4.20)$$

сводится к обыкновенному интегрированию,

$$x(t) = \int f(t) dt + c,$$

¹²Уравнения электродинамики, диффузии, распространения волн и эпидемий, гидро- и аэродинамики — *дифференциальные*.

где c — произвольная константа, которая выбирается обычно из *начального условия* типа $x(0) = x_0$, в связи с чем часто сразу пишут

$$x(t) = x(0) + \int_0^t f(s) ds. \quad (4.21)$$

Вообще говоря, называть (4.20) дифференциальным уравнением даже язык не поворачивается. Дана производная, надо найти первообразную. Что тут нового? Но глубокие воды начинаются с маленьких ручейков. Так что лицезреть только (4.20), выводы делать рано.

Несколько сложнее уравнение, $\dot{x} = f(x)$, т. е. $\frac{dx}{dt} = f(x)$, которое переписывают в виде

$$\frac{dx}{f(x)} = dt,$$

и после интегрирования получают

$$\int \frac{dx}{f(x)} = \int dt = t + c, \quad (4.22)$$

что неявно определяет функцию $x(t)$. Всё благополучно разрешается, если интеграл в (4.22) берётся, что случается, когда задача «хорошо поставлена». Высшими силами, например.

В частности, для $\dot{x} = \lambda x$ легко получаем:

$$\int \frac{dx}{x} = \int \lambda dt, \text{ откуда } \ln x = \lambda t + \text{const, окончательно } x(t) = ce^{\lambda t}.$$

Заметим, что уравнение

$$\dot{x} = f(x, t) \quad (4.23)$$

в общем случае уже не решается ни в элементарных функциях, ни в квадратурах. То есть решение не удаётся записать даже с участием интегралов. При этом «нерешаемыми»¹³ оказываются самые простые с виду уравнения. Например, $\dot{x} = x^2 + t^2$.

¹³Но в этой «нерешаемости» нет ничего страшного, о чём будем говорить уже в курсе *дифференциальных уравнений*.

Рассмотрим для иллюстрации несколько примеров.

• В момент $t = 0$ двигатель лодки выключается. Какое расстояние лодка пройдет по инерции, если скорость в момент выключения — v_0 , а сила сопротивления воды пропорциональна скорости, $F = -\beta v$?

Интегрирование уравнения движения,

$$m \frac{dv}{dt} = -\beta v,$$

даёт

$$\int \frac{dv}{v} = -\frac{\beta}{m} \int dt \Rightarrow \ln v + C = -\frac{\beta}{m} t.$$

Константа определяется из условия $v(0) = v_0$. В итоге

$$v(t) = v_0 e^{-\frac{\beta}{m} t}.$$

Окончательно,

$$S_{max} = \int_0^{\infty} v dt = -\frac{v_0 m}{\beta} e^{-\frac{\beta}{m} t} \Big|_0^{\infty} = \frac{v_0 m}{\beta}.$$

• **Реактивное движение.** В неподвижной системе координат в некоторый момент масса ракеты равна M , скорость — v . Через малый промежуток времени Δt скорость ракеты увеличится на Δv , масса уменьшится на ΔM , причём ΔM будет иметь некоторую скорость V .

◀ Закон сохранения количества движения

$$Mv = (M - \Delta M)(v + \Delta v) + \Delta M \cdot V$$

после деления на ΔM и перехода к пределу при $\Delta M \rightarrow 0$ приводит к

$$M \frac{dv}{dM} + V_g = 0, \quad \text{или} \quad dv = -V_g \frac{dM}{M},$$

где $V_g = V - v$ — скорость истечения газов.

Интегрирование последнего уравнения даёт

$$v = \int dv = -V_g \int \frac{dM}{M} = -V_g \ln M + C.$$

Константа C определяется, например, из условия $M = M_0$ при $v = 0$, что даёт $v = V_g \ln \frac{M_0}{M}$, откуда следует формула Циолковского

$$\frac{M_0}{M} = e^{v/V_g}. \quad \blacktriangleright$$

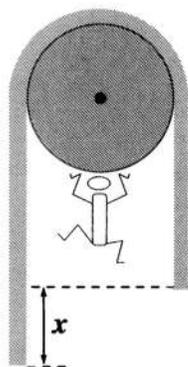
• Веревка длины l перекинута через блок. В начале один из свисающих концов длиннее другого на h . Найдите x в момент t . Трение отсутствует.

◀ При плотности веревки (на единицу длины) равной ρ вес длины x будет $\rho g x$, а масса всей веревки ρl . Закон движения (Ньютона) принимает вид

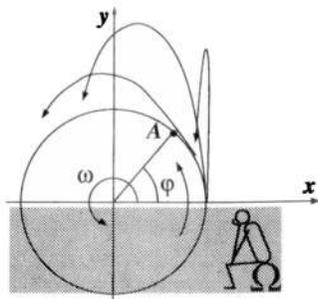
$$\rho l \frac{1}{2} \ddot{x} = \rho g x, \quad \text{т. е.} \quad \ddot{x} = \frac{2g}{l} x.$$

Подстановка $x = e^{\lambda t}$ приводит, с учётом начальных условий $x(0) = h$, $\dot{x}(0) = 0$, к результату

$$x(t) = \frac{h}{2} \left(e^{\sqrt{\frac{2g}{l}} t} + e^{-\sqrt{\frac{2g}{l}} t} \right). \quad \blacktriangleright$$



• Точильное колесо радиуса R , погруженное нижней половиной в воду, вращается с угловой скоростью ω . На какую максимальную высоту забрасываются брызги?



◀ Фиксируем на ободе колеса некоторую точку A . Её координата y меняется по закону

$$y = R \sin \omega t, \quad \text{откуда} \quad \dot{y} = R \omega \cos \omega t.$$

Брызги, срывающиеся с колеса в точке A , взлетают на высоту

$$h = y + \frac{\dot{y}^2}{2g} = R \sin \varphi + \frac{R^2 \omega^2}{2g} \cos^2 \varphi,$$

где $\varphi = \omega t$.

В точке максимума производная h по φ должна быть равна нулю,

$$\frac{dh}{d\varphi} = R \cos \varphi - \frac{R^2 \omega^2}{2g} 2 \cos \varphi \sin \varphi = 0,$$

откуда либо $\cos \varphi = 0$, либо $\sin \varphi = g/R\omega^2$.

Уравнения второго порядка. Как уже отмечалось, малые колебания маятников единообразно описываются уравнением

$$\ddot{x} + \omega^2 x = 0. \quad (4.24)$$

Общим решением (4.24) служит¹⁴

$$x(t) = A \sin(\omega t + \delta),$$

где A — амплитуда, δ — сдвиг по фазе, ω — круговая частота. Задание начальных условий $x(0)$, $\dot{x}(0)$ определяет выбор констант A , δ .

При наличии вязкого трения уравнение (4.19), т. е.

$$\ddot{x} + \gamma \dot{x} + \omega^2 x = 0, \quad (4.25)$$

сопряжено с некоторыми проблемами. Обойтись тригонометрическими функциями при решении (4.25) уже не удастся. Трение влечёт за собой энергетические потери, и колебания должны затухать. В качестве решения напрашивается $e^{\lambda t}$. Поскольку дифференцирование экспоненты даёт $(e^{\lambda t})' = \lambda e^{\lambda t}$, подстановка $e^{\lambda t}$ в (4.25) приводит к $(\lambda^2 + \gamma\lambda + \omega^2)e^{\lambda t} = 0$. Поэтому, если λ — корень уравнения

$$\lambda^2 + \gamma\lambda + \omega^2 = 0, \quad (4.26)$$

то $e^{\lambda t}$ удовлетворяет (4.25). Поэтому общее решение имеет вид

$$x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t},$$

где λ_1, λ_2 — корни (10).

В случае комплексных корней λ_1, λ_2 возникает впечатление, что нашла коса на камень. Но тревога оказывается ложной. Более того, все разрешается наилучшим образом — эффективно

¹⁴ Легко проверяется дифференцированием.

и просто. Если $e^{\lambda t}$ удовлетворяет уравнению (4.25) при $\lambda_{1,2} = \alpha \pm \pm i\beta$, — то уравнению (4.25) удовлетворяет как действительная часть функции

$$e^{\alpha t}(\cos \beta t \pm i \sin \beta t),$$

так и мнимая. Поэтому все действительные решения уравнения (4.25) исчерпываются семейством

$$x(t) = e^{\alpha t}(c_1 \cos \beta t + c_2 \sin \beta t).$$

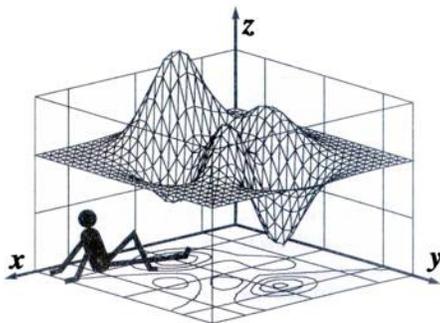
Глава 5

Функции нескольких переменных

*Инстинкт не даёт ходу скороспелым умозаключениям.
Иначе голова окажется полна идеями,
которые цитрамоном не лечатся.*

5.1 В двух словах о векторах

Если x, y — длина, ширина прямоугольника, то площадь $z = xy$ представляет собой функцию двух переменных. В общем не очень сложном случае графиком $z = f(x, y)$ будет некая поверхность.



Изначально имеет смысл ориентироваться всё же на произвольное число переменных. В случае функциональной зависимости

$y = f(x_1, \dots, x_n)$ комплект

$$x = \{x_1, \dots, x_n\}$$

удобно считать *вектором*, мысленно откладывая x_1, \dots, x_n по координатным осям, и коротко записывая $f(x)$.

Поначалу думается, что векторные понятия используются для краткости. Это лишь часть правды. Геометрия пространства имеет «некоординатный» характер, и *координатное насилие* над ней мешает видеть такие простые понятия, как расстояние, направление, ортогональность, без которых «за деревьями леса не видно». В полной мере эти факторы выявляются в «Линейной алгебре», но и в данном контексте без них не обойтись.

Поэтому очень коротко напомним некоторые существенные для дальнейшего пункты. Аналогия с обычным представлением о векторах становится осмысленной при подходящем определении операций:

- *умножение на скаляр*¹, $\lambda x = \{\lambda x_1, \dots, \lambda x_n\}$,
- *сложение*, $x + y = \{x_1 + y_1, \dots, x_n + y_n\}$,
- *скалярное произведение*, $\boxed{x \cdot y = x_1 y_1 + \dots + x_n y_n}$,
эквивалентные обозначения: $x \cdot y = (x, y) = xy = \langle xy \rangle$.

С помощью скалярного произведения вводится *евклидова норма* (длина вектора),

$$\|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + \dots + x_n^2}, \quad (5.1)$$

а с помощью нормы *расстояние* $\|x - y\|$ между векторами x, y .

Множество векторов, на которых введены перечисленные операции, называют *n-мерным евклидовым пространством* и обозначают \mathbb{R}^n .

¹Обратите внимание, что изложение ведётся с расчётом на некоторую догадливость читателя. Разумеется, здесь λ — скаляр, и не составило бы труда написать $\lambda \in \mathbb{R}$, что принято в математических текстах. Но учебник со справочником «в одном флаконе» имеет свои минусы. Плюсы тоже есть, однако, хотя «кашу маслом не испортишь», с некоторого момента *каша с маслом превращается в масло с кашей*.

Скалярное произведение в \mathbb{R}^3 обычно определяется как произведение длин векторов на косинус угла между ними. Потом выясняется, что $(\mathbf{x}, \mathbf{y}) = x_1y_1 + \dots + x_3y_3$ даёт то же самое. В общем случае такое соответствие имеет место благодаря *неравенству Коши–Буняковского*,



$$\mathbf{x} \cdot \mathbf{y} \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|, \quad (5.2)$$

которое даёт возможность ввести понятие косинуса

$$\cos \varphi = \mathbf{x} \cdot \mathbf{y} / \|\mathbf{x}\| \cdot \|\mathbf{y}\|,$$

угла между векторами при любом n .

Неравенство (5.2) доказывается совсем просто. Раскрывая скобки в очевидном неравенстве $(\mathbf{x} - \lambda \mathbf{y})^2 \geq 0$, имеем $\lambda^2 \|\mathbf{y}\|^2 - 2\lambda \mathbf{x} \cdot \mathbf{y} + \|\mathbf{x}\|^2 \geq 0$ для любого λ , — что возможно лишь при неположительности дискриминанта квадратного многочлена — а это и есть (5.2).

Существенную конструктивную роль в многомерном анализе играет понятие *ортогональности*. Векторы \mathbf{x}, \mathbf{y} определяют как *ортогональные*, если их *скалярное произведение* равно нулю, $\mathbf{x} \cdot \mathbf{y} = 0$. Ортогональность в свою очередь позволяет ввести в \mathbb{R}^n понятие *плоскости*, как множества элементов \mathbf{x} , ортогональных вектору \mathbf{a} ,

$$\mathbf{a} \cdot \mathbf{x} = 0, \quad \text{т. е.} \quad a_1x_1 + \dots + a_nx_n = 0.$$

5.2 Предел и сходимость

Векторный язык при изучении функций $y = f(x_1, \dots, x_n)$ оказывается удобен и даже необходим с первых шагов.

5.2.1 Определение. Число A называется *пределом функции* $y = f(\mathbf{x})$ при \mathbf{x} стремлящемся к \mathbf{a} ,

$$f(\mathbf{x}) \rightarrow A \quad \text{при} \quad \mathbf{x} \rightarrow \mathbf{a}, \quad (5.3)$$

если по любому $\varepsilon > 0$ можно указать такое δ , что²

$$|f(\mathbf{x}) - A| < \varepsilon, \quad \text{как только} \quad 0 < \|\mathbf{x} - \mathbf{a}\| < \delta. \quad (5.4)$$

²Обратите внимание в (5.4) на существенность условия $0 < \|\mathbf{x} - \mathbf{a}\|$.

Вся разница с определением предела функции одной переменной заключается в замене меры близости $|x - a|$ на $\|x - a\|$. В результате n условий $|x_i - a_i| \rightarrow 0$ заменяются одним $\|x - a\| \rightarrow 0$. Но вода камень точит. Капля за каплей и пределы вырывают на новую платформу, где мозаика подробностей перестаёт ретушировать суть.

• **Определение.** Число A называется пределом функции $y = f(x)$ при $x \rightarrow \infty$, если по любому $\varepsilon > 0$ можно указать такое $M > 0$, что $|f(x) - A| < \varepsilon$, как только $\|x\| > M$. Вместо $A = \lim_{x \rightarrow \infty} f(x)$ используют также запись

$$A = \lim_{\|x\| \rightarrow \infty} f(x).$$

Заметим, что (ε, δ) -определение предела (5.3), как и в одномерном случае, эквивалентно следующей дефиниции.

• Число A есть предел функции $f(x)$ при $x \rightarrow a$, если

$$f(x_n) \rightarrow A$$

на любой последовательности $x_n \rightarrow a$ ($x_n \neq a$)³.

Здесь, кстати, можно многое повторить в новой обстановке, воспроизводя многоголосие учения о пределах функции одной переменной⁴. Некоторые учебники так и делают, решая заодно государственные задачи типа «чем-то занять население» или «отвлечь от психологически опасных попыток поиска смысла жизни». На самом деле важно понять, что в основе всевозможных определений пределов в разных дисциплинах лежит одна и та же схема. И когда она усваивается, человеку легче самому сконструировать нужное определение, чем выискивать его в тексте. Поэтому здесь можно рекомендовать вернуться к пределам функции одной переменной и посмотреть, как там трансформировались бы понятия при смене формата $y = f(x)$ на

³Опять-таки обратите внимание на принципиальное условие $x_n \neq a$.

⁴Начиная приблизительно из того же положения: векторные последовательности вместо числовых, последовательности Коши, бесконечно малые величины.

$$y = f(x_1, \dots, x_n).$$



5.3 Непрерывность

При переходе к функциям многих переменных не меняется также и определение непрерывности. Не меняется по сути. Изменяется лишь мера близости. Поэтому звучит та же мелодия немного в другой аранжировке.

5.3.1 Функция $f(x)$ непрерывна в точке⁵ $x = a$, если

$$f(x) \rightarrow f(a) \text{ при } x \rightarrow a,$$

т. е. если по любому $\varepsilon > 0$ можно указать такое δ , что

$$|f(x) - f(a)| < \varepsilon, \text{ как только } \|x - a\| < \delta.$$

Аналогия с одномерным случаем сохраняется тотально. Ту же роль играют *последовательности Коши*. Если установлено существование предела $f(x) \rightarrow A$ при $x \rightarrow a$, то $f(x^k) \rightarrow A$ на любой последовательности $x^k \rightarrow a$, и обратно. Далее:

- Из любой ограниченной последовательности x^k можно извлечь сходящуюся подпоследовательность (*лемма Больцано – Вейерштрасса*).
- Непрерывная на ограниченном замкнутом множестве функция $y = f(x)$ ограничена снизу и сверху, и достигает своих минимального и максимального значений (*теорема Вейерштрасса*).
- Функция $f(x)$, непрерывная на ограниченном замкнутом множестве X , автоматически *равномерно непрерывна*⁶ на X (*теорема Кантора*).

⁵И непрерывна на $X \subset \mathbb{R}^n$, если непрерывна в любой точке X .

⁶Функция $f(x)$ равномерно непрерывна на X , если по любому $\varepsilon > 0$ можно указать такое δ , что

$$|f(x) - f(y)| < \varepsilon, \text{ если } \|x - y\| < \delta$$

для любых $x, y \in X$.

5.4 Частные производные

Если у $f(x_1, \dots, x_n)$ зафиксировать все переменные кроме x_i , получится функция одной переменной x_i , её обыкновенную производную называют частной производной $f(x)$ по x_i , и обозначают $\frac{\partial f}{\partial x_i}$, или f'_{x_i} . Избежать недопонимания проще всего, рассмотрев несколько примеров.

$$z = xy^2 \quad \Rightarrow \quad \frac{\partial z}{\partial x} = y^2, \quad \frac{\partial z}{\partial y} = 2xy;$$

$$p = \frac{RT}{V} \quad \Rightarrow \quad p'_T = \frac{R}{V}, \quad p'_V = -\frac{RT}{V^2};$$

$$x = \sin(\omega t + \varphi) \quad \Rightarrow \quad x'_t = \omega \cos(\omega t + \varphi), \quad x'_\varphi = \cos(\omega t + \varphi).$$

• Пусть $z = f(u, v)$. Тогда производная функции одной переменной $z(t) = f(u(t), v(t))$ равна

$$\frac{\partial z}{\partial t} = \frac{\partial f}{\partial u} u'_t + \frac{\partial f}{\partial v} v'_t,$$

что устанавливается переходом к пределу в равенстве

$$\frac{\Delta z}{\Delta t} = \frac{f(u + \Delta u, v + \Delta v) - f(u, v + \Delta v)}{\Delta u} \frac{\Delta u}{\Delta t} + \frac{f(u, v + \Delta v) - f(u, v)}{\Delta v} \frac{\Delta v}{\Delta t}.$$

В общем случае $z = f(x_1, \dots, x_n)$, $x_i = x_i(y_1, \dots, y_m)$,

$$\frac{\partial z}{\partial y_j} = \sum_i \frac{\partial z}{\partial x_i} \frac{\partial x_i}{\partial y_j}.$$

Конечно, определить частную производную — дело нехитрое. Вопрос в другом. Какова роль частных производных в анализе функций многих переменных?

5.5 Приращения и дифференциалы

Приращение функции. Роль частных производных выявляется, когда речь заходит о *приращении функции*



$$\Delta f = f(x + \Delta x) - f(x),$$

что, с точки зрения наглядности, удобнее рассмотреть на примере двух переменных, $z = f(x, y)$. Перепишем для этого случая приращение

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y).$$

в виде

$$\Delta z = \frac{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}{\Delta x} \Delta x + \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \Delta y,$$

откуда следует

$$\Delta z = f'_x(x, y + \Delta y) \Delta x + o(\Delta x) + f'_y(x, y) \Delta y + o(\Delta y),$$

и в итоге

$$\Delta z = \frac{\partial f(x, y)}{\partial x} \Delta x + \frac{\partial f(x, y)}{\partial y} \Delta y + o(\|(x, y)\|),$$

но только в предположении $f'_x(x, y + \Delta y) \rightarrow f'_x(x, y)$ при $\Delta y \rightarrow 0$, что обеспечивается, например, непрерывностью производной $\frac{\partial f}{\partial x}$ в окрестности (x, y) . Прделанное указывает на справедливость следующего результата.

5.5.1 Теорема. Если функция $u = f(x)$ имеет непрерывные в окрестности точки $x = \{x_1, \dots, x_n\}$ частные производные, то справедлива формула приращения



$$\Delta z = \sum_{i=1}^n f'_{x_i}(x) \Delta x_i + o(\|\Delta x\|). \quad (5.5)$$

Непрерывность частных производных здесь не обязательна⁷. Однако просто наличия производных, которой хватает⁸ в одномерном случае, тут недостаточно.

⁷Но вполне приемлема, ибо на практике чаще всего имеет место.

⁸Для справедливости формулы приращения.

Время от времени полезно задумываться о сути используемых и наблюдаемых доказательств. Строго говоря, они являются лишь эмоциональной аргументацией. Степень подробности и траектория рассуждения всегда зависит от того, с чем «противная сторона» готова согласиться. Доказывая, мы оперируем смысловыми категориями. При этом логические прорехи очень трудно контролировать. «Имея семейство множеств, в каждом множестве можно выбрать по элементу» — не самоочевидный ли тезис? Вышло — нет⁹. Так что до синтаксического уровня аргументации человеку никогда не добраться. Хотя некоторые авторы пытаются увлечь население в эту «чёрную дыру».

Дифференциалы высших порядков идут «по тому же следу». Дифференциал от дифференциала даёт второй дифференциал:

$$d^2 f = d(df) = d\left(\sum_i \frac{\partial f}{\partial x_i} dx_i\right) = \sum_i d\left(\frac{\partial f}{\partial x_i}\right) dx_i,$$

т. е.

$$d^2 f = \sum_i \frac{\partial^2 f}{\partial x_i^2} dx_i^2 + 2 \sum_{i \neq j} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j.$$

Продолжая процесс, получаем

$$d^m f = \sum \frac{\partial^m f}{\partial x_1^{m_1} \dots \partial x_n^{m_n}} dx_1^{m_1} \dots dx_n^{m_n}, \quad (5.8)$$

где суммирование ведётся по всем $m_1 + \dots + m_n = m$.

Что касается разложения Тейлора

$$\Delta f = \sum_{m=1}^k \frac{1}{m!} \sum_{\sigma} \frac{\partial^m f}{\partial x_1^{m_1} \dots \partial x_n^{m_n}} \Delta x_1^{m_1} \dots \Delta x_n^{m_n} + o(\|\Delta x\|^k),$$

то оно естественным образом перекликается с (5.8), и может быть получено *разложением Тейлора* в нуле функции одной переменной $\varphi(\tau) = f(x + \tau \Delta v)$, где $\Delta x = \tau \Delta v$.

5.7 Градиент

В многомерном анализе важную роль играют понятия *градиента* и *производной по направлению*.

⁹См. аксиому выбора.

Производная функции $z = f(x, y)$ по направлению *единичного вектора* $s = \{s_x, s_y\}$ определяется как обычная производная по θ :

$$z'_s = z'_\theta = \frac{d}{d\theta} f(x + \theta s_x, y + \theta s_y) = \frac{\partial f}{\partial x} s_x + \frac{\partial f}{\partial y} s_y. \quad (5.9)$$

Правую часть (5.9) естественно интерпретировать как скалярное произведение вектора $s = \{s_x, s_y\}$ на вектор

$$\text{grad } f = \left\{ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\},$$

который называют *градиентом* функции f . Для градиента используют также обозначение ∇f (читается «набла эф»). Таким образом

$$z'_s = s \cdot \nabla f. \quad (5.10)$$

В силу

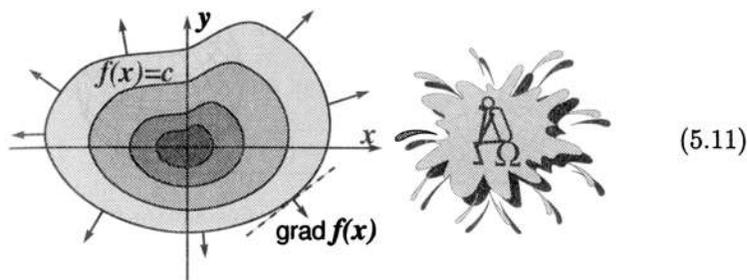
$$s \cdot \text{grad } f = \|\text{grad } f\| \|s\| \cdot \cos \varphi$$

максимум в (5.10) достигается при $\varphi = 0$, т. е. когда вектор s совпадает по направлению с градиентом. При этом значение максимума равно модулю градиента $\|\nabla f\|$, поскольку $\|s\| = 1$. Следовательно, градиент ∇f — это *вектор скорости максимального роста функции* f .

Производная (5.10) обращается в нуль, когда s и $\|\nabla f\|$ ортогональны друг другу, что имеет место, если s направлен вдоль касательной к линии постоянного уровня¹⁰ $f(x, y)$. Стало быть, в каждой точке поверхности $f(x, y) = \text{const}$ градиент ∇f перпен-

¹⁰ Или лежит в касательной плоскости к поверхности постоянного уровня $f(x)$, если размерность аргумента ≥ 3 .

дикулярен этой поверхности, направлен по нормали.



(5.11)

Таким образом, касательная плоскость в точке $\{x_0, y_0\}$ к линии постоянного уровня функции $z = f(x, y)$ описывается уравнением¹¹

$$\frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) = 0, \quad (5.12)$$

а касательная плоскость к поверхности $z = f(x, y)$ в точке $\{z_0, x_0, y_0\}$ определяется уравнением

$$z - z_0 = \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0). \quad (5.13)$$

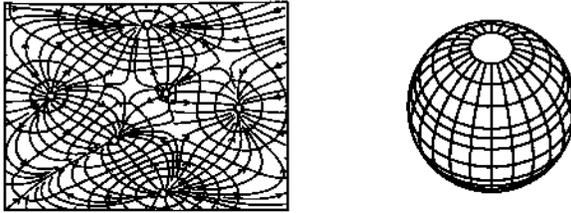
Уравнение (5.13) является непосредственным аналогом (5.12), поскольку касательная плоскость к графику (5.13) — это касательная плоскость к поверхности постоянного уровня функции трёх переменных $u = z - f(x, y)$, градиент которой

$$\left\{ 1, -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y} \right\}.$$

Двумерный случай (аргумент лежит в плоскости), конечно, в силу своей наглядности должен служить первоисточником при изучении многомерного случая. Визуальное представление линий постоянного уровня и перпендикулярных им линий, вдоль которых направлены градиенты, помогает лучше усвоить геометрию внутреннего устрой-

¹¹На рис. (5.11) это пунктирная прямая.

ства функций $z = f(x, y)$, а там и до $z = f(x_1, \dots, x_n)$ недалеко.



При этом благоприятно для воображения поразмышлять над траекториями движения по градиенту

$$\dot{\mathbf{x}} = \nabla f(\mathbf{x}), \quad (5.14)$$

где

$$\nabla f(\mathbf{x}) = \text{grad } f(\mathbf{x}) = \left\{ \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right\}.$$

Один из источников интереса к системам типа (5.14) — численные процедуры поиска локальных максимумов функции $f(\mathbf{x})$.

Отталкиваясь от двумерного случая, легко записать касательные плоскости в общем случае $z = f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$. Уравнение касательной плоскости в \mathbb{R}^n к поверхности постоянного уровня в точке \mathbf{x}_0 :

$$\nabla f(\mathbf{x}) \cdot (\mathbf{x} - \mathbf{x}_0) = 0.$$

Уравнение касательной плоскости в \mathbb{R}^{n+1} к поверхности графика $z = f(\mathbf{x})$ в точке $\{z_0, \mathbf{x}_0\}$, $z_0 = f(\mathbf{x}_0)$:

$$z - z_0 = \nabla f(\mathbf{x}) \cdot (\mathbf{x} - \mathbf{x}_0).$$

Формула конечного приращения в векторных обозначениях приобретает вид

$$\Delta z = \nabla f(\mathbf{x}) \cdot \Delta \mathbf{x} + o(\|\Delta \mathbf{x}\|).$$

Применение *теоремы Лагранжа* к функции скалярного аргумента

$$\varphi(\tau) = f(\mathbf{x} + \tau(\mathbf{y} - \mathbf{x}))$$

даёт

$$\varphi(1) - \varphi(0) = \varphi'(\theta) = \sum_i \frac{\partial f(z)}{\partial x_i} (y_i - x_i) = \nabla f(\mathbf{w}) \cdot (\mathbf{y} - \mathbf{x}),$$

где $\mathbf{w} = \mathbf{x} + \theta(\mathbf{y} - \mathbf{x})$ при некотором $\theta \in (0, 1)$. Поэтому

$$\boxed{f(\mathbf{y}) - f(\mathbf{x}) = \nabla f(\mathbf{w}) \cdot (\mathbf{y} - \mathbf{x})},$$

что называют *теоремой о среднем*.

5.8 О роли повторных пределов

Равенство повторных пределов. Повышенное внимание к повторным пределам типа

$$\lim_{x \rightarrow a} \lim_{y \rightarrow b} f(x, y)$$

и сосредоточенность на изучении равенств

$$\lim_{x \rightarrow a} \lim_{y \rightarrow b} f(x, y) = \lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x, y) \quad \text{✎} \quad (5.15)$$

выглядят весьма странно, ибо думается, что главным действующим лицом в подобном контексте должен быть предел

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y), \quad (5.16)$$

не зависящий от пути стремления $(x, y) \rightarrow (a, b)$. Но существование (5.16) из (5.15) даже не следует, да и обратная импликация не работает, см. далее примеры.

Поэтому (5.15) и (5.16) — явления независимые, хотя и соприкасающиеся. При этом (5.15) имеет важное самостоятельное значение, едва ли не большее чем (5.16). Дело в том, что в подголотной основных операций матанализа (дифференцирования, интегрирования, суммирования рядов) лежат предельные переходы. И вопросы изменения порядка этих операций упираются,

по существу, в равенство повторных пределов. Можно ли менять порядок дифференцирования¹²,

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}, \quad (?)$$

вносить дифференцирование под интеграл,

$$\frac{d}{dy} \int_{\Omega} f(x, y) dx = \int_{\Omega} \frac{\partial f(x, y)}{\partial y} dx, \quad (?)$$

менять местами интегрирование с переходом к пределу,

$$\lim_{y \rightarrow 0} \int_{\Omega} f(x, y) dx = \int_{\Omega} \lim_{y \rightarrow 0} f(x, y) dx, \quad (?)$$

— это всё разные лица равенства (5.15).

Поэтому на первом этапе изучения проблематики разумно сконцентрироваться на ситуациях, где всё на виду. И там особое внимание уделить осмотру «берегов», тех территорий, где желанные и ожидаемые свойства нарушаются.

Примеры

- Вот простейшая ситуация

$$\lim_{x \rightarrow 0} \lim_{y \rightarrow 0} \frac{ax + cy}{bx + dy} = \frac{a}{b}, \quad \lim_{y \rightarrow 0} \lim_{x \rightarrow 0} \frac{ax + cy}{bx + dy} = \frac{c}{d},$$

в которой повторные пределы могут не совпадать.

¹²Функция

$$f(x, y) = \begin{cases} \frac{x^2 - y^2}{x^2 + y^2} xy, & \text{если } x, y \neq 0; \\ 0, & \text{если } x = y = 0, \end{cases}$$

имеет в нуле частные производные $f_{xy}(0, 0) = 1$, $f_{yx}(0, 0) = -1$, что противоречит расхожему представлению о равенстве смешанных производных независимо от порядка дифференцирования. Но в данном случае «не хватает» непрерывности f_{xy} и f_{yx} в окрестности нуля.

- У функции

$$f(x, y) = x \sin \frac{1}{y}, \quad \text{---} \quad (5.17)$$


доопределённой в нуле по непрерывности, $f(0, 0) = 0$, существует двойной предел при $(x, y) \rightarrow (0, 0)$, но только один из посторонних.

- А у функции

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & \text{если } x, y \neq 0; \\ 0, & \text{если } x = y = 0 \end{cases} \quad (5.18)$$

в нуле *существуют и равны* оба повторных предела, но двойного — вообще нет¹³.

Теоремное обеспечение

Повторные пределы ускользают из-под контроля при наличии у функции тех или иных особенностей. В благоприятных ситуациях предосторожности чаще всего излишни.

5.8.1 Теорема. Если функция $f(x, y)$ при $x \rightarrow a$, $y \rightarrow b$ имеет двойной предел A (конечный или бесконечный), и при любом y из некоторой окрестности точки b существует предел $\lim_{x \rightarrow a} f(x, y)$, то существует и

$$\lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x, y) = A.$$

¹³Обратим внимание, функция (5.18) имеет разрыв в нуле, причём довольно экзотический:

- Предела у $f(x)$ при $(x, y) \rightarrow 0$ нет, но при стремлении к нулю вдоль любой прямой $x = ky$ значение f имеет предел $\frac{k}{1+k^2}$.
- Несмотря на разрыв, при любом фиксированном y функция $\varphi(x) = f(x, \text{const})$ — непрерывна, и имеет частные производные всюду, в том числе в нуле, т. е. в точке разрыва!

Другой повторный предел, как следует из примера (5.17), может не существовать. Если однако, в дополнение к условиям теоремы, при любом x из некоторой окрестности точки a существует предел $\lim_{y \rightarrow b} f(x, y)$, то оба повторных предела существуют и равны двойному.

◀ Доказательство теоремы несложно. По любому $\varepsilon > 0$ можно указать такое $\delta > 0$, что

$$|x - a| < \delta, |y - b| < \delta \Rightarrow |f(x, y) - A| < \varepsilon \Rightarrow |g(y) - A| < \varepsilon,$$

где $g(y) = \lim_{x \rightarrow a} f(x, y)$. Поэтому $\lim_{y \rightarrow b} g(y) = A$. ▶

Вот ещё один удобный на практике результат.

5.8.2 Лемма. Пусть при любом x из некоторой окрестности точки a существует предел $\lim_{y \rightarrow b} f(x, y) = \varphi(x)$, а при каждом y из некоторой окрестности b существует предел $\lim_{x \rightarrow a} f(x, y) = \psi(y)$. Если при этом $f(x, y)$ стремится к $\varphi(x)$ равномерно по x , то оба повторных предела существуют и равны друг другу. (?)



О равномерной сходимости

На практике часто рассматривается последовательность функций $f_n(x)$, поточечно сходящаяся к «некой» функции $f(x)$, т. е. $f_n(x) \rightarrow f(x)$ при каждом фиксированном x . В качестве $f_n(x)$ обычно фигурируют частичные суммы какого-либо ряда $\sum_{k=0}^{\infty} \varphi_k(x)$. При этом возникают стандартные вопросы наследования функцией $f(x)$ тех или иных свойств функций $f_n(x)$. Например, непрерывна ли $f(x)$ в точке x_0 , если все $f_n(x)$ непрерывны в x_0 ? Легко видеть, что положительный ответ эквивалентен равенству



$$\lim_{x \rightarrow x_0} \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow x_0} f_n(x).$$

Проблема существования двойного предела даже не возникает.

Ключом к решению в такого рода задачах служит часто та или иная разновидность *равномерной сходимости*. Последовательность функций $f_n(x)$, поточечно сходящихся к $f(x)$ для любого $x \in X$, называют *равномерно сходящейся*, если по любому $\varepsilon > 0$ можно указать такое N , что¹⁴

$$|f_n(x) - f(x)| < \varepsilon, \quad \text{при } n > N, \quad x \in X.$$

5.8.3 Теорема. *Если последовательность непрерывных функций $f_n(x)$ сходится к $f(x)$ равномерно на замкнутом ограниченном множестве X , то $f(x)$ непрерывна на X .*

◀ Равномерная сходимость $f_n(x)$ к $f(x)$ означает

$$f(x) = f_n(x) + \varphi_n(x),$$

где $\varphi_n(x)$ равномерно сходится к нулю. Поэтому

$$|f(x) - f(y)| \leq |f_n(x) - f_n(y)| + |\varphi_n(x)| + |\varphi_n(y)|.$$

Далее по заданному $\varepsilon > 0$ выберем такое n , что (в силу равномерного стремления $\varphi_n(x)$ к нулю)

$$|\varphi_n(z)| < \varepsilon/3$$

для любого $z \in X$. Затем для выбранного n укажем такое δ (в силу непрерывности функций $f_n(x)$), что из $|x - y| < \delta$ будет следовать $|f_n(x) - f_n(y)| < \varepsilon/3$, и в итоге $|f(x) - f(y)| < \varepsilon$. ▶

С феноменом *равномерной сходимости* целесообразно освоиться на примерах и контрпримерах, дабы «чувствовать берега» и придерживаться магистральных путей.

• Стандартной иллюстрацией на тему «как плоха поточечная сходимость в отсутствие равномерной» обычно служит последовательность $f_n(x) = x^n$, которая при любом фиксированном $x \in (0, 1)$ сходится к нулю, при $x = 1$ — к единице. Иными словами, x^n на $[0, 1]$ *поточечно* сходится к разрывной функции.

¹⁴Идея равномерной сходимости может принимать разные облики. Поэтому важно ощущать саму идею, чтобы не запоминать слишком много определений. Вместо дискретного параметра n , например, может фигурировать — непрерывный.

• Функция

$$f(x, y) = \begin{cases} 2 \left(\frac{1}{y} - \frac{x}{y^2} \right), & \text{если } 0 \leq x \leq y; \\ 0, & \text{если } y \leq x \leq 1 \end{cases}$$

при $y \rightarrow 0$ сходится к нулю не равномерно, поэтому следствия типа

$$\lim_{y \rightarrow 0} \int_0^1 f(x, y) dx = 1, \quad \int_0^1 \lim_{y \rightarrow 0} f(x, y) dx = 0$$

не вызывают удивления.

• Но равномерная сходимость сама по себе не гарантирует благополучия, особенно если находится в глубине многослойной задачи. Пусть, например, $f_n(x) = \frac{1}{n}$ при $x \in [0, n]$, и $f_n(x) = 0$ при $x > n$. Очевидно, последовательность функций $\{f_n(x)\}$ равномерно сходится к нулю на $[0, \infty)$, но

$$\lim_{n \rightarrow \infty} \int_0^{\infty} f_n(x) dx = 1.$$

5.9 Интегралы, зависящие от параметра

В прикладных задачах приходится рассматривать функции, возникающие в результате интегрирования. Например,

$$\Phi(y) = \int_a^b f(x, y) dx. \quad (5.19)$$

Поскольку интегрирование обладает *улучшающими свойствами* (из интегрируемой функции делает непрерывную, из непрерывной — дифференцируемую), то многие манипуляции с $\Phi(y)$ можно менять местами с интегрированием при довольно свободных предположениях относительно функции $f(x, y)$.

Вот простые и легко доказываемые факты.

5.9.1 Если

$$f(x, y) \rightarrow \varphi(x) \quad \text{при } y \rightarrow y_0$$

равномерно¹⁵ по x , то

$$\lim_{y \rightarrow y_0} \int_a^b f(x, y) dx = \int_a^b \lim_{y \rightarrow y_0} f(x, y) dx. \quad (?)$$

5.9.2 Пусть функция $f(x, y)$ и частная производная $\partial f(x, y)/\partial y$ непрерывны в прямоугольнике $[a, b] \times [c, d]$. Тогда

$$\frac{d}{dy} \int_a^b f(x, y) dx = \int_a^b \frac{\partial f(x, y)}{\partial y} dx$$

при любом $y \in [c, d]$, что в случае зависимости верхнего предела интегрирования от y дополняется равенством

$$\frac{d}{dy} \int_a^{\varphi(y)} f(x, y) dx = \int_a^{\varphi(y)} \frac{\partial f(x, y)}{\partial y} dx + \varphi'(y) f(\varphi(y), y),$$

где имеется в виду $\varphi(y) \in [a, b]$. (?)

5.9.3 Если функция $f(x, y)$ непрерывна в прямоугольнике $[a, b] \times [c, d]$, то

$$\int_c^d dy \int_a^b f(x, y) dx = \int_a^b dx \int_c^d f(x, y) dy. \quad (?)$$

Выдвигаемые требования к $f(x, y)$ не выглядят слишком жёсткими, но и они могут быть ослаблены. Выходит, что интегрирование для стандартных операций анализа не воздвигает больших преград. Предельные переходы могут свободно проникать «сквозь интеграл» снаружи внутрь, и наоборот. Проблема усложняется при переходе к несобственным интегралам, где приходится опираться на дополнительное условие *равномерной сходимости интегралов*.

¹⁵Равномерная сходимость по x в данном случае означает: по любому $\varepsilon > 0$ можно указать такое δ , что $|f(x, y) - \varphi(x)| < \varepsilon$ при любом x , как только $|y - y_0| < \delta$.

В случае

$$\int_a^\infty f(x, y) dx = \lim_{A \rightarrow \infty} \int_a^A f(x, y) dx \quad (5.20)$$

это означает равномерную сходимость (5.20) по y , т. е. по любому $\varepsilon > 0$ можно указать такое B , что при $A > B$ выполняется неравенство

$$\left| \int_a^\infty f(x, y) dx - \int_a^A f(x, y) dx \right| < \varepsilon$$

для любого y из рассматриваемой области Y .

Для несобственных интегралов на конечном промежутке ключевым также оказывается условие *равномерной сходимости интегралов*, определяемое по аналогии с предыдущим. Выполнение соответствующих условий приходится проверять отдельно либо «в лоб», либо опираясь на различные достаточные признаки. В более сложных случаях используется так называемая *теорема Арцела*, но это уже инструмент из «другой оперы».

Глава 6

Отображения, или операторы

Сложное — и дурак придумает.
Михаил Кошкин¹

6.1 Аргументы и функции как векторы

Чаще всего на практике приходится иметь дело с моделями, в которых и «аргументом», и «функцией» служат наборы числовых переменных.

Линейная модель производства. Пусть x_j обозначает интенсивность j -го технологического процесса, a_{ij} — количество i -го продукта, производимого ($a_{ij} > 0$) или потребляемого ($a_{ij} < 0$) при единичной интенсивности j -го технологического процесса. Тогда суммарное производство (потребление) i -го продукта определяется суммой

$$y_i = \sum_{j=1}^n a_{ij} x_j, \quad i = 1, \dots, m. \quad (6.1)$$

Межотраслевой баланс. Имеется n отраслей, i -я отрасль выпускает i -й продукт в количестве x_i . На выпуск единицы i -го продукта в системе затрачивается $a_{ij} \geq 0$ единиц j -го продукта. Понятно, что при выпуске набора

$$x = \{x_1, \dots, x_n\}$$

¹Главный конструктор легендарного танка Т-34.

чистый выпуск i -го продукта равен $y_i = x_i - \sum_j a_{ij}x_j$. Вопрос о продуктивности модели, таким образом, сводится к положительной разрешимости системы уравнений

$$\begin{array}{c} \text{stick figure} \end{array} \quad x_i - \sum_j a_{ij}x_j = y_i, \quad i = 1, \dots, n. \quad (6.2)$$



В подобного рода ситуациях совокупности функций

$$\begin{cases} y_1 = f_1(x_1, \dots, x_n), \\ \dots \\ y_n = f_n(x_1, \dots, x_n) \end{cases} \quad (6.3)$$

вектору $x = \{x_1, \dots, x_n\}$ сопоставляют вектор $y = \{y_1, \dots, y_n\}$. Стенографический трюк, состоящий в замене (6.3) записью

$$y = f(x), \quad (6.4)$$

переводит мысль на следующий этаж виртуальных манипуляций. Поначалу думается: ну, не может такого быть. Какой толк от перехода (6.3) \rightarrow (6.4) кроме экономии чернил? Тем не менее (6.4) рождает новую категорию мышления. «Бухгалтерская ведомость» (6.3) превращается в компактный атом f , и появляется возможность говорить, что f отображает элемент x в y , в связи с чем f называют *отображением*, или *оператором*². Элементы x и y считают при этом векторами, или точками³ из \mathbb{R}^n . Таким образом, f переводит точки в точки. Пустячок, казалось бы, но он ведёт к революционным метаморфозам в головах пользователей.

Шапки вверх бросать торопиться, конечно, не надо. Процесс медленный, постепенный. Ассортимент выгод пополняется шаг за шагом. В какой-то момент, скажем, обращает на себя внимание композиция fg , т. е. $f(g(x))$,

$$x \xrightarrow{g} y \xrightarrow{f} z,$$



²Или *преобразованием*. Не возбраняется и прежнее название «функция».

³Несмотря на то, что переменные x , y могут иметь разный физический смысл, их можно считать векторами одного пространства в рамках фиксированной системы измерений.

но попробуйте расписать $z = f(g(x))$ в духе (6.3). И беда тут не в том, что писанины много, а в том, что мысли расползаются.

Что касается пределов и непрерывности таких векторнозначных функций f , то в определениях модули меняются на нормы — вот и вся разница. Два примера:

6.1.1 Определение. Вектор b называется пределом отображения $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ при x стремящемся к a (т. е. $f(x) \rightarrow b$ при $x \rightarrow a$), если по любому $\varepsilon > 0$ можно указать такое δ , что

$$\|f(x) - b\| < \varepsilon, \quad \text{если} \quad \|x - a\| < \delta.$$

6.1.2 Определение. Оператор $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ называют непрерывным в точке a , если по любому $\varepsilon > 0$ можно указать такое δ , что

$$\|f(x) - f(a)\| < \varepsilon, \quad \text{если} \quad \|x - a\| < \delta,$$

и непрерывным на $X \subset \mathbb{R}^n$, если он непрерывен в любой точке X .

6.2 Линейные отображения

Простейшим примером векторнозначной функции (6.4) является линейный оператор. Его покомпонентная запись выглядит так:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \dots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n. \end{cases} \quad (6.5)$$

Каждая строчка в (6.5) — это скалярная линейная функция

$$z = \mathbf{a} \cdot \mathbf{x} = a_1x_1 + \cdots + a_nx_n,$$

которая принимает постоянные значения $\mathbf{a} \cdot \mathbf{x} = \beta$ на гиперплоскостях⁴ параллельных плоскости $\mathbf{a} \cdot \mathbf{x} = 0$. Действительно, из $\mathbf{a} \cdot \mathbf{u} = \beta$, $\mathbf{a} \cdot \mathbf{v} = \beta$ следует $\mathbf{a} \cdot (\mathbf{u} - \mathbf{v}) = 0$, т. е. вектор $\mathbf{u} - \mathbf{v}$ параллелен (лежит в) плоскости $\mathbf{a} \cdot \mathbf{x} = 0$.

⁴Множества $\mathbf{a} \cdot \mathbf{x} = \beta$ называют гиперплоскостями, чтобы подчеркнуть, что они не являются линейными пространствами, которым положено вместе с векторами содержать их сумму. Другими словами, гиперплоскость — это плоскость, не проходящая через начало координат.

Таблицу коэффициентов $A = [a_{ij}]$ в (6.5),



$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

называют *матрицей* и коротко пишут $y = Ax$, определяя тем самым умножение матрицы на вектор. При этом в $y = Ax$ под x , y подразумеваются вектор-столбцы. Первый индекс у a_{ij} обозначает номер строки, второй — столбца. Совокупность элементов a_{11}, \dots, a_{nn} называют *главной диагональю* матрицы.

Теория операторов (6.5), или матриц, изучается в «*Линейной алгебре*», каковая является выдающейся математической дисциплиной, сочетающей в себе глубину с изяществом и лёгкостью. Но об этом речь в отдельном курсе. В данном контексте остановимся лишь на некоторых понятиях, необходимых по ходу дела.

Умножение матрицы на скаляр и сложение матриц:

$$\gamma A = [\gamma a_{ij}], \quad A + B = [a_{ij} + b_{ij}].$$

При этом, очевидно,

$$A(\lambda x + \mu y) = \lambda Ax + \mu Ay, \quad (6.6)$$

и это показывает, что на матрицу можно смотреть как на *линейный оператор*⁵. Подобный взгляд автоматически порождает *правило умножения матриц*. Допустим, вектор x под действием оператора B переходит в $z = Bx$, а вектор z под действием оператора A переходит в $y = Az$. Результирующее преобразование x в y определяется матрицей $C = AB$ с элементами c_{ij} , формула вычисления которых определяется обыкновенным приведением подобных,

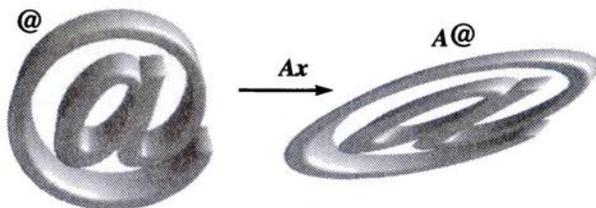
$$y_i = \sum_k a_{ik} \sum_j b_{kj} x_j = \sum_j \sum_k a_{ik} b_{kj} x_j = \sum_j c_{ij} x_j.$$

Следовательно, $C = AB$ есть матрица с элементами

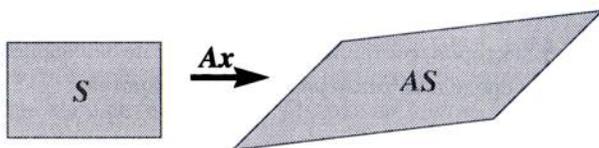
$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}. \quad \text{⚠️} \quad (6.7)$$

⁵(6.6) — как раз характеристическое свойство линейного оператора.

Произведение матриц в общем случае не коммутативно⁶, $AB \neq BA$, но ассоциативно, $A(BC) = (AB)C$, и дистрибутивно, $A(B+C) = AB+AC$, что легко проверяется.



На векторы и фигуры линейное преобразование Ax действует как поворот плюс растяжение. В случае $Ax + b$ добавляется ещё перенос.



Матрица A^T с элементами $a_{ij}^T = a_{ji}$ называется *транспонированной* к A . Матрицу называют *симметрической*, если $A = A^T$, и *кососимметрической*, если $A = -A^T$. Для транспонированной матрицы широко используется также обозначение A' .

При транспонировании произведения:

$$(AB)^T = B^T A^T, \quad (6.8)$$

что элементарно устанавливается:

$$\blacktriangleleft (AB)^T = \left[\sum_k a_{ik} b_{kj} \right]^T = \left[\sum_k a_{jk} b_{ki} \right] = \left[\sum_k b_{ik}^T a_{kj}^T \right] = B^T A^T. \blacktriangleright$$

В данном контексте имеет смысл кое-что уточнить насчёт устройства \mathbb{R}^n , в котором действуют линейные преобразования. Говорят, что множество векторов $\{x^1, \dots, x^k\}$ *линейно зависимо*, если существуют такие коэффициенты $\lambda_1, \dots, \lambda_k$, не все равные нулю, что

$$\lambda_1 x^1 + \dots + \lambda_k x^k = 0.$$

⁶В случае $\boxed{AB = BA}$ говорят, что матрицы A и B *коммутируют*.

В противном случае говорят о *линейной независимости* $\{x^1, \dots, x^k\}$. *Коллинеарные* векторы (лежащие на одной прямой), а также *компланарные* векторы (лежащие в одной плоскости) — линейно зависимы.

Линейно независимое множество $\{e^1, \dots, e^n\}$ называют *базисом*, если любой вектор x можно представить в виде линейной комбинации

$$x = \lambda_1 e^1 + \dots + \lambda_n e^n.$$

Стандартный базис \mathbb{R}^n :

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Число n векторов, составляющих базис (и не зависящее от выбора последнего), определяет *размерность* пространства.

6.3 Обратимые преобразования

Изучению отображений способствуют некоторые удобные понятия. А также неудобные, но общепринятые.

- В ситуации $y = f(x)$ элемент y считается *образом* x , а совокупность элементов x , образом которых служит y , — *прообразом*⁷ y .

- О функции $f : X \rightarrow Y$, отображающей элементы $x \in X$ в элементы $y \in Y$, в случае $f(X) = Y$ говорят как о *сюръективном отображении*, или *отображении «на»*⁸.

- Если образы различных элементов $x_1 \neq x_2$ не совпадают, f называют *инъективным отображением*, или *инъекцией*. В этом случае *уравнение* $y = f(x)$ имеет (если имеет) единственное решение $x = f^{-1}(y)$, где f^{-1} называют *обратным отображением*, которое может быть определено не на всем Y . Но если f ещё и сюръекция, то $f^{-1} : Y \rightarrow X$.

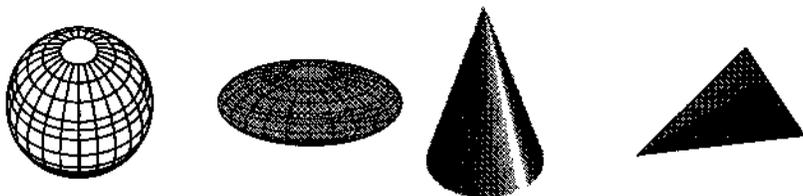
⁷ Либо *полным прообразом*, оставляя за термином «прообраз» любой элемент x , переходящий в y .

⁸ В общем случае $f : X \rightarrow Y$ говорят, что f отображает X «в» Y .

Особого внимания заслуживает понятие *гомеоморфизма*, каковым называют взаимно однозначное (*инъективное* плюс *сюръективное*) и взаимно непрерывное (непрерывны как f , так и f^{-1})⁹ отображение

$$f: X \rightarrow Y \quad (6.9)$$

Пространства, между которыми можно установить гомеоморфизм, считаются *гомеоморфными*.



Отображение называют *локальным гомеоморфизмом*, если у любой пары точек x, y такой, что $f(x) = y$, найдутся окрестности, гомеоморфно отображаемые (с помощью f) друг на друга.

Перечисленное имеет прямое отношение к распространённой задаче исследования систем уравнений вида (6.3), (6.4), т. е.

$$f(x) = y. \quad (6.10)$$

Повторим ещё раз. Если f *сюръективно*, система (6.10) имеет решение $x \in X$ при любом $y \in Y$. Если — *инъективно*, то решение (6.10) всегда единственно (разумеется, при условии существования). А когда — *инъективно* и *сюръективно*, то решение (6.10) при любом $y \in Y$ существует и единственно. Наконец, в случае *гомеоморфизма* f — добавляется ещё непрерывная зависимость решения $x \in X$ от $y \in Y$.

Интересно посмотреть, как сказанное преломляется в частном случае *линейного отображения*

$$f(x) = Ax. \quad (6.11)$$

Здесь *инъективность* обеспечивает «всё остальное». Если

$$Ax^1 \neq Ax^2 \quad \text{при} \quad x^1 \neq x^2, \quad (6.12)$$

⁹Возможность взаимно однозначного соответствия, непрерывного только в одну сторону — кажется сомнительной, но в этом мире случается все, что не запрещено, см. [3].

то (6.11) *сюръективно*¹⁰ и *гомеоморфно* отображает \mathbb{R}^n на \mathbb{R}^n , что означает существование *обратного оператора* A^{-1} .

Но это пока в «божественных понятиях». Так сказать, с точки зрения Всевидящего Ока. Ибо не ясно, как проверять (6.12). Ответ даёт «*Линейная алгебра*». Оператор Ax обратим, когда *детерминант* (*определитель*) матрицы A не равен нулю,

$$\det A \neq 0. \quad (6.13)$$

В случае (6.13) матрицу A называют *невырожденной*. Равносильный признак невырожденности A — *линейная независимость* вектор-строк матрицы¹¹. За информацией о *детерминантах* и *линейной независимости* лучше обратиться к «*Линейной алгебре*», но краткую справку мы всё же дадим.

6.4 Детерминанты, или определители

Единичный куб, построенный на векторах (ребрах)

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (6.14)$$

под действием матрицы (оператора) $A = [a_{ij}]$ переходит в параллелепипед, построенный на вектор-столбцах матрицы A ,

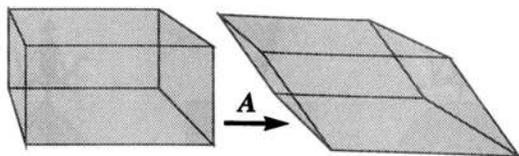
$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad \dots, \quad \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix}.$$

Объём этого параллелепипеда и есть $\det A$ с точностью до знака, т. е. *детерминант* — это *коэффициент искажения объёма* при линейном преобразовании A с учётом сохранения или перемены *ориентации*. Дело в том, что любое тело $\Omega \subset \mathbb{R}^n$ можно разбить на N кубиков со стороной ε и объёмом ε^n , и тем самым сколь угодно точно приблизить объём Ω величиной $N \cdot \varepsilon^n$. Каждый кубик под действием A перейдет в параллелепипед объёма $|\det A| \cdot \varepsilon^n$.

¹⁰Равно как и *сюръективность*, т. е. разрешимость $Ax = y$ при любом y , вытекает за собой (6.12).

¹¹Из линейной независимости вектор-строк вытекает *линейная независимость* вектор-столбцов, и наоборот.

Поэтому A искажает объём любого тела в $|\det A|$ раз¹².



Другие пути определения и вычисления *детерминанта* см. всё-таки в «*Линейной алгебре*». Не так трудно было бы дать подробные пояснения и здесь, но это нарушило бы ритм изложения. Да и знакомиться с некоторыми экспонатами надо в подходящей обстановке, при адекватных декорациях. Там возникает резонанс с духом n -мерного коловращения, синхронизация с внутренней логикой дисциплины. А здесь всё это будет потеряно. Не говоря о том, что «посторонние разговоры» уведут мысль с магистрального пути.

6.5 Эквивалентные нормы

Евклидова норма (5.1), $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$, служит базой для введения в \mathbb{R}^n *расстояния* $\|\mathbf{x} - \mathbf{y}\|$ между векторами \mathbf{x}, \mathbf{y} . Но популярны и другие нормы. Например, *кубическая норма*,

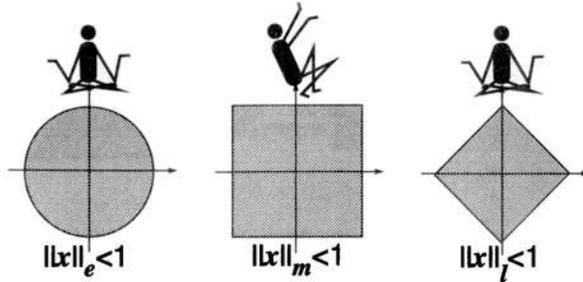
$$\|\mathbf{x}\|_m = \max_i |x_i|, \quad (6.15)$$

и *октаэдрическая норма*,

$$\|\mathbf{x}\|_l = \sum_{i=1}^n |x_i|. \quad (6.16)$$

¹²Из определения детерминанта как *коэффициента искажения объёма* сразу следует $|\det AB| = |\det A||\det B|$, хотя, как известно, модули можно опустить: $\det AB = \det A \det B$.

Вот примеры *единичных шаров* в \mathbb{R}^2 :



Возникает, конечно, естественный вопрос: зачем столько норм? Нельзя ли обойтись одной? В принципе можно, но многообразие норм обеспечивает удобства. Выбор нормы «под задачу», как правило, существенно облегчает решение. Скажем, требуется установить сходимость $x^n \rightarrow 0$, для чего необходимо доказать, что при любом $\varepsilon > 0$ для достаточно больших n выполняется неравенство

$$\|x^n\| < \varepsilon. \quad (6.17)$$

Какую норму в (6.17) выбрать? Может оказаться, что по одной норме (6.17) легко доказывается, по другой — нет. Поэтому выбор нормы — это рычаг, распоряжаясь которым, можно облегчать решение задач. Но есть ещё один принципиальный вопрос, может быть, самый принципиальный. Если последовательность x^n сходится по одной норме, будет ли она сходиться — по другой. В \mathbb{R}^n ответ — «да»¹³.

6.5.1 *С точки зрения сходимости в \mathbb{R}^n все нормы эквивалентны, т. е. для любых двух норм $\|\cdot\|_1, \|\cdot\|_2$ можно указать такие константы α и β , что*

$$\|x\|_1 \leq \alpha \|x\|_2, \quad \|x\|_2 \leq \beta \|x\|_1$$

при любом $x \in \mathbb{R}^n$.

¹³В функциональных пространствах — «не обязательно».

◀ Легко видеть, что любая норма однозначно определяется заданием единичного шара¹⁴. Функция $\varphi(x) = \|x\|_1$ на множестве $\|x\|_2 = 1$ достигает своих минимального ($\gamma > 0$) и максимального ($\delta > 0$) значений. Поэтому

$$\gamma \|x\|_2 \leq \|x\|_1 \leq \delta \|x\|_2$$

при любом $x \in \mathbb{R}^n$. ▶

Нормой матрицы A называют положительное число $\|A\|$, удовлетворяющее тем же самым условиям 1 – 3 (см. сноску), если в них векторы заменить на матрицы. Поскольку матрицы — это по существу линейные операторы в \mathbb{R}^n , дополнительно необходимо согласование норм векторов и матриц, чтобы была возможность оценки образов Ax . Норма матрицы $\|A\|$ называется *согласованной* с нормой вектора $\|x\|$, если для любых A и x

$$\|Ax\| \leq \|A\| \|x\|,$$

и для любых матриц A и B выполняется неравенство $\|AB\| \leq \|A\| \|B\|$.

Чтобы неравенство $\|Ax\| \leq \|A\| \|x\|$ давало хорошую оценку, оно должно быть неулучшаемо. Это мотивирует выделение в самостоятельное понятие *подчинённой нормы* матрицы:

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

Нормам $\|\cdot\|_m$ и $\|\cdot\|_l$ подчинены, соответственно,

$$\|A\|_m = \max_i \sum_{j=1}^n |a_{ij}|, \quad (\text{строчная норма}),$$

и

$$\|A\|_l = \max_j \sum_{i=1}^n |a_{ij}|, \quad (\text{столбцовая норма}).$$

¹⁴В качестве которого может быть «назначено» любое выпуклое центрально-симметричное тело. Для такого вывода надо обратиться к определению нормы. Нормой вектора $x \in \mathbb{R}^n$ называют положительное число $\|x\|$, удовлетворяющее следующим требованиям:

1. $\|x\| = 0$ в том случае, когда $x = 0$;
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника);
3. $\|\alpha x\| = |\alpha| \cdot \|x\|$ для любого $\alpha \in (-\infty, \infty)$.

6.6 Дифференцирование оператора

Вернёмся к нелинейным операторам: Оглядываясь назад, можно заметить, что дифференцируемость до сих пор сводилась к возможности линейной аппроксимации приращения функции $f(\mathbf{x})$,

$$\Delta f = \zeta \cdot \Delta \mathbf{x} + o(\|\Delta \mathbf{x}\|), \quad (6.18)$$

в окрестности рассматриваемой точки. В одномерном случае ζ было численным коэффициентом, равным $f'(\mathbf{x})$. Для скалярной функции n переменных внешне все осталось тем же самым. Вплоть до $\zeta = f'(\mathbf{x})$, но $f'(\mathbf{x})$ — стало *вектором (градиентом)*, а произведение в (6.18) — *скалярным*.

В нашем случае, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, тоже ничего не меняется, внешне. По-прежнему $\zeta = f'(\mathbf{x})$, но теперь $f'(\mathbf{x})$ это матрица размера $n \times n$,

$$f'(\mathbf{x}) = \left[\frac{\partial f_i}{\partial x_j} \right],$$

которую называют *производной Фреше*, или *матрицей Якоби*¹⁵. И теперь $f'(\mathbf{x}) \cdot \Delta \mathbf{x}$ — это умножение матрицы на вектор.

Короче говоря, снова имеем¹⁶

$$\boxed{\Delta f = f'(\mathbf{x}) \cdot \Delta \mathbf{x} + o(\|\Delta \mathbf{x}\|)}, \quad (6.19)$$

причём ничего нового — по сравнению со случаем скалярной функции n переменных — здесь нет. Потому что $\mathbf{y} = f(\mathbf{x})$ в данном случае представляет собой набор n скалярных функций $y_i = f_i(\mathbf{x})$, для каждой из которых по теореме 5.5.1

$$\Delta y_i = \nabla f_i(\mathbf{x}) \cdot \Delta \mathbf{x} + o(\|\Delta \mathbf{x}\|).$$

Объединение n таких неравенств в одну формулу (6.19) никакого нового содержания не привносит, за исключением стенографического. Просто для краткости используется матричный язык. Другое дело, что язык — вещь продуктивная.

¹⁵ Вообще говоря, могут рассматриваться отображения f из \mathbb{R}^n в \mathbb{R}^m при $n \neq m$. В этом случае матрица Якоби будет прямоугольной, размера $m \times n$.

¹⁶ Для справедливости (6.19) необходимо выполнение дополнительных условий типа непрерывности частных производных, см. теорему 5.5.1.

Выгоды использования матричного языка выявляются буквально на каждом шагу. Скажем, умножение матриц Якоби сводит громоздкое дифференцирование сложной функции $y = f(g(x))$ в одну строчку¹⁷

$$y'_x = f'_y \cdot g'_x.$$

6.6.1 Теорема о среднем. Пусть оператор f непрерывно дифференцируем, и в некоторой норме $\|f'(z)\| \leq \lambda$ на отрезке, соединяющем точки x и y . Тогда

$$\|f(x) - f(y)\| \leq \lambda \|x - y\|. \quad (6.20)$$

◀ *Доказательство.* Разобьем отрезок $z = \tau x + (1 - \tau)y$, соединяющий x и y , точками z_i на малые участки (полагая $z_0 = x$, $z_N = y$). В силу (6.19)

$$\|\Delta f_i\| \leq \lambda \|\Delta z_i\| + o(\|\Delta z_i\|),$$

где $\Delta f_i = f(z_{i+1}) - f(z_i)$. Отсюда ясно, что для любого $\varepsilon > 0$ можно выбрать настолько мелкое разбиение, что

$$\|\Delta f_i\| \leq (\lambda + \varepsilon) \|\Delta z_i\|.$$

Но тогда

$$\|f(x) - f(y)\| = \left\| \sum_i \Delta f_i \right\| \leq \sum_i \|\Delta f_i\| \leq (\lambda + \varepsilon) \sum_i \|\Delta z_i\| = (\lambda + \varepsilon) \|x - y\|.$$

Окончательный вывод даёт предельный переход $\varepsilon \rightarrow 0$. ▶

6.7 Принцип сжимающих отображений

Многие уравнения естественно или принудительно сводятся к разрешимости уравнения

$$x = f(x), \quad (x \in X). \quad (6.21)$$

Решение $x^* = f(x^*)$ называют *неподвижной точкой* оператора f .

6.7.1 Определение. Отображение $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ называется *сжимающим*, если существует такое $\lambda < 1$, что

$$\|f(x) - f(y)\| \leq \lambda \|x - y\| \quad \text{для любых } x, y \in \mathbb{R}^n. \quad (6.22)$$

¹⁷Вместо n^2 равенств

$$\frac{\partial y_i}{\partial x_j} = \sum_k \frac{\partial f_i}{\partial g_k} \frac{\partial g_k}{\partial x_j}.$$

6.7.2 Теорема. *Всякое сжимающее отображение $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ имеет неподвижную точку x^* , которая единственна. Последовательные приближения $x^{k+1} = f(x^k)$ сходятся к x^* независимо от начального положения x^0 .*

◀ Покажем, что любая последовательность x^k , определяемая механизмом $x^{k+1} = f(x^k)$, является последовательностью Коши. Очевидно,

$$\|x^n - x^m\| \leq \|x^n - f(x^n)\| + \|f(x^n) - f(x^m)\| + \|f(x^m) - x^m\|,$$

т. е.

$$\|x^n - x^m\| \leq \varphi(x^n) + \lambda \|x^n - x^m\| + \varphi(x^m),$$

где $\varphi(x) = \|x - f(x)\|$. Следовательно,

$$\|x^n - x^m\| \leq \frac{\varphi(x^n) + \varphi(x^m)}{1 - \lambda} \rightarrow 0 \quad (n, m \rightarrow \infty).$$

Здесь мы воспользовались тем, что $\varphi(x)$ на любой последовательности $x^{k+1} = f(x^k)$ убывает до нуля, поскольку $\varphi(x^{k+1}) \leq \lambda \varphi(x^k)$, что вытекает из (6.22). Поэтому $x_k \rightarrow x^*$. В силу (6.22) оператор f непрерывен. Следовательно, $x^n = f(x^*)$. Двух неподвижных точек быть не может, благодаря тому же неравенству (6.22). ▶

6.8 Обратные и неявные функции

При решении системы уравнений (6.10), т. е. $y = f(x)$, существенную роль играет зависимость решения x от y , остаётся ли x единственным решением при изменении y , непрерывно ли зависит от y . В отсутствие непрерывности, например, получалась бы глупость. Параметры y меняются чуть-чуть, а решение x улетает в тартарары. Короче говоря, проблема упирается в обратимость преобразования f в окрестности рассматриваемой точки.

6.8.1 Теорема. *Пусть оператор f , действующий в \mathbb{R}^n , непрерывно дифференцируем в окрестности нуля¹⁸, $f(0) = 0$ и матрица Якоби $f'(0)$ невырождена. Тогда в некоторой окрестности нуля оператор f обратим.*

◀ Исходному $y = f(x)$ эквивалентно уравнение $x = \Phi(x, y)$ с оператором

$$\Phi(x, y) = x + [f'(0)]^{-1}[y - f(x)],$$

¹⁸От $a = f(b)$ к $0 = f(0)$ всегда ведёт замена координат.

имеющим в нуле нулевую производную Фреше,

$$\Phi'_0(0, \mathbf{y}) = I - [f'(0)]^{-1} f'(0) = [0].$$

Поскольку матрица $\Phi'_0(0, \mathbf{y})$ нулевая (при любом \mathbf{y}), а сама производная $\Phi'_0(\mathbf{x}, \mathbf{y})$ непрерывна и не зависит от \mathbf{y} , то при некотором $\varepsilon > 0$ в шаре $\|\mathbf{x}\| < \varepsilon$ можно обеспечить такую малость элементов $\Phi'(\mathbf{x}, \mathbf{y})$, что

$$\|\Phi'(\mathbf{x}, \mathbf{y})\| < \lambda < 1.$$

Поэтому (теорема 6.6.1) в шаре $\|\mathbf{x}\| < \varepsilon$ оператор $\Phi(\mathbf{x}, \mathbf{y})$ сжимает с коэффициентом λ при любом \mathbf{y} , причём, в силу $\Phi(0, 0) = 0$ и той же теоремы 6.6.1, $\Phi(\mathbf{x}, 0)$ переводит шар $\|\mathbf{x}\| < \varepsilon$ в шар меньшего радиуса $\lambda\varepsilon$. Это, в совокупности с непрерывностью по \mathbf{y} , гарантирует, что при малых \mathbf{y} (из некоторого шара $\|\mathbf{y}\| < \delta$) оператор $\Phi(\mathbf{x}, \mathbf{y})$ переводит шар $\|\mathbf{x}\| < \varepsilon$ в себя. Далее остается применить теорему 6.7.2, которая обеспечивает однозначную разрешимость уравнения $\mathbf{y} = f(\mathbf{x})$ в окрестности $\{\|\mathbf{x}\| < \varepsilon, \|\mathbf{y}\| < \delta\}$. ►

Результат 6.8.1 интуитивно понятен. В окрестности нуля уравнение $\mathbf{y} = f(\mathbf{x})$ в силу (6.19) равносильно

$$\mathbf{y} = \left[\frac{\partial f_i(\mathbf{0})}{\partial x_j} \right] \cdot \mathbf{x} + o(\|\mathbf{x}\|),$$

откуда

$$\mathbf{x} = \left[\frac{\partial f_i(\mathbf{0})}{\partial x_j} \right]^{-1} \cdot \{\mathbf{y} - o(\|\mathbf{x}\|)\},$$

и роль невырожденности матрицы $\left[\frac{\partial f_i(\mathbf{0})}{\partial x_j} \right]$ интуитивно просматривается. Однако «дух приближительности» всё же остаётся, а дожимая задачу, приходим в конце концов к необходимости доказательства типа данного выше.

Локальная обратимость отображения представляет собой частный случай следующей более общей задачи. Если в некоторой окрестности точки (x_0, \mathbf{y}_0) , для которой $\Psi(x_0, \mathbf{y}_0) = 0$, векторное уравнение¹⁹

$$\Psi(\mathbf{x}, \mathbf{y}) = 0 \tag{6.23}$$

однозначно разрешимо относительно \mathbf{x} , то говорят, что (6.23) определяет в окрестности (x_0, \mathbf{y}_0) неявную функцию $\mathbf{x}(\mathbf{y})$.

¹⁹ Имеется в виду $\Psi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$.

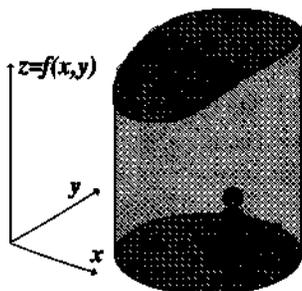
Глава 7

Кратные интегралы

*Спасибо тебе, Господи,
что ты создал всё нужное нетрудным,
а всё трудное — невозможным.*
Георгий Сковорода

7.1 Двойные интегралы

Интегрировать можно (и бывает нужно) не только вдоль прямой, но и по площадям, объёмам и вообще по любым множествам, на которых задана мера.



Объём V цилиндрического тела, стоящего на основании D в плоскости xy и ограниченного в направлении z поверхностью

$$z = f(x, y),$$

можно приблизить суммой

$$V_{\Delta} = \sum_i f(x_i, y_i) \Delta x_i \Delta y_i,$$

равной сумме объёмов вертикальных столбиков высотой $f(x_i, y_i)$, стоящих на малых квадратиках с площадью $\Delta x_i \Delta y_i$.

Предел V_{Δ} при измельчении разбиения фигуры D называется

двойным интегралом функции f по области D и обозначается

$$\boxed{\iint_D f(x, y) dx dy = \lim_{\Delta \rightarrow 0} V_\Delta} \quad (7.1)$$

Заодно возникает интерпретация двойного интеграла как объёма, заключённого под графиком $z = f(x, y)$. Объёмы, где $f(x, y) > 0$ или < 0 , засчитываются со знаком *плюс/минус*. Коротче говоря, все развивается по образцу определённого интеграла от функции одного переменного, только промежутки заменяются плоскими фигурами. Аналогично вводятся *суммы Дарбу* и т.п.¹

В случае $z = f(x, y) \equiv 1$

$$\iint_D dx dy = S_D,$$

где S_D — площадь фигуры D . Таким образом, двойной интеграл может использоваться для вычисления площадей.

Что касается возможного диапазона обозначений, то вместо произведения дифференциалов $dx dy$ иногда пишут дифференциал площади dS , вместо \iint — просто \int , если из контекста ясно, о чём идет речь. Наконец область D может заменяться её описанием типа $x^2 + y^2 \leq 1$.

Сведение двойного интеграла к повторному. Практически — это единственный метод вычисления двойного интеграла, заслуживающий внимания. Снова всплывает теория повторных пределов. Другое дело, что в случае непрерывных функций $f(x, y)$ и простых ограниченных областей D никаких проблем не

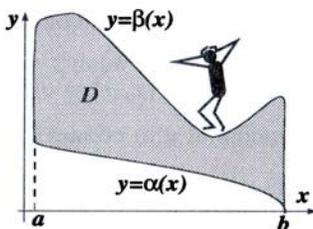
¹Принципиальная тонкость здесь заключается в допуске к рассмотрению только *квадрируемых областей* D , которые характеризуются тем, что *ограничены контуром, имеющим нулевую площадь*. *Площадь кривой*, в свою очередь, определяется (упрощённо говоря) как предел суммы покрывающих кривую квадратиков при бесконечном их измельчении. Для простых областей D никаких проблем не возникает. Более того, придумать «проблемную» область D — не так легко.

возникает. Например, в условиях прямоугольной области повторные интегралы из утверждения 5.9.3 оказываются равны как раз двойному интегралу

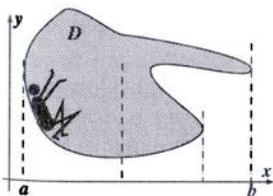
$$\iint_{[a,b] \times [c,d]} f(x,y) dx dy = \int_c^d dy \int_a^b f(x,y) dx = \int_a^b dx \int_c^d f(x,y) dy,$$

что представляет собой не что иное, как равенство повторных пределов двойному² для аппроксимирующей суммы V_Δ .

Рассмотрим теперь область D , ограниченную кривыми $y = \alpha(x)$ и $y = \beta(x)$, переменная x меняется от a до b . Опять-таки, если все фигурирующие функции непрерывны, то, очевидно,



$$\iint_D f(x,y) dx dy = \int_a^b dx \int_{\alpha(x)}^{\beta(x)} f(x,y) dy.$$

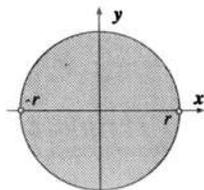


Здесь возможны некоторые трудности в связи с неоднозначностью функций типа $\alpha(x)$ и $\beta(x)$, см. выше. Выход из положения подсказывает рис. слева. Отрезок $[a, b]$ подразделяется на промежутки, в пределах которых область D можно покусочно описывать прежним образом.

² Двойные интегралы могут рассматриваться и на бесконечных областях.

Пример. Проверим рецепт на вычислении площади круга, $x^2 + y^2 \leq r^2$,

$$S = \int_{-r}^r dx \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} dy = 2 \int_{-r}^r \sqrt{r^2-x^2} dx = \pi r^2.$$



7.2 Замена переменных

При вычислении двойных интегралов довольно часто эффективным инструментом является замена переменных

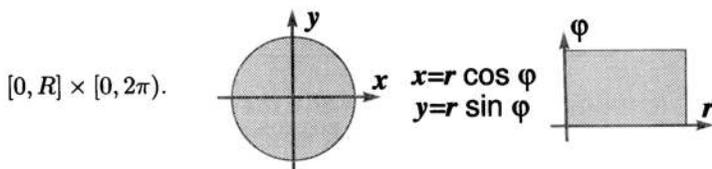
$$\{x, y\} = \theta(\xi, \eta) = \{x(\xi, \eta), y(\xi, \eta)\}, \quad (7.2)$$

упрощающая или область D , или функцию $f(x, y)$.

Например, в полярных координатах $\{r, \varphi\}$,

$$\begin{cases} x = r \cos \varphi, \\ y = r \sin \varphi, \end{cases}$$

круг $x^2 + y^2 \leq R^2$ оказывается прямоугольником



Функция $f(x, y)$ меняется при этом на $f(r \cos \varphi, r \sin \varphi)$. Остается вопрос о соотношении дифференциалов площади $dx dy$ и $dr d\varphi$. Коэффициент искажения объёма — в данном случае площади — в малой окрестности точки $\{r, \varphi\}$ определяется величиной $|\det \theta'(r, \varphi)| = r$,

$$\theta'(r, \varphi) = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{bmatrix} = \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix}.$$

Следовательно,

$$\iint_{x^2+y^2 \leq R^2} dx dy = \int_0^{2\pi} d\varphi \int_0^R r dr = \pi R^2,$$

что с вычислительной точки зрения проще прямолинейного взятия интеграла, использованного выше.

В общем случае замены (7.2) формула перехода к новым переменным такова

$$\iint_{D_{xy}} f(x, y) dx dy = \iint_{D_{\xi\eta}} f(x(\xi, \eta), y(\xi, \eta)) |\det \theta'(\xi, \eta)| d\xi d\eta. \quad (7.3)$$

Наличие модуля под интегралом (7.3) контрастирует с правилом замены переменной в определённом интеграле³

$$\int_a^b f(x) dx = \int_\alpha^\beta f[x(\xi)] x'(\xi) d\xi, \quad (7.4)$$

где $a = x(\alpha)$, $b = x(\beta)$. Очевидно, $|x'(\xi)|$ в (7.4) является коэффициентом искажения длины, но знак модуля отсутствует. Причина заключается в том, что промежуток интегрирования в (7.4) *ориентирован* ($a < b$ либо $a > b$), и от этого зависит знак интеграла. Двойной же интеграл определялся пока на *неориентированных областях*. Однако возможен другой подход. Области придаётся ориентация заданием направления обхода ограничивающего контура — положительного или отрицательного. Соответственно, площади ориентированной области приписывается знак плюс или минус. И тогда в формуле замены переменных модуль с детерминанта снимается.

³ Вытекающим из формулы (4.4) для неопределённого интеграла.

7.3 Кратные интегралы

Тройной интеграл по области D от функции $u = f(x, y, z)$ определяется как предел

$$\boxed{\iiint_D f(x, y, z) dx dy dz = \lim_{\Delta \rightarrow 0} V_\Delta},$$

где

$$V_\Delta = \sum_i f(x_i, y_i, z_i) \Delta x_i \Delta y_i \Delta z_i. \quad (7.5)$$

Аналогия с двойным интегралом полная. Область D режется на малые кубики (со сторонами $|\Delta x_i|, |\Delta y_i|, |\Delta z_i|$), и в качестве отправной точки рассматриваются суммы (7.5). Дальнейшее развитие теории (в том числе интегралов большей кратности) идет по тому же пути, что и двойных интегралов. Суммы Дарбу, повторное интегрирование, несобственные интегралы и т. п.⁴

7.4 Объёмы n -мерных тел

В описании стандартных n -мерных тел обычно фигурирует некий характерный размер a , входящий в неравенства вида $\varphi(\mathbf{x}, a) \leq 0$.

Скажем, n -мерный шар описывается неравенством $x_1^2 + \dots + x_n^2 \leq a^2$.

Функции φ , как правило, однородны, т. е.

$$\varphi(\tau \mathbf{x}, \tau a) = \tau^n \varphi(\mathbf{x}, a),$$

что позволяет с помощью замены $x_j = a \xi_j$ перейти к рассмотрению «единичных» тел ($a = 1$).

При вычислении объёма это приводит к формуле

$$\boxed{V(a) = a^n V(1)}, \quad (7.6)$$

поскольку $\det x'_\xi = a^n$ и

$$V(a) = \int_{V_a} dV_x = \int_{V_1} |\det x'_\xi| dV_\xi.$$

⁴Замена переменных происходит по формуле аналогичной (7.3), где у Θ' — другое количество переменных.

Таким образом, объём n -мерного куба со стороной a равен

$$V(a) = a^n.$$

При вычислении n -мерных интегралов широко применяется индукция по размерности пространства, обыгрывающая симметрию задачи. Например, для вычисления объёма $T_n(h)$ n -мерного симплекса:

$$x_1 \geq 0, \dots, x_n \geq 0; \quad x_1 + \dots + x_n \leq h,$$

к цели быстро ведёт преобразование

$$\begin{aligned} T_n(1) &= \int_0^1 d\xi_n \overbrace{\int \dots \int}^{n-1} d\xi_1 \dots d\xi_{n-1} = \\ &= T_{n-1}(1) \int_0^1 (1 - \xi_n)^{n-1} d\xi_n = \frac{T_{n-1}(1)}{n}, \end{aligned}$$

где все $\xi_i \geq 0$ и

$$T_{n-1}(1 - \xi_n) = (1 - \xi_n)^{n-1} T_{n-1}(1)$$

в силу (7.6). Учитывая $T_1(1) = 1$, последовательно приходим к $T_n(1) = \frac{1}{n!}$.

Окончательно

$$T_n(h) = \frac{h^n}{n!}.$$

При вычислении объёма $B_n(r)$ n -мерного шара:

$$x_1^2 + \dots + x_n^2 \leq r^2$$



работает аналогичный трюк,

$$\begin{aligned} B_n(1) &= \int_{-1}^1 d\xi_n \overbrace{\int \dots \int}^{n-1} d\xi_1 \dots d\xi_{n-1} = \\ &= B_{n-1}(1) \int_{-1}^1 (1 - \xi_n^2)^{\frac{n-1}{2}} d\xi_n = 2B_{n-1}(1) \int_0^{\frac{\pi}{2}} \sin^n \omega d\omega. \end{aligned}$$

В итоге преодоления рутинности для чётной и нечётной размерности пространства получается:

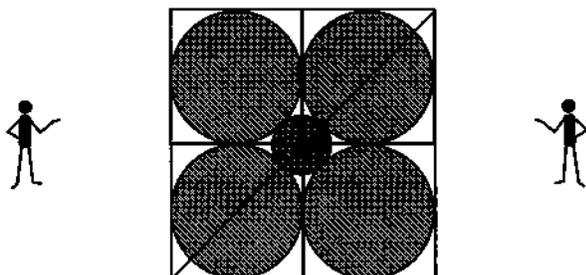


$$B_{2k}(r) = \frac{\pi^k}{k!} r^{2k}, \quad B_{2k+1}(r) = \frac{2^{k+1} \pi^k}{(2k+1)!!} r^{2k+1}. \quad (7.7)$$

7.5 Сюрпризы n измерений

• Из (7.7) следует, что объём n -мерного шара $V_n(r)$ любого фиксированного радиуса (скажем, $r = 1000\text{км}$) при $n \rightarrow \infty$ *стремится к нулю!*

• Пусть квадрат с ребром 2 разбит на 4 единичных квадрата, в каждый из которых вписан круг диаметра $D = 1$. В центральную область вписан круг диаметра d , касающийся остальных кругов. Очевидно, $d = \sqrt{2} - 1$.



Аналогичным образом в \mathbb{R}^3 куб с ребром 2 можно разбить на 8 одинаковых кубов, в каждый из которых вписать шар диаметра $D = 1$. Диаметр центрального вписанного шарика будет равен $d = \sqrt{3} - 1$. Продолжая в том же духе, в \mathbb{R}^n будем иметь диаметр вписанного шарика

$$d = \sqrt{n} - 1.$$

Таким образом, вписанный шарик, начиная с $n = 10$, вылезает за пределы ограничивающего куба, не говоря о том, что, начиная с $n = 5$, радиус «маленького» шарика превосходит радиус «большого».

• *Какое максимальное число вершин может иметь многогранник, у которого каждые две вершины соединены ребром?* На плоскости \mathbb{R}^2 указанным свойством обладает лишь треугольник — максимальное число вершин $m = 3$. В \mathbb{R}^3 — тетраэдр ($m = 4$). В пространстве четырёх измерений и выше такой многогранник *может иметь вершин сколько угодно!*

◀ *Доказательство.* Рассмотрим так называемый *циклический многогранник* в \mathbb{R}^4 с любым числом вершин, лежащих на кривой

$$x(\tau) = \{\tau, \tau^2, \tau^3, \tau^4\}, \quad \tau \geq 0.$$

Кривая $x(\tau)$ обладает поразительным свойством. Через любые её две точки $x(t)$, $x(s)$ можно так провести плоскость H (в данном случае трёхмерную, коль речь идет об \mathbb{R}^4), что все её точки $x(\tau)$ будут находиться по одну сторону H . Покажем это. Рассмотрим функцию

$$\varphi(\tau) = (\tau - t)^2(\tau - s)^2 = \alpha_0 + \alpha_1\tau + \alpha_2\tau^2 + \alpha_3\tau^3 + \alpha_4\tau^4,$$

и в качестве плоскости H возьмём множество точек x , для которых

$$h \cdot x + \alpha_0 = 0, \quad h = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}.$$

Очевидно, $h \cdot x(\tau) + \alpha_0 = \varphi(\tau) \geq 0$, значит, вся кривая $x(\tau)$ лежит по одну сторону H . Кроме того ясно, что лишь две точки $x(t)$, $x(s)$ попадают на самую плоскость H ,

$$\varphi(t) = \varphi(s) = 0.$$

Теперь остается вспомнить, что многогранник является выпуклой оболочкой своих вершин, в данном случае точек

$$x(\tau_1), \dots, x(\tau_m). \tag{7.8}$$

Все плоскости типа H , как говорят, *опорные*. Отрезок, соединяющий $x(\tau_i)$ и $x(\tau_j)$, весь лежит в H , и ни одна его точка не может быть получена как выпуклая комбинация остальных вершин (7.8), поскольку те лежат *строго* по одну сторону H . Поэтому отрезки, соединяющие любые две точки из (7.8), будут ребрами искомого многогранника. ▶

Глава 8

Векторный анализ

*Ability may get you to the top,
but it takes character to keep you there.*

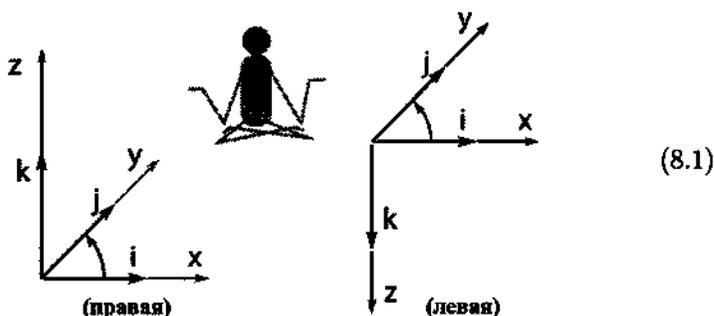
John Wooden

8.1 Координаты и векторы

Вернёмся к тому, что вкратце рассматривалось в разделе 5.1. Векторные представления возникают естественно и вроде бы бесхитростно. По крайней мере изначально многообещающими не выглядят. Но шаг за шагом постепенно выводят на новые этажи виденья, с которых в «покоординатной писанине» начинают просматриваться совершенно новые понятия, которые «за деревьями» детальных описаний были не видны.

При изучении функций $u = f(x, y, z)$ никакого визуального пространства изначально нет. Есть функция и аргументы. Переменные, параметры — что угодно, но никакой геометрии. Пространство возникает как виртуальный инструмент, когда значения x, y, z мысленно откладываются по осям выдуманного пространства. Разумеется, затея с геометрией имеет смысл лишь в том случае, если она хоть что-то даёт. Виденье, понимание или хотя бы вдохновение. Как ни странно, — даёт всё. Аморфные алгебраические факты приобретают наглядную интерпретацию, настроение поднимается, задачи начинают решаться.

Поскольку геометризация анализа идёт «координатно», возникает вопрос о целесообразном выборе системы координат. В большинстве ситуаций наиболее удобны *ортогональные системы координат* со взаимно перпендикулярными осями (x, y, z) , по которым направлены *единичные орты* $\{i, j, k\}$. *Координатами* при этом называют числа, определяющие положение точки $r = \{x, y, z\}$ в пространстве \mathbb{R}^3 и представляющие собой ортогональные проекции r на оси (x, y, z) .



Подготавливая почву для обобщений, естественно нумеровать оси координат и орты как (x_1, x_2, x_3) и $\{e_1, e_2, e_3\}$, что потом облегчает следующий шаг — тройка увеличивается до n , и обобщение \mathbb{R}^3 на \mathbb{R}^n готово. Но в \mathbb{R}^3 привычнее и в некоторых отношениях удобнее система $Oxyz$, в которой $a = \{a_x, a_y, a_z\}$. Поэтому ей тоже будем пользоваться наравне с $Ox_1x_2x_3$.

Однако вернёмся к нашим баранам. Геометрически *вектор* определяют как направленный отрезок¹. При этом *все векторы, которые одинаково направлены и равны по длине, считаются равными, что позволяет ограничиться рассмотрением векторов, исходящих из начала координат*².

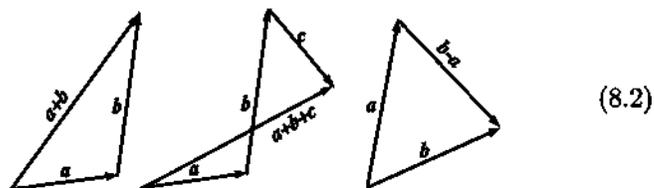
Сумма $x = \{x_1, x_2, x_3\}$ и $y = \{y_1, y_2, y_3\}$ определяется как

$$x + y = \{x_1 + y_1, x_2 + y_2, x_3 + y_3\},$$

¹Но этого мало, см. далее.

²Такова точка зрения в идеологии свободных векторов.

что равносильно сложению векторов по правилу параллелограмма, эквивалентом которого является *правило треугольника*.



Преимущества последнего выявляются при сложении нескольких векторов, рис. (8.2). Каждый следующий слагаемый вектор приставляется началом к концу предыдущего — замыкающий вектор даёт сумму. Вычитание выводится из сложения: $b - a$ определяется как вектор, который в сумме с a даёт b . Этому соответствует простой геометрический трюк: начала a и b совмещаются, а концы соединяются отрезком, направленным к b , что и даёт разность $b - a$, рис. (8.2).

Умножение на скаляр $\lambda x = \{\lambda x_1, \lambda x_2, \lambda x_3\}$ растягивает ($|\lambda| > 1$) или сжимает ($|\lambda| < 1$) вектор x , не меняя направления при $\lambda > 0$, и меняет его на противоположное при $\lambda < 0$.

Определение вектора просто как направленного отрезка ведёт к проблемам, мешая алмазы с пустой породой. Скажем, трёхмерное вращение на угол φ — вектор или не вектор? С одной стороны, ему можно сопоставить направленный по оси отрезок прямой длины φ . Но что толку? Неприятность заключается в том, что вращения около разных осей не складываются по правилу параллелограмма³, — и это даёт отрицательный ответ на исходный вопрос. Поэтому в подобное определение вектора необходимо добавить *требование, чтобы векторы складывались по правилу параллелограмма*. В данном контексте речь идёт об элементах вида $x = \{x_1, x_2, x_3\}$, которые складываются «как надо», по определению. Что касается правил сложения сил, моментов, потоков и других физических величин — это задача из другой оперы.

Конечно, геометрия имеет изначально «некоординатный» характер, и её координатное описание часто уводит мысль, вообще

³Складываются по правилу параллелограмма бесконечно малые вращения, что влечёт за собой векторную природу угловой скорости, см. далее.

говоря, в противоположном направлении. Чисто векторный язык геометрически точнее. Но полезны оба языка, и максимальный эффект даёт их параллельное использование.

Векторы, лежащие на одной прямой (одинаково или противоположно направленные), называют *коллинеарными*; лежащие в одной плоскости, — *компланарными*.

Говорят, что множество векторов⁴ $\{x^1, \dots, x^k\}$ *линейно зависимо*, если существуют такие коэффициенты $\lambda_1, \dots, \lambda_k$, не все равные нулю, что

$$\lambda_1 x^1 + \dots + \lambda_k x^k = 0.$$

В противном случае говорят о *линейной независимости* векторов $\{x^1, \dots, x^k\}$.

Коллинеарные векторы всегда линейно зависимы. Компланарные — линейно зависимы, если их больше двух. (?) Максимальное число линейно независимых векторов определяет *размерность пространства*.

Линейно независимое множество $\{e_1, e_2, e_3\}$ в пространстве считается *базисом*, если любой вектор x можно представить в виде линейной комбинации

$$x = x_1 e_1 + x_2 e_2 + x_3 e_3.$$

Величины x_i называют *координатами* точки $x \in \mathbb{R}^3$.

Стандартный базис \mathbb{R}^3 (единичные векторы, орты, $\{e_1, e_2, e_3\}$ направлены по осям декартовых координат):

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (8.3)$$

Упражнения

1. Из векторов a, b, c можно сложить треугольник, если $a + b + c = 0$.
2. Любые три вектора на плоскости либо четыре — в пространстве — линейно зависимы.

⁴Векторы с индексом мы выделяем жирным шрифтом, чтобы отличить их от координат.

3. Любые два неколлинеарных вектора a , b определяют плоскость⁵, все точки которой могут быть записаны в виде

$$r = \lambda a + \mu b.$$

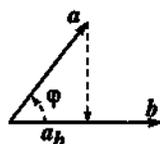
4. В случае $\lambda + \mu = 1$ точка $r = \lambda a + \mu b$ лежит на прямой, проходящей через концы векторов a и b . При *дополнительном* условии $\lambda, \mu \geq 0$ точка $r = \lambda a + \mu b$ лежит на отрезке, соединяющем a и b .
5. Если a , b , c — радиус-векторы вершин треугольника ABC , то

$$r = \frac{1}{3}(a + b + c)$$

является радиус-вектором точки пересечения медиан.

Проекция a_b вектора a на вектор b , направление b , определяется формулой

$$a_b = |a| \cos \varphi,$$



где φ — угол между a и b , неважно по малой или большой дуге измеренный — в силу $\cos \varphi = \cos(2\pi - \varphi)$.

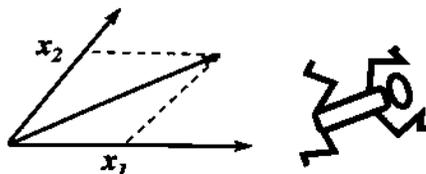
Проекция a_x , a_y , a_z на декартовы оси x , y , z представляют собой как раз координаты вектора a в системе $Oxyz$.

В некоторых ракурсах теории важным оказывается деление систем координат на две категории. Прямоугольную систему $Oxyz$ называют *правой*, если буравчик при вращении от x к y движется вдоль z — рис. (8.1), и *левой*, если ось z направлена противоположно — рис. (8.1).

Косоугольные координаты. Иногда декартовой называют косоугольную систему координат общего вида, в которой координаты определяются разложением вектора по двум (на плоскости) или по трем (в пространстве) направлениям, которые не обязательно взаимно перпендикулярны. Атрибутом «декартовости» при этом является наличие базиса. Примером недекар-

⁵ Проходящую через три точки: нуль, a и b .

товых координат могут служить (нелинейные) полярные координаты (вектор характеризуется длиной r и углом φ с выделенным направлением).



В общем случае непрямоугольных координат понятие ориентации опирается на следующее определение. Упорядоченная тройка некопланарных векторов a, b, c , исходящих из точки O , называется правой, если для наблюдателя, расположенного в нуле, обход концов a, b, c в указанном порядке происходит по часовой стрелке. В противном случае тройка a, b, c — левая. Соответственно классифицируются базисы.

Если не оговорено противное, под декартовыми — подразумеваются далее прямоугольные координаты.

8.2 Скалярное произведение

Скалярное произведение определяется как произведение длин векторов на косинус угла между ними⁶,

$$x \cdot y = |x| \cdot |y| \cos \varphi \quad (8.4)$$

Другими словами, $x \cdot y$ есть произведение длины $x = |x|$ на проекцию вектора y на вектор x , т. е.⁷

$$xy = x \cdot y_x.$$

⁶Из-за четности и периодичности косинуса, $\cos(2\pi - \varphi) = \cos \varphi$, — не важно, как измеряется угол.

⁷Точка в обозначении скалярного произведения часто опускается. Для обозначения xy используется также запись (x, y) или $\langle x, y \rangle$.

Скалярное произведение удовлетворяет обычным свойствам умножения:

$$\begin{aligned} \mathbf{x}\mathbf{y} &= \mathbf{y}\mathbf{x}, & (\text{коммутативность}) \\ \mathbf{x}(\mathbf{y} + \mathbf{z}) &= \mathbf{x}\mathbf{y} + \mathbf{x}\mathbf{z}. & (\text{дистрибутивный закон}) \end{aligned}$$

Второе вытекает из равенства проекции суммы векторов — сумме проекций.

Отметим попутно, что длина $|\mathbf{a}|$ (ещё говорят — норма, и пишут $\|\mathbf{a}\|$) вектора $\mathbf{a} = \{a_x, a_y\}$ по теореме Пифагора равна

$$\|\mathbf{a}\| = \sqrt{a_x^2 + a_y^2},$$

т. е. $\mathbf{a}^2 = \mathbf{a} \cdot \mathbf{a} = a_x^2 + a_y^2$. В пространстве, соответственно,

$$\|\mathbf{a}\|^2 = a^2 = a_x^2 + a_y^2 + a_z^2,$$

откуда

$$(\mathbf{a} \pm \mathbf{b})^2 = (a_x \pm b_x)^2 + (a_y \pm b_y)^2 + (a_z \pm b_z)^2,$$

что приводит к

$$\mathbf{a} \cdot \mathbf{b} = \frac{(\mathbf{a} + \mathbf{b})^2 - (\mathbf{a} - \mathbf{b})^2}{4} = a_x b_x + a_y b_y + a_z b_z.$$

Последнюю формулу

$$\boxed{\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z} \quad (8.5)$$

иногда принимают за исходное определение скалярного произведения, выводя (8.4) в качестве следствия (8.5).

Произведение $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| \cdot |\mathbf{b}| \cos \varphi$ отражает широко распространённый в природе способ взаимодействия векторов, характеризуемый умножением длины одного вектора на проекцию другого: $|\mathbf{a}| b_x$. Работа, например силы \mathbf{F} на перемещении \mathbf{s} , равна

$$\mathbf{F} \cdot \mathbf{s} = |\mathbf{F}| \cdot |\mathbf{s}| \cos \varphi,$$

поскольку «работает» только составляющая \mathbf{F} , направленная вдоль \mathbf{s} .

Поток жидкости через площадку площади ΔS равен скалярному произведению $v \Delta S$, где v — скорость течения (в районе ΔS), а вектор ΔS считается направленным по нормали к площадке. Физически опять-таки понятная ситуация: количество жидкости, протекающей через ΔS , определяется нормальной составляющей потока и не зависит от касательной составляющей.

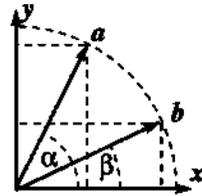
Произведение единичных векторов

$$\mathbf{a} = \{\cos \alpha, \sin \alpha\}, \quad \mathbf{b} = \{\cos \beta, \sin \beta\}$$

сразу даёт формулу «косинуса разности»,

$$\mathbf{a}\mathbf{b} = \cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta,$$

поскольку $|\mathbf{a}| = |\mathbf{b}| = 1$.



Равенство нулю скалярного произведения $\mathbf{a}\mathbf{b}$ означает — в силу $\cos(\mathbf{a}, \mathbf{b}) = 0$ — *ортогональность* ненулевых векторов \mathbf{a} и \mathbf{b} . Поэтому $\mathbf{a}\mathbf{r} = 0$ (при фиксированном \mathbf{a} и текущем $\mathbf{r} = \{x, y, z\}$) представляет собой уравнение плоскости, *ортогональной вектору \mathbf{a} и проходящей через начало координат*.

Уравнение

$$\mathbf{a}\mathbf{r} = \delta \tag{8.6}$$

задаёт *гиперплоскость*⁸, не проходящую через начало координат при $\delta \neq 0$, и означает, что проекция радиус-вектора $\mathbf{r} = \{x, y, z\}$ любой точки (8.6) на направление \mathbf{a} одна и та же. Если дополнительно, $|\mathbf{a}| = 1$, то эта проекция численно равна δ .

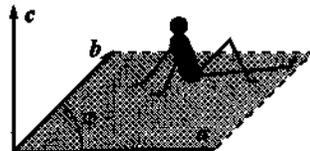
8.3 Векторное произведение

Ещё одной важной операцией является *векторное произведение*,

$$\mathbf{c} = \mathbf{a} \times \mathbf{b},$$

где вектор \mathbf{c} по длине равен

$$|\mathbf{c}| = |\mathbf{a}| |\mathbf{b}| \sin \varphi$$



⁸ *Гиперплоскость* отличается от *плоскости* тем, что не проходит через начало координат.

т. е. площади параллелограмма, построенного на векторах \mathbf{a} , \mathbf{b} , а направление \mathbf{c} определяется по «правилу буравчика».

Векторное произведение не ассоциативно,

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c}),$$

антикоммумутативно,

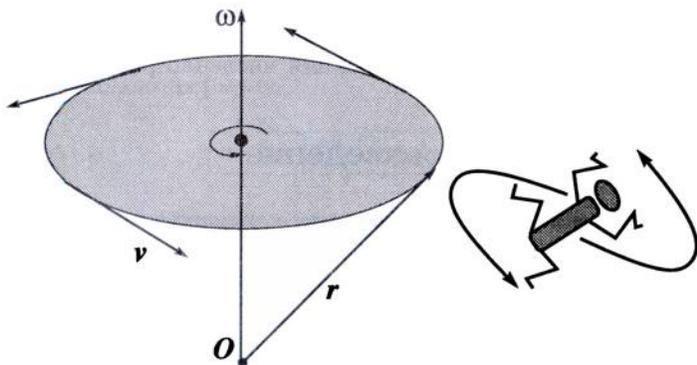
$$\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x},$$

но справедлив дистрибутивный закон,

$$(\mathbf{x} + \mathbf{y}) \times \mathbf{u} = \mathbf{x} \times \mathbf{u} + \mathbf{y} \times \mathbf{u}, \quad (8.7)$$

проверка которого требует некоторых технических усилий — зато окупается впоследствии многочисленными выгодами.

С помощью (8.7) в одно касание доказывается, что угловая скорость — вектор! Факт *важный и крайне неожиданный*, если учесть, что трёхмерные повороты не коммутируют.



При использовании (8.7) усилие мысли почти не требуется. Линейная скорость \mathbf{v} конца радиус-вектора \mathbf{r} при вращении вокруг оси, проходящей через начало координат $\mathbf{0}$, с угловой скоростью $\boldsymbol{\omega}$ равна

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}, \quad (8.8)$$

где (пока авансом) вектор $\boldsymbol{\omega}$ направлен по оси вращения в сторону, определяемую по «правилу буравчика».

Если тело участвует в двух вращениях⁹ ω^1 и ω^2 , то линейные скорости

$$\mathbf{v}^1 = \omega^1 \times \mathbf{r} \quad \text{и} \quad \mathbf{v}^2 = \omega^2 \times \mathbf{r}$$

складываются как векторы, и результирующая скорость оказывается равной, в силу (8.7),

$$\mathbf{v} = \omega^1 \times \mathbf{r} + \omega^2 \times \mathbf{r} = (\omega^1 + \omega^2) \times \mathbf{r},$$

что и даёт нужный вывод. Результирующее движение происходит с векторной угловой скоростью $\omega = \omega^1 + \omega^2$.

• При замене координат $\{x, y, z\}$ на $\{-x, -y, -z\}$ обычные векторы (их называют ещё *полярными*) меняются на противоположные (a на $-a$). Векторы типа $c = a \times b$ (их называют *аксиальными*) — не меняются. Такого рода метаморфозы вообще характерны для векторного исчисления. Например, дифференциальные операции над обычными векторами порождают вроде бы объекты той же природы, но это поверхностное впечатление. Скорость — вектор, градиент — вектор, и разница не видна, пока исследование проводится на базе ортогональных систем координат. В неортогональных системах происходит «сбой» — формулы перехода при замене переменных оказываются различны. Выявляется так называемая *контравариантная* природа скорости и *ковариантная* природа градиента. Но это уже епархия *тензорного исчисления*, где изучается кое-что незаметное в рамках анализа.

• В правой системе координат:

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}. \quad (?)$$

• Координаты вектора $\mathbf{a} \times \mathbf{b}$ равны: (?)

$$(\mathbf{a} \times \mathbf{b})_x = a_y b_z - a_z b_y,$$

$$(\mathbf{a} \times \mathbf{b})_y = a_z b_x - a_x b_z,$$

$$(\mathbf{a} \times \mathbf{b})_z = a_x b_y - a_y b_x.$$

⁹Здесь и далее мы стараемся приписывать верхние индексы векторам, а нижние — координатам.

8.4 Приложения к механике

Моменты. Моментом силы F , приложенной в точке r относительно точки O , называется вектор

$$M = r \times F.$$

Аналогично определяется момент количества движения:

$$N = r \times mv.$$

Дифференцируя N , получаем¹⁰

$$\dot{N} = \dot{r} \times mv + r \times m \dot{v} = r \times m \dot{v} = r \times F = M,$$

поскольку $m \dot{v} = F$ и $\dot{r} \times m v = v \times m v = 0$.

Если система состоит из жёстко скрепленных между собой материальных точек, то

$$N = \sum r_i \times m_i v_i, \quad M = \sum r_i \times F_i.$$

Для каждой точки справедливо уравнение $\dot{N}_i = M_i$, что после суммирования по всем i даёт

$$\boxed{\frac{dN}{dt} = M} \quad (8.9)$$

В случае вращения системы вокруг некоторой фиксированной оси с переменной угловой скоростью $\omega(t)$ скорость каждой частицы равна $v_i = r_i \omega$, где r_i расстояние до оси вращения,

$$N = \left(\sum m_i r_i^2 \right) \omega = J \omega.$$

¹⁰Обоснование правил дифференцирования скалярного и векторного произведения векторов $a(t)$ и $b(t)$, зависящих от времени,

$$\boxed{\frac{d(a \cdot b)}{dt} = \frac{da}{dt} \cdot b + a \cdot \frac{db}{dt}}, \quad \boxed{\frac{d(a \times b)}{dt} = \frac{da}{dt} \times b + a \times \frac{db}{dt}}.$$

предоставляется в качестве упражнения. Напоминаем, что в случае векторного произведения порядок сомножителей принципиален.

Сумма $J = \sum m_i r_i^2$ — есть момент инерции системы (тела).

Таким образом, (8.9) с учётом $N = J\omega$ приводит к уравнению вращательного движения¹¹

$$\boxed{J\dot{\omega} = M}, \quad \text{равносильно,} \quad \boxed{J\ddot{\varphi} = M}.$$

Кинетическая энергия с помощью момента инерции записывается как

$$W_k = \sum \frac{m_i v_i^2}{2} = \frac{1}{2} J \omega^2.$$

• Рассмотрим движение планеты вокруг Солнца. Поскольку сила притяжения F центральная, то её момент действия на планету $M = r \times F$ равен нулю, и, как следствие,

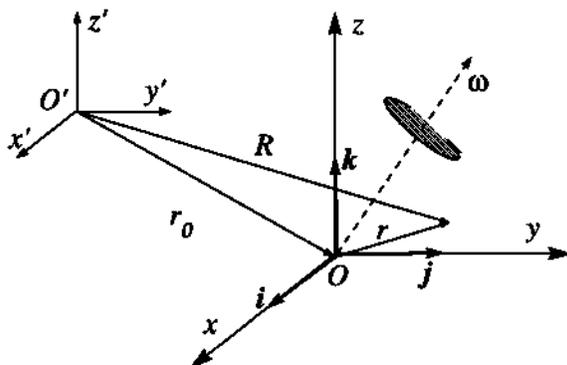
$$\dot{N} = M = 0 \quad \rightarrow \quad N = \text{const.}$$

С другой стороны, если за время dt радиус-вектор заметает площадь dS , то

$$N = r \times mv = mr \times (\omega \times r) = mr^2 \omega = \frac{mr^2 d\varphi}{dt} = 2m \frac{dS}{dt},$$

что доказывает второй закон Кеплера: $\dot{S} = \text{const.}$

Движение в неинерциальных системах. Пусть система координат $Oxyz$ движется относительно $Ox'y'z'$. Движение состоит в изменении расстояния $OO' = r_0$ и вращения $Oxyz$ с угловой скоростью ω



¹¹Получается, что момент инерции для вращательного движения служит аналогом массы (при использовании обобщённой координаты φ).

Дифференцируя $R = r_0 + r$, получаем абсолютную скорость

$$v_a = \dot{R} = \dot{r}_0 + \dot{r}.$$

Дифференцирование $r = x \cdot i + y \cdot j + z \cdot k$ даёт

$$\dot{r} = (\dot{x} \cdot i + \dot{y} \cdot j + \dot{z} \cdot k) + \left(x \cdot \frac{di}{dt} + y \cdot \frac{dj}{dt} + z \cdot \frac{dk}{dt} \right),$$

где $v_0 = \dot{x} \cdot i + \dot{y} \cdot j + \dot{z} \cdot k$ — *относительная скорость*. Что касается вторых скобок, то единичные орты i, j, k вращаются, не меняя длины, поэтому

$$\frac{di}{dt} = \omega \times i, \quad \frac{dj}{dt} = \omega \times j, \quad \frac{dk}{dt} = \omega \times k.$$

Следовательно,

$$x \cdot \frac{di}{dt} + y \cdot \frac{dj}{dt} + z \cdot \frac{dk}{dt} = \omega \times r,$$

что в сумме с \dot{r}_0 называют *переносной скоростью*: $v_e = \dot{r}_0 + \omega \times r$.

Таким образом,

$$\boxed{\dot{r} = v_0 + \omega \times r} \quad (8.10)$$

и $v_a = v_0 + v_e$, т. е.

$$v_a = v_0 + \dot{r}_0 + \omega \times r.$$

Абсолютное ускорение

$$w_a = \dot{v}_a = \dot{v}_0 + \dot{r}_0 + \dot{\omega} \times r + \omega \times \dot{r},$$

где

$$\dot{v}_0 = (\ddot{x} \cdot i + \ddot{y} \cdot j + \ddot{z} \cdot k) + \left(\dot{x} \cdot \frac{di}{dt} + \dot{y} \cdot \frac{dj}{dt} + \dot{z} \cdot \frac{dk}{dt} \right) = w_0 + \omega \times v_0,$$

а составляющая $w_0 = \ddot{x} \cdot i + \ddot{y} \cdot j + \ddot{z} \cdot k$ представляет собой *относительное ускорение*.

Учитывая (8.10), получаем

$$w_a = w_0 + w_e + w_k.$$

Здесь

$$w_k = 2\omega \times v_0$$

ускорение Кориолиса, а

$$w_e = r_0 + \omega \times (\omega \times r) + \dot{\omega} \times r$$

переносное ускорение, составляющая которого $\omega \times (\omega \times r)$ является центростремительным ускорением¹².

Разумеется, все это с налёту не освоишь. А неторопливо, вперемешку с поисками невесты и вращением головой по ходу буравчика — занимает семестр, а то и два. Не говоря о том, что сила Кориолиса пускает жизнь в другую колею.

8.5 Дивергенция

Пусть в \mathbb{R}^3 задана вектор-функция (векторное поле)

$$a(r) = a(x, y, z) = \{a_x, a_y, a_z\}.$$

Можно считать, например, что речь идет о течении жидкости, и $a(r)$ — скорость течения в точке r .

Далее рассматриваются *поверхностные интегралы* $\int_S a \, dS$, представляющие собой *поток вектора* $a(r)$ (жидкости) *через поверхность* S . Интеграл определяется следующим образом.

Ориентированная поверхность S разбивается на малые участки ΔS_i , причём ΔS_i считается *вектором*, длина которого численно равна площади участка, а направление совпадает с направлением нормали в произвольно выбранной точке $x \in \Delta S_i$. В модели с текущей жидкостью скалярное произведение $a \cdot \Delta S_i$ равно как раз количеству жидкости, проходящему через площадку ΔS_i в единицу времени. Предел $\sum a \cdot \Delta S_i$ при $\Delta S_i \rightarrow 0$

¹² Двойное векторное произведение $a \times (b \times c)$ в общем случае вычисляется по правилу «абец» равно бец минус цаб»:

$$a \times (b \times c) = b(a \cdot c) - c(a \cdot b).$$

называют поверхностным интегралом $\int_S \mathbf{a} dS$, который записывают также в форме

$$\int_S \mathbf{a} dS = \int_S (a_x dydz + a_y dx dz + a_z dx dy),$$

где дифференциалы

$$dydz = dS_x, \quad dx dz = dS_y, \quad dx dy = dS_z$$

представляют собой проекцию площадки (вектора) dS на соответствующие координатные плоскости.

Из физических соображений ясно, чего здесь можно ожидать. Если течение жидкости не имеет ни источников, ни стоков, то интеграл $\int_S \mathbf{a} \cdot dS$ по любой замкнутой поверхности S будет нулевым, ибо сколько втекло внутрь S , столько обязано вытечь¹³. Обычно, чтобы подчеркнуть, что интегрирование идет по замкнутой поверхности, используют запись $\oint_S \mathbf{a} dS$.

При наличии источников/стоков было бы естественно полагаться на *плотность* их распределения, которую обозначают $\operatorname{div} \mathbf{a}$, либо $\nabla \mathbf{a}$, и называют *дивергенцией*, или *расхождением* вектора \mathbf{a} . В этом случае в пределах объема V , ограниченного поверхностью S , возникает или поглощается жидкость в количестве $\int_V \operatorname{div} \mathbf{a} dV$. При этом очевидно¹⁴

$$\int_S \mathbf{a} dS = \int_V \operatorname{div} \mathbf{a} dV,$$

что называют *теоремой (формулой) Гаусса — Остроградского*, которая даёт возможность преобразовывать поверхностный интеграл в объемный, и наоборот.

¹³ Имеется в виду несжимаемая жидкость.

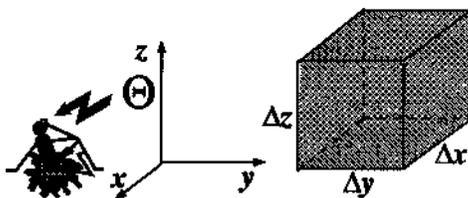
¹⁴ Сколько возникло в V , столько должно вытечь через S .

Из сказанного ясно, что дивергенция \mathbf{a} в точке r есть предел

$$\operatorname{div} \mathbf{a} = \lim_{V \rightarrow 0} \frac{\oint_S \mathbf{a} \cdot d\mathbf{S}}{V},$$

плотность потока вектора \mathbf{a} через замкнутую поверхность S , которая ограничивает бесконечно малый объём V , окружающий точку r .

Необходимо, конечно, убедиться, что определение не «стреляет в молоко». Предел обязан существовать — иначе разговор пустой. Обоснование возможно посредством вычисления самого предела. Для простоты, мы ограничимся случаем, когда тело объёма V представляет собой кубик со сторонами Δx , Δy , Δz , параллельными координатным осям.



Разность потоков вектора \mathbf{a} через грани параллельные плоскости Oyz , очевидно, равна

$$\left[\frac{\partial a_x}{\partial x} \Delta x + o(\Delta x) \right] \Delta y \Delta z.$$

Аналогично по другим граням, — что после суммирования даёт

$$\left[\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \right] \Delta V + o(\Delta V),$$

где $\Delta V = \Delta x \cdot \Delta y \cdot \Delta z$. Переход к пределу при $\Delta V \rightarrow 0$ с учётом определения дивергенция теперь приводит к формуле¹⁵

$$\operatorname{div} \mathbf{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}.$$

Что касается теоремы Гаусса — Остроградского, то в рамках содержательной интерпретации она очевидна и не требует доказательства. На

¹⁵ Если кто-то убеждён в существовании предела, то проведённые выкладки показывают, чему этот предел равен. Ибо, если предел существует, то его можно искать на любой подпоследовательности. В противном случае придётся убеждаться, что значение предела не зависит от вида ΔV , что больше подходит для индивидуального размышления.

приме у формалиста, правда, этот номер не проходит. Обоснование может опираться на стандартную процедуру. По любому $\varepsilon > 0$ объём V разбивается на столь малые элементы V_i , что

$$\left| \frac{1}{V_i} \oint_{S_i} \mathbf{a} \cdot d\mathbf{S} - \operatorname{div} \mathbf{a} \right| < \varepsilon.$$

Суммирование $\oint_{S_i} \mathbf{a} \cdot d\mathbf{S}$ даёт $\oint_S \mathbf{a} \cdot d\mathbf{S}$, поскольку потоки через внутренние перегородки S_i взаимно сокращаются. В итоге

$$\left| \oint_S \mathbf{a} \cdot d\mathbf{S} - \int_V \operatorname{div} \mathbf{a} dV \right| < \varepsilon \sum V_i = \varepsilon V,$$

что с учётом произвольности ε обеспечивает нужный результат.

- Если поверхность S ограничивает объём V , то в случае $\mathbf{a}(\mathbf{r}) = \mathbf{r}$:

$$\oint_S \mathbf{r} \cdot d\mathbf{S} = 3V.$$

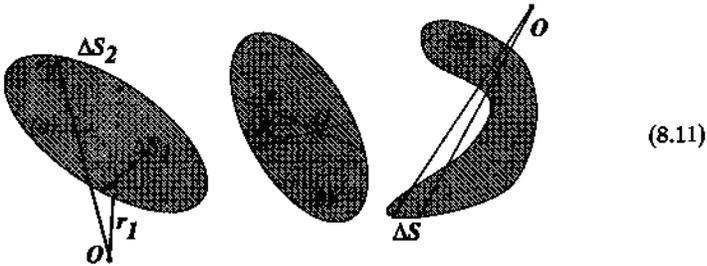


Рис. (8.11) представляет несколько возможных вариантов. В ситуации слева

$$\mathbf{r}_2 \cdot \Delta \mathbf{S}_2 - \mathbf{r}_1 \cdot \Delta \mathbf{S}_1 + o(\dots) = 3V,$$

что после суммирования и перехода к пределу даёт «то, что нужно». Остальные варианты рассматриваются аналогично.

• **Уравнение теплопроводности.** Вывод уравнения в данном контексте сам по себе не представляет особого интереса. Речь о другом — об адекватности и предпочтительности векторного языка.

Пусть $T(\mathbf{r}, t)$ обозначает температуру тела в точке \mathbf{r} в момент времени t , c — теплоемкость, $\rho(\mathbf{r})$ — плотность.

Допустим, элемент объёма dV массой $\rho(r)dV$ за время dt нагревается на $\frac{\partial T}{\partial t} dt$ градусов. Для этого необходимо количество тепла

$$dQ = c\rho dV \frac{\partial T}{\partial t} dt.$$

Интегрирование по произвольному объёму V даёт количество тепла

$$Q = \int_V c\rho dV \frac{\partial T}{\partial t} dt,$$

идущее на нагревание (или охлаждения — тогда знак минус) этого объёма.

Если теперь через $q(r, t)$ обозначить удельный поток тепла в точке r в момент t , тогда то же самое количество тепла Q можно посчитать другим способом:

$$Q = - \oint_S q \cdot dS dt = - \int_V \operatorname{div} q dV dt.$$

В результате бухгалтерия даёт

$$\int_V \left(c\rho \frac{\partial T}{\partial t} + \operatorname{div} q \right) dV = 0,$$

что, в силу произвольности V , приводит к

$$c\rho \frac{\partial T}{\partial t} + \operatorname{div} q = 0.$$

При естественном предположении $q = -k \operatorname{grad} T$ получается стандартное уравнение теплопроводности

$$\boxed{c\rho \frac{\partial T}{\partial t} - \operatorname{div} (k \operatorname{grad} T) = 0}.$$

8.6 Оператор Гамильтона

Оператором Гамильтона называют дифференциальный оператор «набла»

$$\nabla = i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z},$$

представляющий собой довольно удобный инструмент. При записи дифференциальных операций с ним можно манипулировать

как с вектором. Градиент φ , например, получается «умножением» вектора ∇ на скаляр φ :

$$\nabla\varphi = \mathbf{i} \frac{\partial\varphi}{\partial x} + \mathbf{j} \frac{\partial\varphi}{\partial y} + \mathbf{k} \frac{\partial\varphi}{\partial z}.$$

Дивергенция оказывается равной «скалярному произведению» векторов ∇ и \mathbf{a} :

$$\nabla \cdot \mathbf{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}.$$

Удобства на этом не заканчиваются. Оператор ∇ позволяет компактно записывать и другие дифференциальные операции — например, $\operatorname{rot} \mathbf{a} = \nabla \times \mathbf{a}$ (см. след. раздел), — что наталкивает на мысль о поиске определённого единства дифференциальных операций векторного анализа, см. [5].

Отметим ещё, что под скалярным произведением $\nabla \cdot \nabla$ подразумевают дифференциальный оператор Лапласа,

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

представляющий собой последовательное взятие от скалярной функции φ градиента, а потом — дивергенции:

$$\operatorname{div} \operatorname{grad} \varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2}.$$

8.7 Циркуляция

Криволинейный интеграл $\int_L P(x, y) dx$ вдоль кривой L на плос-

кости Oxy , где $P(x, y)$ — скалярная функция, заданная на L , определяется следующим образом. Кривая L , ведущая из точки $A = \{x_0, y_0\}$ в точку $B = \{x_N, y_N\}$, подразделяется точками $\mathbf{r}_i = \{x_i, y_i\}$ на малые участки, и предел суммы

$$\sum_i P(x_i, y_i) \Delta x_i, \quad \text{где } \boxed{\Delta x_i = x_i - x_{i-1}} \quad (8.12)$$

при бесконечном измельчении разбиения полагается равным интегралу $\int_L P(x, y) dx$.

Замечание. В (8.12) вместо $P(x_i, y_i)$ логичнее было бы написать $P(\zeta_i, \eta_i)$, где $\{\zeta_i, \eta_i\}$ — произвольная точка на i -м участке кривой. Но в естественных предположениях на результирующий предел это не влияет.

Важно обратить внимание на другое. Вместо (8.12) более ожидаемо, что ли, было бы рассматривать суммы

$$\sum_i P(x_i, y_i) |\Delta r_i|, \quad \text{где} \quad \boxed{\Delta r_i = r_i - r_{i-1}} \quad (8.13)$$

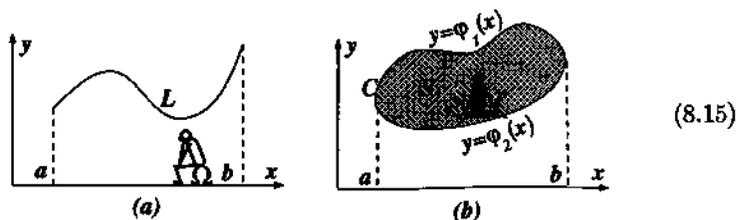
с последующим переходом к пределу. Такие *криволинейные интегралы* тоже имеют юридический статус (их называют интегралами *первого рода*, тогда как определённые выше — второго). Но интегралы на базе (8.12) нужнее.

Интегралы на базе (8.12) возникают в более общей задаче, когда задано *векторное поле* $\mathbf{a}(\mathbf{r}) = \{P(\mathbf{r}), Q(\mathbf{r})\}$ и надо интегрировать *скалярное произведение* $\mathbf{a}(\mathbf{r}) \cdot d\mathbf{r}$, т. е.

$$\int_L \mathbf{a}(\mathbf{r}) \cdot d\mathbf{r} = \int_L P dx + Q dy. \quad (8.14)$$

Если $\mathbf{a}(\mathbf{r})$ — сила, зависящая от $\mathbf{r} \in L$, то интеграл (8.14) — это работа, совершаемая силой $\mathbf{a}(\mathbf{r})$ по перемещению точки вдоль кривой L . И эта физическая интерпретация — один из главных источников интереса к интегралам вида (8.14).

• Если $y = f(x)$ описывает кривую L , рис. (8.15)-(а), то *криволинейный интеграл* $\int_L y dx$ ничем не отличается от обыкновенного



определённого интеграла $\int_a^b f(x)dx$. Отличия возникают, когда L не описывается однозначной функцией $y = f(x)$, но и тогда сохраняются обходные пути, состоящие в поэтапном интегрировании на участках, где зависимость y от x однозначна.

При этом, если L представляет собой замкнутый контур C , соответствующий интеграл по C называют *контурным* и обозначают как $\oint_C ydx$. Интересно, что $\oint_C ydx$ равен площади S , которую ограничивает контур C , рис. (8.15)-(б). Действительно, интегрирование y по C сводится к интегрированию $\varphi_1(x)$ от a до b , что даёт площадь под графиком $y = \varphi_1(x)$, плюс интеграл $\varphi_2(x)$ от b до a , что даёт *минус-площадь* под графиком $y = \varphi_2(x)$.

Если контур C находится не в плоскости Oxy , а в пространстве $Oxyz$, то при взятии того же самого интеграла $\oint_C ydx$ — поскольку от z ничего не зависит — контур C надо спроектировать в плоскость Oxy , и далее, как в предыдущем случае, интегрировать y по проекции контура C_z . Получится

$$\oint_C ydx = S_z,$$

где S_z — площадь, ограниченная C_z . Вместо S_z может получиться также $-S_z$, если учитывается ориентация.

Для любой интегрируемой функции $f(x)$: $\oint_C f(x)dx = 0$. (?)

Рассмотрим теперь общий случай криволинейного интеграла

$$\oint_C \mathbf{a}(\mathbf{r}) \cdot d\mathbf{r} = \oint_C (a_x dx + a_y dy + a_z dz), \quad (8.16)$$

называемого *циркуляцией* поля $\mathbf{a}(\mathbf{r})$ по контуру C .

Если $\mathbf{a} = \text{град}\varphi$, то интеграл (8.16) по любому замкнутому контуру равен нулю. Действительно,

$$d\varphi = \nabla\varphi \cdot d\mathbf{r},$$

поэтому вдоль любой кривой L , соединяющей точки a и b ,

$$\int_L \nabla \varphi \cdot d\mathbf{r} = \varphi(b) - \varphi(a),$$

что в случае замкнутого контура ($a = b$) даёт нуль.

Аналогом теоремы Гаусса – Остроградского при вычислении интеграла (8.16) служит теорема Стокса:

$$\boxed{\oint_C \mathbf{a}(\mathbf{r}) \cdot d\mathbf{r} = \int_S \text{rot } \mathbf{a} \cdot d\mathbf{S}}, \quad (8.17)$$

где

$$\boxed{\text{rot } \mathbf{a} = \nabla \times \mathbf{a}},$$

т. е.

$$\nabla \times \mathbf{a} = \mathbf{i} \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right) + \mathbf{j} \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x} \right) + \mathbf{k} \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right).$$

В плоском случае, когда $\mathbf{a}(\mathbf{r})$ не зависит от z , $\mathbf{a}(\mathbf{r}) = \{P(r), Q(r), 0\}$, формула Стокса (8.17) переходит в формулу Грина

$$\boxed{\oint_C P dx + Q dy = \int_S \left\{ \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right\} dx dy}. \quad (8.18)$$

Ротор называют также *вихрем*. Он характеризует плотность завихрения векторного поля. В случае *потенциального поля* (поля градиента)

$$\text{rot grad } \varphi \equiv 0,$$

что следует из представления $\nabla \times \nabla \varphi$ (либо из теоремы Стокса).

Вычисление ротора поля скоростей (8.10) вращающегося тела,

$$\mathbf{v} = \mathbf{v}_o + \boldsymbol{\omega} \times \mathbf{r},$$

приводит к показательному результату:

$$\boxed{\text{rot } \mathbf{v} = 2\boldsymbol{\omega}},$$

который свидетельствует о связи ротора с вращением.

Вместо формулы $\nabla \times \mathbf{a}$ для определения ротора можно использовать различные предельные переходы. Например, проекция ротора на направление \mathbf{h} :

$$\operatorname{rot}_{\mathbf{h}} \mathbf{a} = \lim_{\substack{C \\ S \rightarrow 0}} \frac{\oint_C \mathbf{a} \cdot d\mathbf{r}}{S}, \quad (8.19)$$

где контур C , ограничивающий площадь S , стягивается в точку, при этом вектор площади S стремится по величине к нулю, а по направлению к \mathbf{h} .

При таком определении ротора *теорема Стокса* становится очевидной, а связь с предыдущим определением ротора легко устанавливается примерно так.

Если C стягивается в нуль, то, в силу

$$a_x(\mathbf{r}) = a_x(0) + \frac{\partial a_x}{\partial x}x + \frac{\partial a_x}{\partial y}y + \frac{\partial a_x}{\partial z}z + o(\|\mathbf{r}\|),$$

где производные берутся в нуле, имеем

$$\oint_C a_x dx = a_x(0) \oint_C dx + \frac{\partial a_x}{\partial x} \oint_C x dx + \frac{\partial a_x}{\partial y} \oint_C y dx + \frac{\partial a_x}{\partial z} \oint_C z dx + \oint_C o(\|\mathbf{r}\|) dx,$$

что после учёта

$$\oint_C dx = \oint_C x dx = 0, \quad \oint_C y dx = -S_z, \quad \oint_C z dx = S_y$$

и

$$S_z = S \cos(\mathbf{h}, z), \quad S_y = S \cos(\mathbf{h}, y),$$

приводит к

$$\frac{1}{S} \oint_C a_x dx = -\frac{\partial a_x}{\partial y} \cos(\mathbf{h}, z) + \frac{\partial a_x}{\partial z} \cos(\mathbf{h}, y) + o(\cdot).$$

Аналогичные выкладки по другим координатам дают в итоге (8.19).

Глава 9

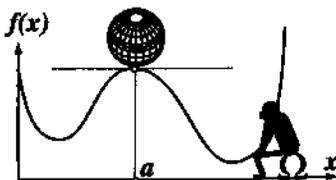
Гладкая оптимизация

*Жизнь под предлогом оптимизации
маскирует непонимание сути.*

9.1 Безусловный экстремум

Тяга к оптимизации спрятана глубоко в природе человека. Может быть потому, что придаёт многому осмысленный вид. «Угол падения равен углу отражения» — навеивает скуку, а «свет распространяется по кратчайшему пути» — вдохновляет. Но позыв больше идеологический. Укрепляет дух и направляет мысль. Конкретные задачи не любит. Рассказывать тем не менее приходится — как будто всё наоборот. Иногда и жить так приходится.

Протуберанцы раскопок экстремальных задач уходят за горизонт, однако начинать естественно с привычного и простого. С *теоремы Ферма* 3.5.1 в первую очередь. У функции $y = f(x)$ в *локальном максимуме* в точке $x = a$ производная $f'(a) = 0$.



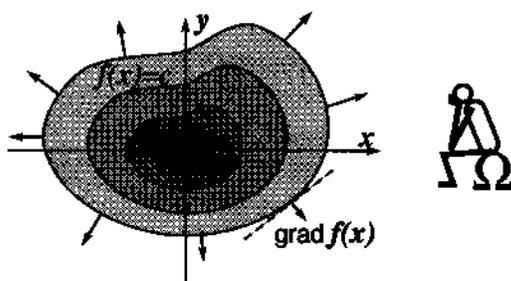
Следовательно в случае векторного аргумента, если $f(x)$ имеет локальный максимум в нуле, то функция одной переменной

$$f(0, \dots, 0, x_i, 0, \dots, 0)$$

тоже имеет локальный максимум в нуле. Поэтому все частные производные $\partial f(0)/\partial x_i$ обязаны быть нулевыми, что означает обращение в нуль градиента.

9.1.1 Теорема.¹ Если $f(x)$ в точке $x = x^*$ дифференцируема и принимает локально экстремальное (максимальное или минимальное) значение, то градиент $\nabla f(x^*) = 0$.

У функции $y = f(x_1, x_2)$ линии постоянного уровня в районе локального минимума $x = 0$ выглядят примерно так



Замкнутые контуры стягиваются к центру, в котором градиент «вынужден» быть направлен сразу во все стороны, что возможно лишь при его обнулении.

Заметим, функции $f(x)$ предполагаются далее *гладкими*. Здесь под гладкостью подразумевается непрерывная дифференцируемость, каковая существенна в данном контексте. Функция $f(x) = 2x^2 + x^2 \sin(1/x^2)$, например, имеет в нуле строгий минимум и равную нулю производную, но (1) в сколь угодно малой окрестности нуля эта производная принимает как положительные, так и отрицательные сколь угодно большие значения.

¹Теорема даёт лишь необходимое условие оптимума, но его чаще всего достаточно для решения задачи, ибо условие $\nabla f(x^*) = 0$ для экстремальной точки — обязательно. Поэтому искомым критическое значение x^* попадает в число решений системы уравнений $\nabla f(x) = 0$. Последняя обычно имеет либо одно, либо несколько решений, среди которых уже трудно заблудиться.

Точки, в которых градиент обращается в нуль, называют *критическими*. Их многообразие, разумеется, не исчерпывается экстремумами.

9.2 Достаточные условия

В одномерном случае *разложение Тейлора*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots$$

показывает, что в критической точке a , $f'(a) = 0$, достигается минимум, если $f''(a) > 0$, и — максимум, если $f''(a) < 0$. В общей ситуации тоже можно судить о характере критической точки по квадратичной части разложения Тейлора,

$$f(\mathbf{x}^* + \Delta \mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*) \cdot \Delta \mathbf{x} + \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} \Delta x_i \Delta x_j + o(\|\Delta \mathbf{x}\|^2).$$

Если второй дифференциал, называемый *гессеианом*, положительно определен, т. е.

$$\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} \Delta x_i \Delta x_j > 0 \text{ при } \Delta \mathbf{x} \neq 0,$$

то в критической точке \mathbf{x}^* — минимум², ибо в малой окрестности \mathbf{x}^* будет $f(\mathbf{x}^* + \Delta \mathbf{x}) > f(\mathbf{x}^*)$, потому что $\nabla f(\mathbf{x}^*) = 0$.

Если *матрица Гессе*³ $\left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]$ вырождена, ситуация принципиально усложняется. В скалярном случае о характере критической точки можно судить по первому ненулевому члену ряда Тейлора. Для функций n переменных дело обстоит совершенно иным образом, и там сам чёрт ногу сломит⁴.

²В случае отрицательной определённости — максимум.

³*Матрица Гессе* — это *матрица Якоби* отображения $\nabla f(x)$.

⁴Предвидеть трудности, но не их масштаб, довольно легко. Вопрос решается в *теории катастроф* [6].

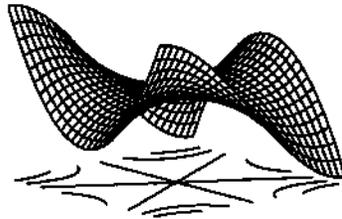
Островок порядка характеризуется невырожденными критическими точками, в которых невырождены матрицы Гессе. Классическая лемма Морса гарантирует существование в некоторой окрестности V невырожденной критической точки, т. е. локальной системы координат x_1, \dots, x_n такой, что

$$f = x_1^2 + \dots + x_l^2 - x_{l+1}^2 - \dots - x_n^2, \quad x \in V. \quad (9.1)$$

Функция вида (9.1) называется морсовским l -седлом. В случае $l = 0$ имеем максимум, при $l = n$ — минимум.

Лемма Морса даёт полную классификацию невырожденных критических точек. Для вырожденных критических точек такой классификации нет. Простой пример вырожденной критической точки — обезьянье седло, которое описывается функцией

$$f(x, y) = x^3 - 3xy^2$$



9.3 Условный экстремум

Оптимизация обычно сталкивается с необходимостью соблюдения некоторых условий. Скажем, так разделить ресурс R на части $\{x_1, \dots, x_n\}$, чтобы добиться максимального эффекта $\sum_i \varphi_i(x_i) \rightarrow \max$, но при соблюдении, разумеется, ограничения $\sum_i x_i \leq R$.

Ограничений (условий оптимизации) в задаче может быть много, но идеологически важно разобраться с простейшей ситуацией одного ограничения. Допустим, максимизация⁵

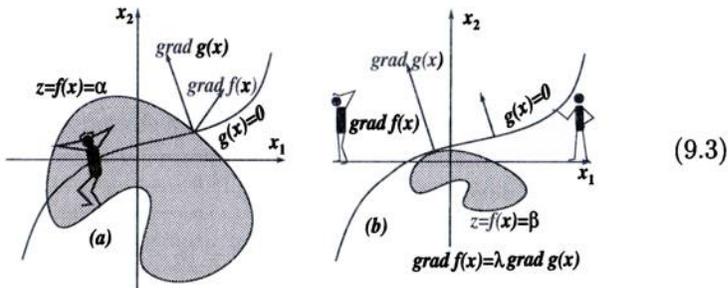
$$f(x) \rightarrow \max$$

сдерживается одним ограничением, которое задаётся уравнением

$$g(x) = 0, \quad x = \{x_1, \dots, x_n\}. \quad (9.2)$$

⁵Гладкость рассматриваемых функций везде подразумевается.

Если точка \mathbf{x} находится на поверхности $g(\mathbf{x}) = 0$, то в направлении плюс-минус градиента $\nabla g(\mathbf{x})$ двигаться нельзя — иначе точка \mathbf{x} сразу сойдет с поверхности $g(\mathbf{x}) = 0$, и условие нарушится, рис. (9.3)-(a). Можно смещаться в касательной плоскости (на бесконечно малую $\Delta \mathbf{x}$), т. е. перпендикулярно $\nabla g(\mathbf{x})$. Если градиент $\nabla f(\mathbf{x})$ не коллинеарен $\nabla g(\mathbf{x})$, т. е. не совпадает по направлению с $\pm \nabla g(\mathbf{x})$, то у $\nabla f(\mathbf{x})$ есть составляющая в касательной плоскости к $g(\mathbf{x}) = 0$, вдоль которой можно сдвинуться и увеличить тем самым значение $f(\mathbf{x})$.



Этой последней возможности «сдвинуться и увеличить значение $f(\mathbf{x})$ » нет: либо в вырожденном случае⁶, либо когда градиенты коллинеарны,

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}) \quad \text{при некотором } \lambda \neq 0. \quad (9.4)$$

Тогда в точке \mathbf{x} происходит касание поверхностей $g(\mathbf{x}) = 0$ и $f(\mathbf{x}) = \beta$, рис. (9.3)-(b). При этом достигается (возможно) *условный минимум*. Условие (9.4), таким образом, является *необходимым*, — но только в невырожденной ситуации.

Особые случаи — стандартная головная боль. Соломинка обходится дороже стога сена, и средств на другие надобности перестает хватать. Поэтому обычно не жалеют сил на достижение единообразия, позволяющего обходиться без «реверансов».

⁶ Вырожденный случай охватывает «малоинтересные» ситуации: либо $\nabla g(\mathbf{x}) = 0$, либо в точке \mathbf{x} на поверхности $g(\mathbf{x}) = 0$ достигается безусловный минимум $f(\mathbf{x})$, и направление градиента $\nabla g(\mathbf{x})$ не играет роли.

В данном случае, например, избежать упоминания вырожденной ситуации можно так. Если \mathbf{x}^* — точка условного максимума, то найдутся такие числа λ_0, λ_1 , не равные нулю одновременно, что

$$\lambda_0 f(\mathbf{x}^*) + \lambda_1 g(\mathbf{x}^*) = 0.$$

Но это имеет свои минусы. Во-первых, смешивая воедино разнородные явления, формулировка теряет лицо. Во-вторых, рассуждения всё равно требуют разграничения возможностей.

Задачи, в которых можно положить $\lambda_0 = 1$, характеризуют как *регулярные*. *Регулярными* называют и соответствующие точки минимума.

Итак, правило поиска условного минимума заключается в решении системы (9.4)⁷ совместно с уравнением (9.2). На практике это оформляется немного иначе. Задача

$$f(\mathbf{x}) \rightarrow \max, \quad g(\mathbf{x}) = 0, \quad (9.5)$$

заменяется поиском экстремальной точки \mathbf{x} *лагранжиана*

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x}) \quad (9.6)$$

при надлежащем выборе λ . Знак перед λ принципиального значения не имеет (пока).

В итоге все сводится к решению системы $n + 1$ уравнений:

$$g(\mathbf{x}) = 0 \quad \text{и} \quad \frac{\partial f(\mathbf{x})}{\partial x_i} - \lambda \frac{\partial g(\mathbf{x})}{\partial x_i} = 0, \quad i = 1, \dots, n,$$

относительно $n + 1$ неизвестных x_1, \dots, x_n, λ . Коэффициент λ называют *множителем Лагранжа*.

Пример. Забегая несколько вперед⁸, решим задачу

$$\sum_{i=1}^n r_i \sqrt{x_i} \rightarrow \max, \quad \sum_{i=1}^n x_i = X, \quad (9.7)$$

⁷ Это n скалярных уравнений, поскольку градиент — вектор.

⁸ «Забегание вперед» связано с присутствием ограничений $x_i \geq 0$.

которую можно интерпретировать как оптимальное распределение ресурса X по n ячейкам с эффективностью $r_i\sqrt{x_i}$ от использования ресурса в количестве x_i .

Действуя по регламенту, получаем:

$$L(x, \lambda) = \sum_{i=1}^n r_i\sqrt{x_i} - \lambda \left(\sum_{i=1}^n x_i - X \right),$$

$$\frac{\partial L}{\partial x_i} = \frac{r_i}{2\sqrt{x_i}} - \lambda = 0, \quad \sum_{i=1}^n x_i = X, \quad i = 1, \dots, n.$$

Решение последней системы уравнений даёт оптимальное распределение:

$$x_i^* = X \frac{r_i^2}{\sum_j r_j^2}, \quad (9.8)$$

а также значение множителя Лагранжа $\lambda^* = \frac{1}{2} \sqrt{\sum_j r_j^2 / X}$, с которым пока не ясно, что делать (см. раздел 9.6).

Итоговая формула (9.8) получается в результате бескритерийного решения вышестоящей системы уравнений — и нуждается по этой причине в обосновании того, что найден действительно максимум. В первую очередь требует проверки возможность более эффективного решения на границе неотрицательного ортанта. Такая возможность исключена, так как производная в нуле любой функции $r_i\sqrt{x_i}$ бесконечна и, следовательно, у каждой ячейки в нуле бесконечна скорость роста эффективности. Поэтому какое-то количество ресурса всем надо давать — иначе заведомо можно получить выигрыш, передав малую часть ресурса той ячейке, которая ничего не получила.

Таким образом, решение должно лежать внутри рассматриваемой области и обязано улавливаться методом множителей Лагранжа. Поскольку «метод» других стационарных решений не обнаружил — на этом можно ставить точку.

Рассмотрим теперь ситуацию двух ограничений

$$g_1(x) = 0 \quad \text{и} \quad g_2(x) = 0.$$

Как уже отмечалось в начале раздела, чтобы оставаться на поверхности $g(x) = 0$, можно двигаться только перпендикулярно градиенту $\nabla g(x)$. В данном случае надо оставаться на обеих поверхностях одновременно, чего можно достичь, смещаясь перпендикулярно как $\nabla g_1(x)$, так и $\nabla g_2(x)$. Другими словами, двигаться можно лишь перпендикулярно плоскости, проходящей через оба градиента. Если градиент целевой функции $\nabla f(x)$ лежит вне этой плоскости, то у него есть составляющая в разрешенном

направлении, и значение $f(\mathbf{x})$ можно улучшать. В противном случае,

$$\nabla f(\mathbf{x}) = \lambda_1 \nabla g_1(\mathbf{x}) + \lambda_2 \nabla g_2(\mathbf{x}) \quad \text{при некоторых } \lambda_1, \lambda_2 \neq 0,$$

разрешённых направлений нет, и соответствующая точка \mathbf{x} — кандидат на решение. В итоге вместо (9.6) лагранжианом будет

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda_1 g_1(\mathbf{x}) - \lambda_2 g_2(\mathbf{x}).$$

Разумеется, как и выше, нужны оговорки насчёт вырожденных случаев, к каковым теперь добавляется возможность касания в точке \mathbf{x} поверхностей $g_1(\mathbf{x}) = 0$ и $g_2(\mathbf{x}) = 0$, при которой градиенты $\nabla g_1(\mathbf{x})$ и $\nabla g_2(\mathbf{x})$ оказываются линейно зависимы.

9.4 Общий случай

Решая общую задачу

$$f(\mathbf{x}) \rightarrow \min, \quad g_1(\mathbf{x}) = 0, \quad \dots, \quad g_m(\mathbf{x}) = 0, \quad (9.9)$$

оптимизируют *лагранжиан*

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_{j=1}^m \lambda_j g_j(\mathbf{x}) \quad (9.10)$$

по \mathbf{x} при надлежащем выборе $\lambda_1, \dots, \lambda_m$.

В результате поиск *условного экстремума* (9.9) сводится к решению системы уравнений

$$\frac{\partial f}{\partial x_k} = \sum_{j=1}^m \lambda_j \frac{\partial g_j}{\partial x_k}, \quad g_1(\mathbf{x}) = 0, \quad \dots, \quad g_m(\mathbf{x}) = 0,$$

$$k = 1, \dots, m.$$

Картину портят *вырожденные случаи*, на рассмотрение которых уходит много энергии с очень низким КПД. Малая эффективность такой работы определяется двумя обстоятельствами. Во-первых, многие исключения тривиальны и скучны. Во-вторых, *вырожденные случаи* на фоне типичных ситуаций — всего лишь «точка в континууме».

Последнее обстоятельство создаёт впечатление, что уйти от рассмотрения исключительных вариантов совсем просто. Достаточно «слегка пошевелить» условия. Подобный трюк изредка экономит усилия, но иногда анализ вырожденных случаев неизбежен. Например, при изучении зависимости поведения системы от параметра. Меняющийся параметр натывается на аномалию, которую «не обойдешь», и это заставляет бросить всё и заняться патологией.

9.4.1 Теорема. Если точка \mathbf{x}^* есть решение задачи (9.9), то найдутся такие числа $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*$, не все равные нулю, что

$$\lambda_0^* \nabla f(\mathbf{x}^*) = \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*). \quad (9.11)$$

Условие (9.11) вместе с $g_1(\mathbf{x}^*) = 0, \dots, g_m(\mathbf{x}^*) = 0$ образуют систему $n + m$ уравнений относительно $n + m + 1$ неизвестных

$$x_1, \dots, x_n; \lambda_0^*, \lambda_1^*, \dots, \lambda_m^*.$$

Однородность по λ (произвольное ненулевое λ_k можно положить равным единице) позволяет систему считать замкнутой (число неизвестных и уравнений совпадает).

Таким образом, замена лагранжиана (9.10) функцией

$$L_0(\mathbf{x}, \lambda) = \lambda_0 f(\mathbf{x}) - \sum_{j=1}^m \lambda_j g_j(\mathbf{x})$$

позволяет в формулировке теоремы 9.4.1 обойтись без оговорок. На практике обычно полагают $\lambda_0 = 1$, и если в процессе решения что-тостораживает, начинают разбираться в чем дело.

Любые предположения, позволяющие исключить в (9.11) случай $\lambda_0 = 0$ — например линейная независимость градиентов $\nabla g_1(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)$ — называют *условиями регулярности*.

9.5 Нелинейное программирование

Задача *нелинейного программирования*,

$$\begin{aligned} f(\mathbf{x}) &\rightarrow \min, \\ g_i(\mathbf{x}) &= 0, \quad i = 1, \dots, m, \\ h_j(\mathbf{x}) &\leq 0, \quad j = 1, \dots, r, \end{aligned} \quad (9.12)$$

отличается от (9.9) наличием дополнительных ограничений в виде неравенств. Ограничения в (9.12) удобно записывать в векторном виде: $g(\mathbf{x}) = 0$, $h(\mathbf{x}) \leq 0$. Элементы \mathbf{x} , удовлетворяющие ограничениям задачи, называются *допустимыми решениями*.

Присутствие в ограничениях неравенств типа $\mathbf{x} \geq 0$, — вещь характерная. Более того, с помощью дополнительных переменных задача (9.12) приводится к виду

$$f(\mathbf{x}) \rightarrow \min, \quad g(\mathbf{x}) = 0, \quad h(\mathbf{x}) + z = 0, \quad z \geq 0, \quad (9.13)$$

где из неравенств фигурирует только неотрицательность координат.

Если решение \mathbf{x}^* задачи (9.12) равносильно (9.13), удовлетворяет неравенству $h_j(\mathbf{x}^*) < 0$, то ограничение $h_j(\mathbf{x}^*) \leq 0$ роли не играет, и его можно не принимать в расчёт. Такого сорта *ограничения* называют *пассивными*, в отличие от *активных*, характеризующих выходом решения на границу: $h_j(\mathbf{x}^*) = 0$.

9.5.1 Теорема Каруши — Джона. Если точка \mathbf{x}^* есть решение задачи (9.12), то найдутся такие числа λ_0^* , λ_1^* , ..., λ_m^* и μ_1^* , ..., μ_r^* , не все равные нулю, что⁹

$$\lambda_0^* \nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^r \mu_j^* \nabla h_j(\mathbf{x}^*) = 0, \quad (9.14)$$

⁹За кадром имеется в виду лагранжиан

$$L_0(\mathbf{x}, \lambda, \mu) = \lambda_0 f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^r \mu_j h_j(\mathbf{x}).$$

Обратите внимание на знаки.

причем $\lambda_0^* \geq 0$, все $\mu_j^* \geq 0$ и

$$\mu_j^* h_j(x^*) = 0, \quad j = 1, \dots, r. \quad (9.15)$$

Равенства (9.15) называются *условиями дополняющей нежёсткости*, и проистекают из отмеченного выше соображения о несущественности пассивных ограничений. Если $h_j(x^*) \neq 0$, то соответствующий множитель Лагранжа $\mu_j^* = 0$, в силу (9.15), — что равносильно исключению неравенства $h_j(x) \leq 0$ из списка ограничений. Гарантия положительности коэффициентов μ_j^* у активных ограничений очевидна в связи со знаком неравенств $h_j(x) \leq 0$.

Уравнения (9.14), (9.15) вместе с m ограничениями $g_i(x) = 0$ и учётом однородности по множителям Лагранжа — приводят в соответствие число уравнений и неизвестных. Поскольку теорема 9.5.1 даёт лишь необходимые условия, которым обязано удовлетворять решение задачи (9.12), ни о чем больше можно не упоминать. Как правило, эти условия позволяют получить желаемое. Но возможны также осложнения, связанные с обстоятельствами, которые принято характеризовать как вырожденные.

Например, • функция $f(x)$ имеет безусловный минимум на границе; • градиенты функций, определяющих ограничения, оказываются линейно зависимы; • нет допустимых решений, т. е. отсутствуют точки x , удовлетворяющие ограничениям. В последнем случае система уравнений (9.14), (9.15) плюс $g(x) = 0$ может иметь решение, но оно не удовлетворяет ограничениям, и задача (9.12) оказывается неразрешимой. Теорема 9.5.1 формально справедлива, поскольку исходное предположение просто не выполняется.

В первых двух случаях теорема также остается справедливой, но эффективность её падает, потому что соответствующая система уравнений песет на себе следы не вполне нормальной ситуации. Решений оказывается слишком много, и без понимания того, что происходит за кадром, выбрать нужное — не так просто, особенно, если речь идет о «слепом» машинном поиске.

Многие руководства, пытаясь помочь неопытным пользователям, раскладывают возникающие варианты по полочкам в форме многочисленных утверждений. Эффект получается отрицательный. Ибо возникает иллюзия наличия рецептов на все случаи жизни, но этими рецептами при непонимании предмета обязательно пользуются невпопад. Сказывается «слепое чтение» теорем, при котором из поля зрения ускользают многие предположения из-за непонимания их роли, не говоря об общей потере ориентации, когда все мешается воедино. Поэтому при столкновении с негипичными ситуациями требуется понимание «подводных течений».

9.6 Интерпретация множителей Лагранжа

Изучая *множители Лагранжа*, важно не упустить главное. В результате решения экстремальных задач определяются «ненужные» коэффициенты $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*$ и μ_1^*, \dots, μ_r^* . Но они, понятно, говорят о чем-то, что может представлять дополнительный интерес, а то и первоочередной.

Секрет частично раскрыт *условиями дополняющей нежёсткости*: если $\mu_j^* = 0$, то соответствующее ограничение $h_j(\mathbf{x}) \leq 0$ пассивно. С целью дальнейшего «расследования» возьмём задачу с одним ограничением

$$f(\mathbf{x}) \rightarrow \min, \quad g(\mathbf{x}) = b, \quad (9.16)$$

отличающуюся от (9.5) наличием числового параметра b .

Пусть $\{\mathbf{x}^*(b), \lambda^*(b)\}$ обозначает экстремальную точку соответствующего (9.16) лагранжиана $L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda[g(\mathbf{x}) - b]$. Дифференцируя $L(b) = L[\mathbf{x}^*(b), \lambda^*(b)]$, получаем

$$\frac{dL}{db} = \sum_i \frac{\partial f}{\partial x_i} \frac{dx_i^*}{db} - \frac{d\lambda^*}{db} [g(\mathbf{x}^*(b)) - b] - \lambda^*(b) \left[\sum_i \frac{\partial g}{\partial x_i} \frac{dx_i^*}{db} - 1 \right]. \quad (9.17)$$

В (9.17) оба выражения в квадратных скобках равны нулю по причине $g(\mathbf{x}^*(b)) \equiv b$. Поэтому $\frac{dL}{db} = \frac{df}{db}$, где $\frac{df}{db} = \sum_i \frac{\partial f}{\partial x_i} \frac{dx_i^*}{db}$.

С другой стороны, группируя в (9.17) слагаемые иначе:

$$\frac{dL}{db} = \sum_i \left[\frac{\partial f}{\partial x_i} - \lambda^*(b) \frac{\partial g}{\partial x_i} \right] \frac{dx_i^*}{db} + \lambda^*(b),$$

получаем $\frac{dL}{db} = \lambda^*(b)$. В итоге:

$$\boxed{\frac{df}{db} = \lambda^*(b)}. \quad (9.18)$$

Таким образом, $\lambda^*(b)$ показывает скорость роста оптимального значения целевой функции f по параметру b . Иными словами, представляет собой *цену ограничения*. Ещё говорят о *чувствительности* оптимума по отношению к ограничению.

В задаче распределения ресурса (9.7), например, $\lambda^* = \frac{1}{2} \sqrt{\sum_j r_j^2 / X}$ даёт зависимость суммарного эффекта от использования ресурса X системой в целом. Это значение λ^* при использовании в качестве реальной цены на ресурс имеет то преимущество, что индивидуальная прибыль

$$r_i \sqrt{x_i} - \lambda^* x_i$$

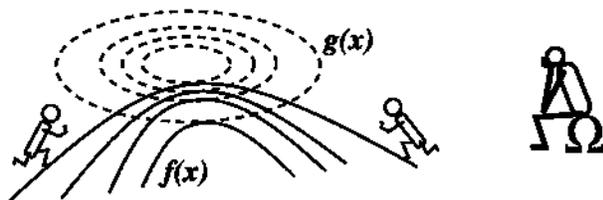
каждого элемента достигает максимума при тех значениях ресурса x_i^* , которые оптимальны по суммарному критерию, т. е. для всего «коллектива».

В случае m ограничений $g_j(x) - b_j = 0$ чуть более громоздкие выкладки приводят к обобщению (9.18):

$$\boxed{\frac{\partial f}{\partial b_j} = \lambda_j^*} \quad (9.19)$$

9.7 Двойственные задачи

Вернёмся к задаче (9.16) с одним ограничением. На рисунке



изображены линии постоянного уровня функций $f(x)$ и $g(x)$. Поскольку, как уже отмечалось, градиенты в экстремальной точке должны быть коллинеарны, решение задачи x^* будет точкой касания линий постоянного уровня

$$f(x) = c \quad \text{и} \quad g(x) = b.$$

При таком угле зрения понятно, что нет разницы, идет ли речь о задаче « $f(x) \rightarrow \min, g(x) = b$ » или о *двойственной задаче* « $g(x) \rightarrow \max, f(x) = c$ ».

¹⁰При условии его оптимального распределения внутри системы.

Смена минимума на максимум при переходе к двойственной задаче на данном этапе рассмотрения проблемы в определённой степени условна. Проще говорить об экстремуме в той и другой задаче. Есть две функции, f и g , и получается, что нет разницы, какую из них оптимизировать при поддержании значения второй — на заданном уровне. Уточнение сорта оптимизации (минимум или максимум) возникает при некотором углублении в задачу, когда в поле зрения попадает «правильное» построение лагранжиана (за счет «правильного» выбора знака, см. замечание к теореме 9.5.1).

В случае более общей задачи

$$f(\mathbf{x}) \rightarrow \min, \quad g_1(\mathbf{x}) = 0, \quad \dots, \quad g_m(\mathbf{x}) = 0$$

возникает аналогичная ситуация. С точки зрения поиска стационарных точек нет особой разницы, какая функция в наборе

$$\{f(\mathbf{x}), g_1(\mathbf{x}), \dots, g_m(\mathbf{x})\}$$

оптимизируется, а какие — описывают ограничения.

Глава 10

Аналитические функции

*В Поднебесной нет задач,
есть только паруса фантазий.*

10.1 О роли комплексных чисел

Выход в комплексную плоскость вскрывает внутренние механизмы многих теорий. Скажем, теория матриц без комплексных чисел (КЧ) — вообще не теория. Начинается-то все с естественных вопросов¹. Как найти систему координат, в которой матрица A как линейный оператор стала бы диагональной, т. е. как найти преобразование T , обеспечивающее трансформацию A в наиболее простую форму:



$$A' = T^{-1}AT = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{bmatrix}. \quad (10.1)$$

◀ В случае (10.1) A' на векторы нового базиса $\{f'_1, \dots, f'_n\}$ действует предельно просто: $A'f'_j = \lambda_j f'_j$, что после умножения слева на T , справа на T^{-1} , даёт $TA'T^{-1} \cdot Tf'_jT^{-1} = \lambda_j Tf'_jT^{-1}$, т. е. $Af_j = \lambda_j f_j$, где $f_j = Tf'_jT^{-1}$. ▶

Таким образом, для определения чисел λ_j и векторов f_j надо решить уравнение $(A - \lambda I)x = 0$, каковое может иметь ненулевые ре-

¹Разговор пока для создания подходящего настроения. Поэтому технические детали играют роль декораций, на которые не надо обращать пристального внимания.

шения $x = f_j$ лишь в случае, когда матрица $A - \lambda I$ вырождена, т. е. $\det(A - \lambda I) = 0$, что называют *характеристическим уравнением* матрицы A , его решения λ_j — *собственными значениями*, а соответствующие ненулевые решения f_j — *собственными векторами* матрицы A .

Далее линейный анализ подходит к *фундаментальному поворотному моменту*. Если уравнение $\det(A - \lambda I) = 0$ имеет n различных действительных корней λ_j , то все слишком хорошо. Каждому λ_j соответствует свой собственный вектор f_j . Эти векторы составляют базис новой системы, в которой матрица $T^{-1}AT$ принимает диагональный вид, а матрица перехода T образуется столбцами f_j .

Угроза потери этих благ из-за того, что полином n -й степени редко имеет n действительных корней, подталкивает к выходу в комплексную плоскость \mathbb{C} . Там характеристическое уравнение всегда имеет n корней. Но если параметр λ комплексный, то к рассмотрению надо допускать комплексные векторы, иначе нет смысла говорить о ненулевых решениях x уравнения $(A - \lambda I)x = 0$.

Затем возникает необходимость в сдаче следующей позиции. Столбцы f_j (теперь комплексные) образуют в совокупности матрицу перехода T , которая приводит A к диагональной форме $T^{-1}AT$, причём у $T^{-1}AT$ на диагонали могут стоять в том числе комплексные λ_j . Получается, что надо допускать к рассмотрению комплексные матрицы. Поэтому проще всего — и целесообразнее — с самого начала все считать комплексным. Конечно, не все так просто, как поначалу кажется, — подробности в [2, т. 3] — но комплексный вариант не только спасает теорию матриц от блуждания впотьмах, но и открывает многие ракурсы, которые иначе остались бы незамеченными. Спектр матрицы в \mathbb{C} , например, гарантирует в \mathbb{R}^n асимптотическую устойчивость системы $\dot{x} = Ax + b$, если $\operatorname{Re} \lambda_j < 0$ для всех j . Система $\dot{x} = Ax + b$ изначально «действительна», но её поведение оказывается следствием «комплексных пружин». И такое положение вещей характерно для действительного анализа вообще.

Вычисление k -й производной функции $y = \frac{1}{x^2+1}$ весьма обременительно. В то же время, представление функции в виде

$$y = \frac{i}{2} \left(\frac{1}{x+i} - \frac{1}{x-i} \right)$$

легко даёт

$$y^{(k)} = \frac{i}{2} (-1)^k k! \left[\frac{1}{(x+i)^{k+1}} - \frac{1}{(x-i)^{k+1}} \right],$$

что легко переводится в действительную форму.

КЧ резко упрощают теорию колебаний, электричество. Законы волновой оптики странным образом оказываются согласованы с правилами умножения комплексных чисел. Потом вдруг обнаруживается, что изучение линейных дифференциальных уравнений без КЧ вообще невозможно. Постепенно становится ясно, что это не просто удобная вещь, а *фантастически удобная*. Дальше — больше. Получается, без комплексных чисел никуда нет ходу. Даже в теорию вероятностей. Даже в теорию чисел.

Тогда размышления устремляются в другое русло. В чем причина? Почему добавление к колебанию фиктивной мнимой части всё упрощает? Почему, чтобы найти радиус сходимости ряда

$$a_0 + a_1x + a_2x^2 + \dots$$

для вещественного x надо изучать поведение ряда в комплексной плоскости (причём иначе — невозможно)? Откуда у вспомогательного инструмента такая колоссальная мощь и проникающая способность? Вот в чем источник подозрения, что какая-то космическая тайна стоит за кадром.

Сегодня на все эти вопросы есть простые и ясные ответы, которые начали созревать и оформляться относительно недавно — в девятнадцатом веке. До этого в суждениях о комплексных числах даже таких великих математиков, как Лейбниц и Эйлер, преобладали эмоции. И очень трудно сказать, чего им не хватало, ибо сегодняшние ответы в общих чертах они бы могли сформулировать и тогда. Разница едва ли не идентичных формулировок была бы в наполненности несколько иным содержанием, в других акцентах, в другом понимании тех же самых слов. Последние двести лет дали массу интересных примеров, породили

серию других «нереальных» инструментов. Кроме того, вымерли многие несостоятельные идеи. Все это вместе взятое наполнило старые формы новым содержанием.

Первый слой понимания сути комплексных чисел внешне достаточно прост. Введение операций сложения и умножения для натурального ряда инициирует цепную реакцию. Обратные операции — вычитание и деление — влекут за собой необходимость появления отрицательных чисел, потом дробей. Аналогичная ситуация возникает в связи с извлечением корней, что приводит к появлению мнимой единицы, а потом — с учётом сложения и умножения — комплексных чисел.

Иными словами, на целых положительных числах вводятся арифметические операции, после чего натуральный ряд расширяется до максимального игрового поля. Такое поле дают комплексные числа. Процесс расширения однозначен, и дальше пути, кстати, нет².

На следующем витке спирали возникает вопрос о широком распространении комплексных чисел. Почему они неожиданно всплывают то там, то здесь? Ответ прост. Потому, что вся математика стоит на арифметических операциях, для которых комплексные числа — неизбежный финал расширения натурального ряда. Поэтому игра идет на едином общем поле.

10.2 Дифференцируемость

Функцией комплексного переменного $z = x + iy$ можно было бы считать любую функцию вида

$$f(z) = u(x, y) + iv(x, y),$$

но тогда, с точностью до декораций, все свелось бы к изучению отображений на плоскости, $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$\begin{bmatrix} x \\ y \end{bmatrix} \xrightarrow{f} \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix}.$$



²Например, кватернионы (трехмерные числа) рассматривают иногда как обобщение КЧ, но это вопрос словоупотребления. Умножение кватернионов некоммутативно ($ab \neq ba$), и потеря тех или иных свойств «арифметических операций» неизбежна при любом расширении поля комплексных чисел.

Аналитические функции, изучаемые в ТФКП, выделяет особое понятие *дифференцируемости*, каковая определяется как предельность приращения функции в виде

$$\Delta f(z) = f(z + \Delta z) - f(z) = A \Delta z + o(|\Delta z|), \quad (10.2)$$

где A — комплексное число, зависящее от точки z , которое называют *производной* $f(z)$ в точке z , и обозначают $f'(z)$, либо $\frac{df}{dz}$.

Может показаться, что ничего особенного в таком определении нет, но это только на первый взгляд. Если бы речь шла, например, просто об отображении плоскости, то определение дифференцируемости f совпадало бы по записи с (10.2), но A было бы матрицей,

$$A = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}. \quad (10.3)$$

Мы же хотим, чтобы A было комплексным числом³. Это очень сильно ограничивает класс рассматриваемых функций.

Определение производной записывают также в виде

$$f'(z) = \lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}, \quad (10.4)$$

но в (10.4) требуется дополнительно независимость предела от пути стремления Δz к нулю, что приводит к *необходимости* выполнения для $f = u + iv$ *условий Коши — Римана*:

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}} \quad \begin{array}{c} \text{⚙} \\ \Omega \end{array} \quad (10.5)$$

Необходимость (10.5) устанавливается совсем просто.

³ Другими словами, чтобы действие матрицы A на Δz было эквивалентно умножению на комплексное число (растяжение плюс поворот). Такие матрицы, как известно, имеют вид $k \cdot \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$, откуда следует необходимость условия (10.5).

◀ В случае $\Delta z = \Delta x + i0$ и $\Delta x \rightarrow 0$ вычисление предела (10.4) даёт

$$\frac{u(x + \Delta x, y) - u(x, y)}{\Delta x} + i \frac{v(x + \Delta x, y) - v(x, y)}{\Delta x} \rightarrow \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = f'(z),$$

в случае $\Delta z = 0 + i \Delta y$ и $\Delta y \rightarrow 0$ тот же предел (10.4) оказывается равным

$$\frac{u(x, y + \Delta y) - u(x, y)}{i \Delta y} + i \frac{v(x, y + \Delta y) - v(x, y)}{i \Delta y} \rightarrow \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y} = f'(z),$$

Оба варианта обязаны совпадать, что и приводит к (10.5). ▶

(!) Независимость предела (10.4) от характера стремления Δz к нулю — позволяет вычислять производную $f'(z)$ любым удобным способом:

$$f'(z) = \frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \quad \text{либо} \quad f'(z) = -i \frac{\partial f}{\partial y} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}.$$

С учётом (10.5) это даёт дополнительно:

$$f'(z) = \frac{\partial u}{\partial x} - i \frac{\partial u}{\partial y}, \quad f'(z) = \frac{\partial v}{\partial y} + i \frac{\partial v}{\partial x}. \quad (10.6)$$

10.2.1 Необходимые условия Коши — Римана (10.5) в случае непрерывности частных производных функций u, v в некоторой области D достаточны для дифференцируемости f в D , т. е. в каждой точке $z \in D$.

◀ В силу непрерывности частных производных⁴:

$$\Delta u = u(x + \Delta x, y + \Delta y) - u(x, y) = \frac{\partial u}{\partial x} \Delta x + \frac{\partial u}{\partial y} \Delta y + o(|\Delta z|).$$

Аналогично для Δv . С учётом условий Коши — Римана это приводит к

$$\Delta f(z) = \Delta u + i \Delta v = \left(\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) (\Delta x + i \Delta y) + o(|\Delta z|),$$

что совпадает с определением (10.2). ▶

(!) Требование непрерывности частных производных в п. 10.2.1 — необязательно. Достаточно дифференцируемости u и v как функций действительных переменных в каждой точке $(x, y) \in D$, для чего одного лишь существования частных производных не хватает [2, т. 1].

⁴Ниже Δ обозначает (пока) приращение, а не оператор Лалласа.

10.2.2 *Функции, дифференцируемые в некоторой области D , называются аналитическими в D .*

Тавтология терминов может показаться странной. Зачем для дифференцируемой функции выдумывать ещё одно имя? Но тут есть исключительное оправдание. Оказывается, *аналитические функции обладают совершенно неожиданным свойством*. Дело заключается в следующем. Обыкновенный анализ приучает к мысли, что функция может быть дифференцируема либо один раз, либо два, либо n раз. Как будет показано далее, имеет место удивительный факт:

Аналитическая функция, дифференцируемая по определению всего один раз, оказывается бесконечно дифференцируемой.

10.3 Примеры

Несмотря на довольно жёсткие рамки дифференцируемости, аналитические функции весьма многочисленны и разнообразны.

1. Функция $f(z) = z^n$ при любом натуральном n дифференцируема, её производная $f'(z) = nz^{n-1}$ легко вычисляется:

$$\frac{(z + \Delta z)^n - z^n}{\Delta z} = nz^{n-1} + (\dots) \cdot \Delta z \rightarrow nz^{n-1}$$

при $\Delta z \rightarrow 0$. Независимость от характера стремления Δz к нулю очевидна.

2. Сумма и произведение аналитических в D функций $f_1(z)$ и $f_2(z)$ — аналитичны в D . Функция $f_1(z)/f_2(z)$ аналитична в D всюду, где $f_2(z) \neq 0$. Композиция аналитических функций, $f(g(z))$, аналитична.

◀ Обоснование слово в слово повторяет обычные доказательства аналогичных утверждений (по новоду дифференцируемости) в случае действительного аргумента. ▶

3. Функция $f(z)$, представляемая сходящимся в некотором круге ($|z| < R$) рядом

$$f(z) = \sum_{n=0}^{\infty} a_n z^n,$$

аналитична при $|z| < R$. Пример дается пока авансом⁵.

⁵ До рассмотрения вопросов сходимости комплексных рядов.

4. Положим

$$e^z = e^x(\cos y + i \sin y) = e^{x+iy}.$$

На действительной оси функция совпадает с e^x . Условия Коши — Римана легко проверяются. (?) Это может служить пока формальным определением комплексной экспоненты. Более глубокие мотивы такого определения вскроются позже.

5. Функция $f(z) = |z|^2 = x^2 + y^2$ не аналитична. Условия Коши — Римана выполняются только при $z = 0$.

Из перечисленных примеров может возникнуть впечатление, что почти все функции аналитичны. На самом деле, конечно, «почти все — не аналитичны». Но практика сталкивает нас в большинстве случаев с аналитическими функциями: синусы, логарифмы, дробно-рациональные выражения и т. п.

С оазисом аналитических функций соседствуют и обескураживающие примеры. Не аналитична функция $\bar{z} = x - iy$. Сопряжение нарушает аналитичность и в общем случае. «Зеркальный образ»

$$\overline{f(z)} = u(x, y) - iv(x, y)$$

и сама функция $f(z) = u(x, y) + iv(x, y)$ не могут обе удовлетворять условиям Коши — Римана, если не говорить о тривиальном исключении $f(z) \equiv \text{const}$.

Нигде не дифференцируема, скажем, функция $f(z) = 2x + iy$, что указывает на широкое поле контрпримеров, если в таковых есть потребность. Однако теория, как и Вселенная, растёт в благоприятных направлениях.

10.4 Простейшие свойства

Остановимся на некоторых свойствах аналитических функций, где понимание достигается малой кровью.

1. Чем меньше окрестность точки z , тем точнее выполняется соотношение $\Delta f = f'(z) \Delta z$, которое означает, что все Δz поворачиваются на один и тот же угол и растягиваются в одно и то же число раз. Растяжение и поворот определяются множителем $f'(z)$. Поэтому *малые фигуры переводятся функцией f в подобные* (углы сохраняются, размеры пропорционально изменяются — разумеется, с точностью до «о малых»). Это служит основанием называть такие преобразования *конформными*.

Разумеется, если кривые пересекаются в плоскости z под углом α , то их образы при отображении f пересекаются в точности под углом α без всяких «о малых».

2. Тот факт, что аналитическое преобразование $w = f(z)$ локально представляет собой растяжения и повороты, влечет за собой обязательное сохранение ориентации. Иначе говоря, образы всех маленьких вихрей оказываются закручены в ту же сторону, в силу чего невозможны *отражения*. Именно поэтому дифференцируемость $f(z)$ исключает дифференцируемость $f(\bar{z})$. При этом двукратное сопряжение $f(\bar{z})$ сохраняет аналитичность.

3. Задание частных производных определяет функцию с точностью до константы. Поэтому задание одной из функций $u(x, y)$, $v(x, y)$, в силу условий Коши — Римана, с точностью до константы определяет вторую.

4. В силу условий Коши — Римана

$$\nabla u \cdot \nabla v = u_x v_x + u_y v_y = -u_x u_y + u_y u_x \equiv 0, \quad (10.7)$$

т. е. *градиенты функций $u(x, y)$ и $v(x, y)$ ортогональны в любой точке z* , а значит, ортогональны линии постоянного уровня⁶ функций u и v .

5. Дифференцируя условия Коши — Римана и исключая одну из функций u или v , легко убедиться, что обе удовлетворяют *уравнению Лапласа*:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0, \quad \text{☝}$$

и являются таким образом *гармоническими функциями*, хорошо известными благодаря другим источникам происхождения.

Кто знаком с дифференциальными уравнениями в частных производных, может сразу сделать неожиданный (для аналитических функций) вывод. Задание поведения гармонической функции на границе области полностью определяет её значения внутри области. Для дифференциальных уравнений такая связь естественна и привычна, но для ТФКП — несколько шокирующая.

⁶ Семейства кривых $u(x, y) = C$, $v(x, y) = C$.

Все, чем занимается теория аналитических функций, — как выясняется — это изучение решений одного дифференциального уравнения $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$. Специалисты по диффурам удивляются противоположным образом: «Неужели теория одного диффура — это целая ТФКП?»⁷

Гидродинамическая модель. Вектор скорости

$$\mathbf{v} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix}$$


двумерного безвихревого потока — без источников и стоков — обязан удовлетворять двум уравнениям

$$\operatorname{div} \mathbf{v} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \quad \operatorname{rot} \mathbf{v} = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0. \quad (10.8)$$

Чтобы (10.8) совпадало с условиями Коши — Римана, достаточно взять за исходную функцию

$$\overline{f(z)} = u(x, y) - i v(x, y),$$

каковая «видится в зеркале» при наблюдении за $f(z)$. В данном случае аналитической должна быть $\overline{f(z)}$, но не $f(z)$, — если желательно говорить об условиях Коши — Римана.

Линии тока в модели совпадают с траекториями дифференциального уравнения $\frac{dz}{dt} = \overline{f(z)}$. Взаимосвязь моделей (ТФКП и гидродинамики) двусторонняя. Интуитивные представления о ламинарных течениях позволяют кое-что заключать об аналитических функциях, и «наоборот».

10.5 Контурные интегралы

Комплексный интеграл $\int_L f(z) dz$ вдоль кривой L можно рассматривать как предел суммы

$$\sum_k f(z_k) \Delta z_k$$

⁷Здесь есть некоторая натяжка в ту и другую сторону.

при стандартном измельчении разбиения либо, пользуясь тем, что криволинейные интегралы действительных функций определены как

$$\int_L f(z)dz = \int_L udx - vdy + i \int_L udy + vdx,$$

что получается после перемножения $f = u + iv$ и $dz = dx + idy$.

10.5.1 Теорема Коши. Пусть функция $f(z)$ аналитична в односвязной области D . Тогда интеграл от $f(z)$ по любому замкнутому контуру C , лежащему в D , равен нулю⁸.

◀ Доказательство моментально даёт применение к

$$\oint_C f(z)dz = \oint_C udx - vdy + i \oint_C udy + vdx,$$

формулы Грина (8.18) с последующим учётом того, что u и v удовлетворяют условиям Коши — Римана. ▶

В условиях теоремы 10.5.1, если кривые $L_1 \in D$ и $L_2 \in D$ имеют одни и те же концевые точки a, b , то

$$\int_{L_1} f(z)dz = \int_{L_2} f(z)dz, \quad (10.9)$$



что сразу получается интегрированием по замкнутому контуру (10.9). Но для этого (1) *внутри контура, образованного кривыми L_1 и L_2 , не должно быть особых точек, в которых $f(z)$ недифференцируема*⁹. Другими словами, интеграл $\int_L f(z)dz$ не

⁸ Теорема Коши остаётся справедливой и в более общем случае. Область D может быть неодносвязной, но тогда контур C должен быть таков, чтобы его можно было стянуть в точку по области D . Либо интегрировать надо по совокупности контуров, ограничивающих D , см. [2, т. 9].

⁹ Иначе нарушается предположение теоремы 10.5.1 об аналитичности функции $f(z)$ в области D .

зависит от пути интегрирования, а определяется лишь начальной и конечной точками. Поэтому вместо $\int_L f(z)dz$ можно писать

$$\int_a^b f(z)dz. \text{ При этом очевидно свойство аддитивности: } \int_a^b = \int_a^c + \int_c^b.$$

Положим

$$F(z) = \int_a^z f(w)dw,$$

где подразумевается путь интегрирования по произвольной кривой, лежащей в D и ведущей из a в z .

Легко показать, что функция $F(z)$ аналитична, причём

$$\boxed{F'(z) = f(z)}.$$

Действительно,

$$F(z + \Delta z) - F(z) = \int_z^{z+\Delta z} f(w)dw,$$

откуда

$$\frac{F(z + \Delta z) - F(z)}{\Delta z} - f(z) = \frac{1}{\Delta z} \int_z^{z+\Delta z} [f(w) - f(z)]dw.$$

Интеграл $\int_z^{z+\Delta z} f(w)dw$ стремится к нулю при $\Delta z \rightarrow 0$, что и обеспечивает $F'(z) = f(z)$. Множество первообразных $\Phi(z)$, $\Phi'(z) = f(z)$, называют неопределённым интегралом функции $f(z)$. Очевидно, первообразные отличаются друг от друга на константу. Поэтому для любой из них

$$\int_a^b f(z)dz = \Phi(b) - \Phi(a).$$

Примеры.

• Функция z^n при любом $n = 1, 2, \dots$ аналитична везде. Поэтому интеграл от z^n по любому замкнутому контуру равен нулю.

• Пусть C — окружность радиуса R с центром в нуле. Тогда

$$\oint_C \frac{dz}{z} = 2\pi i. \quad \text{(10.10)}$$

Это важный факт. Интеграл легко вычисляется в полярных координатах,

$$z = R(\cos \varphi - i \sin \varphi) = Re^{-i\varphi},$$

а так как $R = \text{const}$, то

$$dz = R(-\sin \varphi - i \cos \varphi)d\varphi = iRe^{-i\varphi}d\varphi.$$

Поэтому

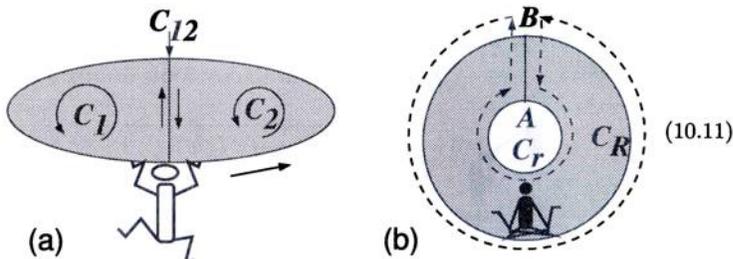
$$\oint_C \frac{dz}{z} = i \int_0^{2\pi} d\varphi = 2\pi i.$$

Интеграл оказывается не равным нулю, но противоречия с теоремой 10.5.1 нет, поскольку функция $1/z$ не аналитична в нуле, и контур C нельзя стянуть в точку в $\mathbb{R}^2 \setminus \{0\}$. Именно последнее обстоятельство является критичным для справедливости теоремы Коши. Область D на самом деле может быть какой угодно. Если контур C можно стянуть в точку в области D , то интеграл от $f(z)$ по C будет равен нулю.

Для дальнейшего важно ещё отметить *аддитивность* контурного интеграла в следующем смысле. Если контуры C_1 и C_2 имеют общий участок C_{12} , после выбрасывания которого образуется контур C , причём C_{12} проходится в противоположных направлениях при интегрировании по C_1 и C_2 , то

$$\oint_{C_1} f(z)dz + \oint_{C_2} f(z)dz = \oint_C f(z)dz.$$

Это не что иное как свойство слагаемых противоположных знаков взаимно сокращаться. Ситуация на рис. (10.11)-(а) абсолютно прозрачна.



На рис. (10.11)-(b) изображен более сложный пример. Функция $f(z)$ аналитична в кольце, ограниченном малой C_r и большой окружностью C_R . Интеграл по замкнутому контуру $C_{r,R}$, проходящему C_R против часовой стрелки, потом по BA , затем C_r по часовой стрелке, затем возврат в исходную точку по AB , — равен нулю. Поэтому

$$\oint_{C_R} f(z) dz = \oint_{C_r} f(z) dz, \quad (10.12)$$

что является следствием

$$\oint_{C_R} f(z) dz + \oint_{C_{r,R}} f(z) dz = \oint_{C_r} f(z) dz.$$

10.6 Интеграл Коши

Пусть функция $f(z)$ аналитична внутри контура C и на самом контуре C . Тогда $\frac{f(w)}{w-z}$ как функция w аналитична на C и внутри C за исключением точки $w = z$. Поэтому

$$\oint_{C_r} \frac{f(w)}{w-z} dw = \oint_{C_r} \frac{f(w)}{w-z} dw, \quad (10.13)$$

где C_r обозначает окружность радиуса r с центром в z . Равенство (10.13) можно получить так же, как и (10.12), поскольку в обосновании (10.12) никакой роли не играло, что C_R окружность.

Поскольку $f(w) = f(z) + o(w-z)$, то

$$\oint_{C_r} \frac{f(w)}{w-z} dw = f(z) \oint_{C_r} \frac{dw}{w-z} + \oint_{C_r} \frac{o(w-z)}{w-z} dw. \quad (10.14)$$

При $r \rightarrow 0$ второй интеграл справа в (10.14) обнуляется, а первое слагаемое, в силу (10.10), оказывается равным $f(z) \cdot 2\pi i$. В результате



$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(w)}{w-z} dw, \quad (10.15)$$

что называют *интегральной формулой Коши*.

В случае, когда C является окружностью C_r , представляет интерес перейти в (10.15) к интегрированию в полярных координатах:

$$w = z + re^{i\varphi}, \quad dw = ire^{i\varphi} d\varphi.$$

В результате получается формула

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} f(z + re^{i\varphi}) d\varphi,$$

или, равносильно,

$$f(z) = \frac{1}{2\pi r} \oint_{C_r} f(w) ds, \quad (10.16)$$

где ds — дифференциал дуги окружности C_r .

Выходит, что $f(z)$ в точности равно среднему значению f на окружности C_r с центром в z .

10.6.1 Принцип максимума модуля¹⁰. Функция $f(z)$, не равная константе, и аналитическая на замыкании области Ω , — может достигать максимума своего модуля $|f(z)|$ только на границе области Ω .

◀ Доказательство легко вытекает из факта (10.16), ибо при равенстве среднего значения f на окружности C_r значению f в центре C_r — максимум модуля $|f|$ в центре — невозможен. Поэтому во внутренней точке области невозможен локальный максимум модуля, тем более — глобальный. ▶

Бесконечная дифференцируемость. Уникальность ситуации, которую улавливает интегральная формула Коши, заключается в том, что (10.15) можно продифференцировать по параметру z и получить



$$f'(z) = \frac{1}{2\pi i} \oint_C \frac{f(w)}{(w-z)^2} dw, \quad (10.17)$$

потом снова продифференцировать (10.17) по z , получить формулу для второй производной,

$$f''(z) = \frac{2}{2\pi i} \oint_C \frac{f(w)}{(w-z)^3} dw,$$

¹⁰ Принцип максимума имеет многочисленные следствия, широко применяемые в качестве подручных инструментов.

и так до бесконечности,

$$f^{(n)}(z) = \frac{n!}{2\pi i} \oint_C \frac{f(w)}{(w-z)^{n+1}} dw. \quad (10.18)$$

Перенос дифференцирования под знак интеграла здесь легко обосновывается, поскольку дифференцирование под интегралом саму функцию f вообще не затрагивает.

В результате выявляется уникальная вещь. *Аналитическая функция оказывается бесконечно дифференцируемой сама по себе, без каких бы то ни было дополнительных требований.* Интеграл Коши даёт наиболее простой ключ к пониманию этого факта. Функция $f(z)$ в силу (10.15) определяется интегрированием самой себя, а поскольку интеграл обладает улучшающими свойствами¹¹, то $f(z)$ обязана быть лучше, чем того можно ожидать. Выход из положения один — быть «бесконечно хорошей».

10.7 Регулярность

10.7.1 Теорема. *Если функция $f(z)$ аналитична в некоторой окрестности W точки a , то она представима рядом Тейлора,*

$$f(z) = f(a) + f'(z)(z-a) + \dots + \frac{f^{(n)}(a)}{n!}(z-a)^n + \dots, \quad (10.19)$$

в любом шаре $|z-a| < r$, принадлежащем этой окрестности¹².

◀ Для доказательства возьмём максимальное r , при котором контур C_r принадлежит W , и рассмотрим сумму бесконечно убывающей геометрической прогрессии

$$\frac{1}{w-z} = \frac{1}{w-a} + \frac{z-a}{(w-a)^2} + \dots + \frac{(z-a)^n}{(w-a)^{n+1}} + \dots, \quad (10.20)$$

где $w \in C_r$, а z лежит внутри C_r , и потому знаменатель прогрессии $(z-a)/(w-a)$ по модулю меньше единицы. В результате ряд (10.20) на C_r сходится равномерно. Умножая (10.20) на $f(w)$ и интегрируя по C_r , с учётом (10.18) получаем (10.19). ▶

¹¹Из непрерывной функции делает дифференцируемую.

¹²Ситуация кардинально отличается от той, которая имела место в случае функций действительной переменной. Там ряд Тейлора мог сходиться вообще к другой функции (см. раздел 3.7).

10.7.2 Определение. *Функция, представимая рядом (10.19), сходящимся в некоторой окрестности точки a , называется регулярной, или голоморфной в точке a .*

Казалось бы, ещё один перехлест в нагромождении терминов. Дифференцируемость в окрестности z , аналитичность и регулярность — просто совпадают. Но это в ТФКП. В анализе на действительной прямой это совершенно различные понятия. Кроме того, в самой ТФКП за каждым понятием стоят свои собственные особые признаки, что удобно для расстановки акцентов. Упоминание регулярности функции, например, подчеркивает её разложимость в ряд Тейлора¹³.

10.7.3 Теорема. *Любой степенной ряд с ненулевым радиусом сходимости является рядом Тейлора своей суммы.*

Разумеется, это переформулировка теоремы 10.7.1. Однако новая форма часто привносит дополнительные краски, и «процесс» начинает развиваться другим путём. Благодаря теореме 10.7.3 на степенные ряды можно смотреть как на ряды Тейлора. При этом фундаментальную роль играет следующий простой результат.

10.7.4 Теорема Абеля. *Если ряд $\sum_{n=0}^{\infty} c_n z^n$ сходится при каком-то z , равным по модулю r , то при условии $|z| < r$ он сходится абсолютно, а при $|z| \leq r - \varepsilon$ ($\varepsilon > 0$) — ещё и равномерно.*

◀ Из сходимости $c_0 + c_1 z + c_2 z^2 + \dots$ вытекает $c_n z^n \rightarrow 0$, а значит и ограниченность: $|c_n r^n| < M$. Поэтому для $|z| < r$ будет

$$\sum_{n=0}^{\infty} |c_n z^n| = \sum_{n=0}^{\infty} |c_n r^n| \cdot \left| \frac{z}{r} \right|^n < M \sum_{n=0}^{\infty} \left| \frac{z}{r} \right|^n < \infty,$$

что обеспечивает абсолютную сходимость в круге $|z| < r$. Аналогичная выкладка даёт требуемую равномерную сходимость. ▶

Использование в ТФКП рядов Тейлора в качестве стандартного инструмента удобно, но не обязательно. Использование других рядов

¹³ Тем не менее, на прилагательные аналитическая, регулярная и голоморфная (функция) надо смотреть как на синонимы. Это главное. Чтобы фразы типа «аналитическая функция голоморфная в такой-то области» — не ставили в тупик. Особенно поначалу. Потом уже синонимам можно придавать оттенки.

имеет свои плюсы и минусы. Широкое распространение при изучении аналитических функций получили, например, *ряды Лорана*

$$f(z) = \sum_{n=-\infty}^{\infty} c_n(z-a)^n = \sum_{n=0}^{\infty} c_n(z-a)^n + \sum_{n=1}^{\infty} \frac{c_{-n}}{(z-a)^n}. \quad (10.21)$$

Первое слагаемое справа в (10.21) представляет собой обыкновенный степенной ряд, сходящийся в некотором круге радиуса R_1 . Второе слагаемое после замены $w = \frac{1}{1-z}$ тоже превращается в обыкновенный степенной ряд, который сходится в плоскости w в круге радиуса R_2 , а в плоскости z — вне круга радиуса $1/R_2$.

Поэтому область сходимости (10.21) является кольцо, рис. справа, что и служит причиной популярности рядов Лорана. Чтобы накрыть изолированную особую точку требуется бесконечно много «кругов Тэйлора» либо одно «кольцо Лорана».



Несложные выкладки показывают, что коэффициенты в (10.21) могут быть определены при любом n по единой формуле

$$c_n = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z-a)^{n+1}} dz,$$

C — произвольный замкнутый контур, лежащий в кольце

$$\frac{1}{R_2} < |z-a| < R_1$$

и охватывающий точку a .

Коэффициент

$$c_{-1} = \frac{1}{2\pi i} \oint_C f(z) dz \quad (10.22)$$

в разложении (10.21) называется *вычетом* и обозначается $\text{res}(f, a)$.

10.7.5 Теорема. Пусть функция $f(z)$ однозначна и аналитична на замкнутом контуре C и внутри него за исключением конечного числа точек a_1, \dots, a_k . Тогда

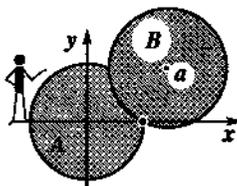
$$\oint_C f(z) dz = 2\pi i [\text{res}(f, a_1) + \dots + \text{res}(f, a_k)]. \quad (10.23)$$

◀ Аддитивность контурного интеграла сводит \oint_C к сумме контурных интегралов по маленьким окружностям с центрами в особых точках. Это и даёт (10.23). ▶ Инструмент (10.23) широко применяется для решения многочисленных задач.

10.8 Аналитическое продолжение

Благодаря теореме 10.7.1 степенные ряды в ТФКП играют роль универсального инструмента. Но при этом есть свои трудности. Например, функция $\frac{1}{1-z}$ представляется степенным рядом

$$\frac{1}{1-z} = 1 + z + z^2 + \dots,$$



(10.24)

который сходится при $|z| < 1$ (круг A). В то же время функция $\frac{1}{1-z}$ аналитична везде за исключением $z = 1$. Поэтому её можно разложить в ряд в любой точке $a \neq 1$. Получится другой ряд Тейлора, который будет сходиться в другом круге B (10.24). В каком соотношении находятся эти ряды и сама функция $\frac{1}{1-z}$? За кадром здесь просматривается очевидная идея.

Если на области A (речь идет уже не обязательно о кругах) задана аналитическая функция $f_1(z)$, на B — $f_2(z)$, и на пересечении $A \cap B$ значения функций совпадают, то было бы хорошо $f_2(z)$ считать продолжением $f_1(z)$ на более широкую область (равно как и $f_1(z)$ — продолжением $f_2(z)$). Чтобы такой трюк был законным, не хватает «малости» — гарантий единственности продолжения.

10.8.1 Теорема единственности. Пусть функции $f(z)$ и $g(z)$ аналитичны в области D , и их значения совпадают в различных точках $a_n \in D$, причём $a_n \rightarrow a \in D$ при $n \rightarrow \infty$. Тогда $f(z) \equiv g(z)$ в D .

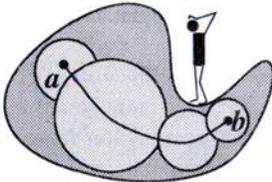
◀ *Доказательство.* Очевидно, аналитическая в D функция $h(z) = f(z) - g(z)$ в точках a_n обращается в нуль, а потому и $h(a) = 0$, в силу $a_n \rightarrow a$ и непрерывности h . Поэтому в разложении

$$h(z) = c_0 + c_1(z - a) + c_2(z - a)^2 + \dots \quad (10.25)$$

коэффициент $c_0 = 0$. Следовательно, $h(z) = (z - a)h_1(z)$. Но тогда

$$h(a_n) = 0 \quad \rightarrow \quad h_1(a_n) = 0.$$

Снова, по непрерывности, $h_1(a) = 0$, что означает $c_1 = 0$. И так далее. Все коэффициенты в разложении h оказываются нулевыми. Поэтому $h(z) \equiv 0$ в круге сходимости ряда (10.25), т. е. $f(z) \equiv g(z)$.



Окончательный вывод о тождественности нулю $h(z)$ в D вытекает из возможности попасть в любую точку $w \in D$ по цепочке пересекающихся кругов, рис. слева, в каждом из которых $h(z)$ разлагается в свой ряд Тейлора, и повторяется предыдущее рассуждение (для некоторой новой последовательности a_n из предыдущего круга, где уже $h(z) \equiv 0$). ►

Из доказательства легко усмотреть, что аналитичность¹⁴ влечет за собой *изолированность нулей*. Но главный вывод заключается, если можно так сказать, в «голографическом устройстве» аналитических функций. Целое отражается в любом маленьком фрагменте. *Задание функции на кусочке дуги определяет все остальное.*

Поэтому функции e^x , $\sin x$, $\ln x$ на действительной прямой автоматически порождают соответствующие функции в комплексной плоскости. Вопрос лишь в том, как их найти. Когда выдвигаются какие-нибудь дополнительные требования, проблема может оказаться сложной, но принципиально она всегда разрешима. Разложение в степенные ряды позволяет гарантированно расширить область определения аналитической функции до максимальных размеров. Иногда это реализуется совсем просто. Скажем, в ситуациях

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!},$$

в которых ряды имеют неограниченный радиус сходимости, достаточно поменять действительное x на комплексное z — и аналитические функции e^z и $\sin z$ готовы. В случае логарифма,

$$\ln x = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(x-1)^n}{n}$$

¹⁴ Имеется в виду функция, не равная тождественно нулю.

проблема сложнее. Конечно, x здесь можно заменить на z , но это даст лишь ряд Тейлора, сходящийся в круге единичного радиуса, тогда как логарифм определен в гораздо более широком диапазоне¹⁵. Надежный выход из положения заключается в расширении области определения пошаговым переходом к другим рядам по цепочкам пересекающихся кругов. Это обязательно приведет к успеху, но только «идеологическому». Логарифм везде, где возможно, будет определен и даже вычислим, но единой формулы это не даст. Простая формула,

$$\ln z = \int_1^z \frac{dw}{w},$$

получается совсем из других соображений. Но это уже вопрос удачи, а не систематического подхода.

В сказанном важно обратить внимание на следующий аспект. *Любая аналитическая функция может быть продолжена до функции, определённой на максимально возможной области.* Для $\sin z$ — это вся плоскость, для $\frac{z}{1-z}$ — плоскость с выколотой точкой $z = 1$. Но так или иначе появляется возможность вообще не упоминать область определения, считая $f(z)$ определённой везде, где возможно. Это ещё одна удивительная особенность ТФКП. *Область определения задаётся самой функцией.*

Функции, аналитичные на всей комплексной плоскости, называются *целыми*, и с ними меньше всего хлопот. Любое их разложение в ряд Тейлора сходится при любом z . Там же, где необходимо прыгать по цепочкам кругов, — может возникать *проблема многозначности*, связанная с тем, что продолжение по различным путям даёт разные результаты¹⁶. Такая ситуация вполне типична. Самые обыкновенные функции \sqrt{z} , $\ln z$ — многозначны, и это заставляет разрабатывать инструменты для борьбы с неудобствами, см. [2, т. 9].

10.9 Теорема Руше

Сначала несколько слов об особых точках, как о засасывающих воронках. Их никто, как правило, не хочет изучать, поскольку — «не до аномалий». Но они иногда требуются не сами по себе, а как инструмент решения проблем,

¹⁵ Логарифм определен везде за исключением действительной полуоси $(-\infty, 0)$. Говорят: на комплексной плоскости с разрезом по $(-\infty, 0)$.

¹⁶ В этом случае внутри контура, образуемого разными путями, обязательно имеется особая точка.

где патология нависает виртуально, как «угроза». Кроме того, аномалия — обычно наиболее информативная часть явления в целом.

Точку $z = a$ называют *изолированной особой точкой* функции $f(z)$, если $f(z)$ аналитична и однозначна в некотором кольце

$$0 < |z - a| < r,$$

а в самой точке a — не дифференцируема. Если дифференцируемость в a восстанавливается при доопределении значения функции в этой точке, то a называют *устраняемой особой точкой*¹⁷.

Если все коэффициенты c_{-n} разложения Лорана $f(z)$ в a равны нулю при $n > k$, но $c_{-k} \neq 0$, точку a называют *полосом k -го порядка*.

О нулях и полюсах аналитической функции $f(z)$ удобно судить по *логарифмическим вычетам*, каковыми называют обычные вычеты (10.22) производной логарифма

$$[\ln f(z)]' = \frac{f'(z)}{f(z)}.$$

При этом оказывается:

$$\frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z)} dz = N - P, \quad (10.26)$$

где N и P соответственно — полное число нулей и полюсов $f(z)$ внутри контура C . Каждый нуль и полюс считается столько раз, каков его порядок.

Интеграл (10.26) имеет также другую естественную интерпретацию¹⁸:

$$\frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z)} dz = \frac{1}{2\pi} \Delta_C \arg f(z), \quad (10.27)$$

где $\Delta_C \arg f(z)$ — изменение аргумента $f(z)$ при обходе контура C точкой z .

Таким образом,



$$N - P = \frac{1}{2\pi} \Delta_C \arg f(z),$$

¹⁷Особая точка a функции $f(z)$ устранима в том случае, когда лорановское разложение $f(z)$ в окрестности a не содержит отрицательных степеней.

¹⁸Вывод формул (10.26), (10.27) можно использовать в качестве упражнения, либо см. [2, т. 9, гл. 6].

что называют *принципом аргумента*. В частности, если $f(z)$ аналитична внутри контура C , то в области, ограниченной контуром C , функция $f(z)$ имеет

$$N = \frac{1}{2\pi} \Delta_C \arg f(z) \quad \text{нулей.} \quad \text{!} \quad (10.28)$$

10.9.1 Теорема Руше. Если функции $f(z)$, $g(z)$ аналитичны в области D и непрерывны на замыкании \bar{D} , причём $|f(z)| > |g(z)|$ на контуре C , ограничивающем D , то $f(z)$ и $f(z) + g(z)$ имеют в D одинаковое число нулей.

◀ На C , в силу $|f(z)| > |g(z)| \geq 0$,

$$|f(z) + g(z)| \geq |f(z)| - |g(z)| > 0,$$

функции f и $f + g$ нулей не имеют. Далее,

$$\begin{aligned} \frac{1}{2\pi} \Delta_C \arg[f(z) + g(z)] &= \frac{1}{2\pi} \Delta_C \arg \left\{ f(z) \left[1 + \frac{g(z)}{f(z)} \right] \right\} = \\ &= \frac{1}{2\pi} \Delta_C \arg f(z) + \frac{1}{2\pi} \Delta_C \arg \left[1 + \frac{g(z)}{f(z)} \right], \end{aligned}$$

причём $\arg \left[1 + \frac{g(z)}{f(z)} \right] = 0$, потому что вектор $1 + \frac{g(z)}{f(z)}$ не может сделать оборота вокруг нуля, оставаясь в правой полуплоскости в силу $\left| \frac{g(z)}{f(z)} \right| < 1$. Поэтому

$$N_{f+g} = \frac{1}{2\pi} \Delta_C \arg[f(z) + g(z)] = \frac{1}{2\pi} \Delta_C \arg f(z) = N_f. \quad \blacktriangleright$$

10.9.2 Основная теорема алгебры. Любой многочлен $P_n(z)$ n -й степени имеет на комплексной плоскости \mathbb{C} ровно n корней.

◀ В применении к ситуации

$$f(z) = z^n, \quad g(z) = a_{n-1}z^{n-1} + \dots + a_1z + a_0,$$

теорема Руше сразу даёт необходимое заключение, потому что все требования выполняются на кругах достаточно большого радиуса. ▶

Глава 11

Функциональные ряды

*Многие неприятности
в гильбертовых пространствах
не помещаются.*

Представление функций $f : \mathbb{R} \rightarrow \mathbb{R}$ в виде бесконечных сумм,

$$f(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \dots + c_n\varphi_n(x) + \dots, \quad (11.1)$$

играет в анализе существенную роль. *Функциональные ряды* (11.1) интерпретируются обычно как запись $f(x)$ в координатной форме, c_n — n -я координата, $\varphi_n(x)$ — направляющий вектор n -й оси. Постепенно выясняется, что за этим стоит гораздо больше, чем просто внешнее сходство.

Принципиальной особенностью разложений (11.1) является возможность приближения функций $f(x)$ *частичными суммами*

$$f_n(x) = c_0\varphi_0(x) + \dots + c_n\varphi_n(x). \quad \img alt="A small cartoon drawing of a person looking through a telescope." data-bbox="638 728 704 759"/>$$

Вопрос сходимости $f_n(x) \rightarrow f(x)$ обычно подменяют вопросом стремления к нулю хвоста ряда

$$\img alt="A small cartoon drawing of a person standing." data-bbox="216 853 275 906"/> $\sum_{k=n+1}^{\infty} c_k\varphi_k(x) \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty, \quad (11.2)$$$

что влечёт за собой сходимость $f_n(x)$ к некоторой функции, но к «эф ли от икс?» — это другая задача. Дело в том, что проблемы нет, когда само разложение (11.1) служит определением $f(x)$. Если же $f(x)$ определяется как-то иначе, то необходимо ещё обосновать, что правая часть (11.1) сходится именно к $f(x)$.

11.1 Равномерная сходимость

Под равномерной сходимостью ряда



$$\sum_{n=0}^{\infty} c_n \varphi_n(x) \quad (11.3)$$

подразумевается равномерная сходимость к нулю его хвоста (11.2) на рассматриваемом отрезке $[a, b]$. Равномерной сходимостью ряда на бесконечной (либо незамкнутой) области X будем называть его равномерную сходимость на любом ограниченном замкнутом подмножестве X .

11.1.1 Теорема. Пусть все функции $\varphi_n(x)$ непрерывны на отрезке $[a, b]$, и ряд (11.3) сходится равномерно. Тогда его сумма $f(x)$ — непрерывная функция.

◀ Результат сразу следует из теоремы 5.8.3. ▶

11.1.2 Теорема. Пусть функции $\varphi_n(x)$ непрерывны на $[a, b]$, и ряд (11.3) сходится равномерно. Тогда ряд (11.3) можно почленно интегрировать, т. е. интеграл от суммы ряда $f(x)$ равен

$$\int_a^b f(x) dx = \sum_{n=0}^{\infty} c_n \int_a^b \varphi_n(x) dx$$



◀ Доказательство элементарно. Из равномерной сходимости к нулю хвоста ряда следует стремление к нулю интеграла хвоста

$$\left| \int_a^b f(x) dx - \sum_{k=0}^n c_k \int_a^b \varphi_k(x) dx \right| \rightarrow 0$$

при $n \rightarrow \infty$. ▶

11.1.3 Теорема. Пусть функции $\varphi_n(x)$ непрерывно дифференцируемы на $[a, b]$ и равномерно сходятся как ряд (11.3), так и ряд, составленный из производных:

$$\sum_{n=0}^{\infty} c_n \frac{d\varphi_n(x)}{dx}. \quad (11.4)$$

Тогда почленно продифференцированный ряд (11.4) сходится к производной $f'(x)$, т. е.

$$f'(x) = \sum_{n=0}^{\infty} c_n \varphi_n'(x).$$

Если сюжет (11.3) разворачивается на комплексной плоскости C , тогда равномерно сходящийся ряд аналитических функций можно почленно интегрировать и дифференцировать без всяких дополнительных условий. Доказательство имеет технический характер. Некоторое удивление может вызывать отсутствие требования равномерной сходимости ряда (11.4), которое в данном случае выполняется автоматически благодаря аналитичности. Причина заключается в следующем. Если остаток ряда $r_n(z)$ на контуре C меньше ε по модулю, то в силу $r_n'(z) = \frac{1}{2\pi i} \oint_C \frac{r_n(w)}{(w-z)^2} dw$ справедлива равномерная на C оценка $|r_n'(w)| < \frac{2\pi\rho^2}{L} \varepsilon$, где L — длина контура C , а ρ — максимальное расстояние от z до контура C , см. главу 10.

Коль равномерная сходимость рядов выходит в данном контексте на авансцену, естественно, повышается роль соответствующих критериев. Вот очень простой признак равномерной сходимости, который зачастую выручает.

11.1.4 Если $|c_n \varphi_n(x)| \leq \gamma_n$ при любом $x \in X$ и числовой ряд $\sum_{n=0}^{\infty} \gamma_n$ сходится, то ряд (11.1) на X сходится равномерно.

11.2 Степенные ряды

В анализе широко используются степенные ряды

$$\sum_{n=0}^{\infty} c_n z^n, \quad \text{и} \quad \text{и} \quad (11.5)$$

как в ситуации $z \in \mathbb{R}$, так и в случае $z \in \mathbb{C}$. Ключом к сходимости (11.5) является *теорема Абеля* 10.7.4, каковая в проекции на действительную ось звучит так.

11.2.1 Если ряд (11.5) сходится при каком-то $x = r$, то при условии $|x| < |r|$ он сходится абсолютно, а при $|x| \leq |r - \varepsilon|$ ещё и равномерно (при любом $\varepsilon > 0$).

Из 11.2.1 сразу следует, что область сходимости степенного ряда — это всегда промежуток $(-R, R)$, в котором ряд сходится, причём всегда абсолютно, а на любом сегменте $[-r, r] \subset (-R, R)$ — равномерно. В результате можно гарантировать, что на $(-R, R)$ ряд сходится к непрерывной функции и его можно почленно интегрировать и дифференцировать¹.

Радиус сходимости R может быть конечным или бесконечным. Что касается сходимости ряда на концах промежутка $(-R, R)$, то ситуации могут быть разные (встречаются все мыслимые).

Примеры

1. Ряды

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots, \quad (11.6)$$

$$\sin x = x - \frac{x^3}{3!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots, \quad (11.7)$$

сходятся при любом x , что моментально выводится как из признака Коши 2.10.4, так и из признака Даламбера 2.10.5.

2. Те же признаки Коши и Даламбера легко позволяют установить сходимость степенных рядов

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots, \quad (11.8)$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^5}{5} + \dots, \quad (11.9)$$

но уже при $|x| < 1$.

Во всех перечисленных случаях признаки Коши и Даламбера позволяют точно уловить границы сходимости, т. е. заодно определяют радиусы сходимости рядов, что в каком-то смысле закономерно. В общем случае радиус сходимости ряда (11.5) вычисляется с помощью верхнего предела по формуле

$$R^{-1} = \overline{\lim}_{n \rightarrow \infty} \sqrt[n]{|c_n|}.$$

¹Из-за того что степенные ряды на действительной прямой попадают в юрисдикцию ТФКП и пользуются «аналитическими благами».

11.3 Ортогональные разложения

Степенные ряды (11.5) представляют собой один из самых бесхитростных способов разложения функций по осям

$$1, x, x^2, x^3, \dots$$

По конечномерным пространствам ясно, что систему координат надо выбирать с умом, чтобы потом не кусать локти. Во многих отношениях выгодны, например, ортогональные оси. Но что такое ортогональность в бесконечномерных пространствах? Чтобы ввести соответствующее понятие, начинать надо со *скалярного произведения*. Будем считать величину



$$(f, g) = \int_a^b f(x)g(x)dx \quad (11.10)$$

скалярным произведением функций f и g . Легко проверить, что (11.10) действительно определяет операцию, которая удовлетворяет обычным свойствам скалярного произведения.

Аналогия с конечномерным случаем на этом не заканчивается. Функции считаются векторами, а равенство (11.1) их координатным представлением. Конечно, здесь не всё так просто, как выглядит, но будем пока придерживаться оптимистической ноты. Систему векторов (*базис*)

$$\{\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x), \dots\} \quad (11.11)$$

назовем *ортогональной*, если

$$(\varphi_n, \varphi_m) = \int_a^b \varphi_n(x)\varphi_m(x)dx = 0 \quad \text{при } m \neq n, \quad (11.12)$$

и *ортонормальной*, если дополнительно

$$\|\varphi_n\| = (\varphi_n, \varphi_n)^{\frac{1}{2}} = \left(\int_a^b \varphi_n^2(x)dx \right)^{\frac{1}{2}} = 1. \quad (11.13)$$

В случае ортонормальной системы (11.11) после умножения (11.1) на $\varphi_n(x)$ и интегрирования от a до b , получается²

$$\int_a^b f(x)\varphi_n(x)dx = \sum_{k=0}^{\infty} c_k \int_a^b \varphi_k(x)\varphi_n(x)dx = c_n, \quad (11.14)$$

что даёт формулу для определения коэффициентов c_n (координат $f(x)$).

Теорема Пифагора в рассматриваемой ситуации

$$\|f\|^2 = \int_a^b f^2(x)dx = \sum_{n=0}^{\infty} c_n^2 \quad (11.15)$$

называется *равенством Парсеваля*³.

В приложениях широко используются различные ортогональные системы. Например, известная тригонометрическая система

$$1, \cos x, \sin x, \dots, \cos nx, \sin nx, \dots, \quad (11.16)$$

ортогональная на⁴ $[-\pi, \pi]$, служит основанием для знаменитых *рядов Фурье*. Ортогональность (11.16) легко проверяется, например,

$$\int_{-\pi}^{\pi} \cos nx \cos kx dx = \frac{1}{2} \int_{-\pi}^{\pi} \left(\cos \frac{n+k}{2}x + \cos \frac{n-k}{2}x \right) dx = 0, \quad n \neq k.$$

Нормирование⁵ (11.16) даёт ортонормальную систему:

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \dots$$

²С учётом (11.12) и (11.13).

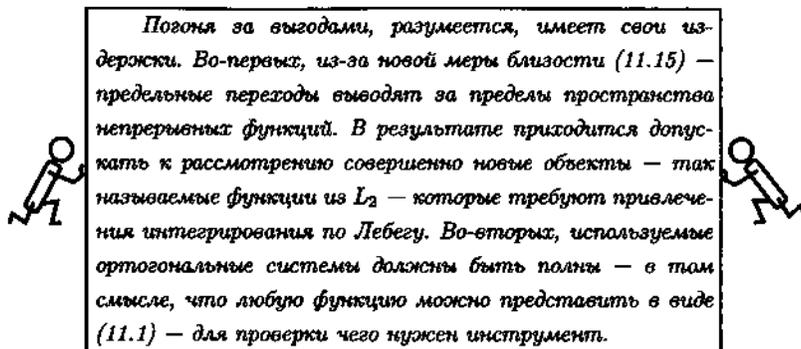
³Из (11.15) следует, что разложение (11.1) любой интегрируемой с квадратом функции имеет коэффициенты $c_n \rightarrow 0$.

⁴Любо на любом другом промежутке длины 2π .

⁵Умножение φ_n на подходящие множители для обеспечения (11.13).

Достаточно широко используются ортогональные на $[-1, 1]$ полиномы Лежандра⁶

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2 - 1)^n}{dx^n}.$$



Эти проблемы удовлетворительно решаются, но в рамках другой дисциплины — функционального анализа. В то же время ортогональные ряды существенны для классического анализа. Поэтому затронутая тематика исторически застряла на полпути. В своей основе она излагается в рамках обычного анализа, но при этом обходятся стороной острые углы.

Пространство $L_2[a, b]$ — это пространство функций $f(x)$, для которых существует интеграл от квадрата $f^2(x)$, и скалярное произведение определяется по формуле (11.10). К необходимости такого определения однозначно подталкивает всё, что сказано по поводу ортогональных разложений. В то же время, если подразумевается обыкновенное интегрирование, то данное определение L_2 внутренне противоречиво, поскольку фундаментальные по норме (11.13) последовательности могут сходиться к неинтегрируемым по Риману функциям. То есть — к функциям, норма которых не определена из-за невозможности интегрирования. Ситуация в свое время была не из лёгких. Выход из положения нашёл Лебег, разработавший более общую теорию интегрирования, см. [2, т. 5].

Интегралы Лебега и Римана совпадают, если оба существуют. Причём интеграл Лебега — кроме «спокойствия души» — ничего не даёт, и в части

⁶Популярны также ортогональные на $[-1, 1]$ полиномы Чебышева $T_n(x) = \cos n \arccos x$, но там в основе лежит иначе определяемое скалярное произведение.

вычислений о нём можно даже не упоминать. Принципиальную важность имеет сама возможность интегрирования по Лебегу. Это сводит концы с концами. Примерно как иррациональные числа. В приближённых вычислениях они не используются, но, заделывая брешни, превращают вещественную прямую в нормальное игровое поле. Аналогичную «философскую» роль в функциональных пространствах играет теория Лебега.

Полнота ортогональных систем — это тоже проблема функционального анализа — и пытаюсь расставить все точки над *i*, можно увязнуть на другой территории. Но в общих чертах здесь многое понятно. Решающую роль играют *аппроксимационные теоремы Вейерштрасса* типа 3.6.2.

Аналогичное утверждение имеет место для функций из L_2 с заменой неравенства $|f(x) - P_n(x)| < \epsilon$ на $\|f(x) - P_n(x)\| < \epsilon$, где норма определяется равенством (11.13). Поэтому система $\{1, x, x^2, \dots\}$ — полна в L_2 , равно как и любая её ортогонализация, сохраняющая возможность разложения (11.1) с любыми коэффициентами. Остаётся заметить, что «история» почти дословно повторяется для тригонометрических полиномов, что гарантирует полноту систем, лежащих в основе рядов Фурье.

- Коэффициенты c_n , обеспечивающие минимум

$$\int_a^b \left[f(x) - \sum_{k=0}^n c_k \varphi_k(x) \right]^2 dx \rightarrow \min, \quad \text{🏃} \quad (11.17)$$

определяются формулой (11.14). (?)

Оптимальность коэффициентов c_n с точки зрения (11.17) плюс полнота системы $\{\varphi_n(x)\}$ — это как раз та комбинация, которая позволяет гарантировать, что разложение (11.1) сходится именно к $f(x)$, если c_n считаются по формулам (11.14). Подразумевается, конечно, сходимость в среднем, т. е. по норме L_2 .

11.4 Механизм производящих функций

Чтобы люди не перепутались, им дают имена. Бесконечную последовательность чисел $a_0, a_1, \dots, a_n, \dots$ тоже можно поместить под одну крышу, заранее индивидуально пометив все a_n . Один из вариантов «кошелька» для $\{a_n\}$ — ряд

$$f(s) = \sum_{n=0}^{\infty} a_n s^n. \quad (11.18)$$

Если в (11.18) $f(s)$ — не просто обозначение ряда, а его компактная сумма типа

$$\frac{1}{1-s} = 1 + s + s^2 + \dots \quad \text{либо} \quad \ln(1+s) = s - \frac{s^2}{2} + \frac{s^3}{3} - \dots,$$

манипуляция оказывается эффективной для работы с $\{a_n\}$.

Достаточно вспомнить соотношение

$$(1+s)^n = 1 + C_n^1 s + C_n^2 s^2 + \dots + C_n^n s^n,$$

из которого извлекается масса тождеств для биномиальных коэффициентов. Аналогичная техника работоспособна в очень широком диапазоне.

Степенной характер производящего ряда необязателен. Возможно использование рядов Фурье $\sum_{n=0}^{\infty} c_n \varphi_n(s)$ с ортогональной системой функций $\{\varphi_n(s)\}$. Например, $\sum_{n=0}^{\infty} c_n e^{ins}$. Конечно, в теории рядов Фурье мысль направляется от функции к ряду, но движение навстречу возникает само собой.

Многие комбинаторные задачи ведут к производящим функциям, хорошо известным по другим сюжетам. И наоборот, разложение элементарных функций в ряд Тэйлора даёт последовательности коэффициентов, имеющие естественную комбинаторную интерпретацию. Взаимосвязь местами прозрачна, кое-где — загадочна.

11.5 Ряды Фурье

Рядами Фурье называют различные вариации разложения функций по ортогональным тригонометрическим системам, например по (11.16). Нормализация и подсчёт коэффициентов разложения по формуле (11.14) в этом случае приводят к представлению

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \quad (-\pi < x < \pi), \quad (11.19)$$

где

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx.$$

Разложение (11.19) можно представить в эквивалентной форме



$$f(x) = \sum_{n=0}^{\infty} A_n \sin(nx + \delta_n)$$

при очевидном пересчёте коэффициентов.

◦ Проверить справедливость разложений:

$$\begin{aligned} \bullet \quad x &= 2 \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sin nx}{n} \quad (-\pi < x < \pi). \\ \bullet \quad x^2 &= \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} (-1)^n \frac{\cos nx}{n^2} \quad (-\pi \leq x \leq \pi). \end{aligned}$$

◦ Представление (11.19) легко переносится (заменой переменных) на функции, заданные на произвольном отрезке. Например, в случае $[-l, l]$

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{l} + b_n \sin \frac{n\pi x}{l} \right), \quad (11.20)$$

где

$$a_n = \frac{1}{l} \int_{-l}^l f(t) \cos \frac{n\pi t}{l} \, dt, \quad b_n = \frac{1}{l} \int_{-l}^l f(t) \sin \frac{n\pi t}{l} \, dt. \quad (11.21)$$

Частичная сумма

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx)$$

сходится к $f(x)$ — но только в среднем, т. е. по норме L_2 .

Другие вопросы тоже легко решаются в L_2 . Причина понятна. Идеология, формулы подсчёта коэффициентов — все происходит из L_2 .

Никто, правда, не запрещает интересоваться, как обстоят дела в случае непрерывных либо кусочно-непрерывных функций. Тут, конечно, приходится городить новый огород, потому что хозяйство прибыло из L_2 , а его насильно акклиматизируют в других условиях. Все меняется с первой строчки. Равенство в (11.19)

становится проблематично, потому что теперь не ясно, куда ряд сходится. Потом, конечно, выясняется, что — «более-менее, куда надо», но с оговорками. Чтобы в точке x ряд сходился к $f(x)$ — накладывается, например, *условие Дини*: сходимость интеграла

$$\int_{-\varepsilon}^{\varepsilon} \left| \frac{f(x+u) - f(x)}{u} \right| du$$



при некотором $\varepsilon > 0$.

В местах разрывов ряд сходится к полусумме значений функции слева и справа. А разрывов не удаётся избежать даже в «непрерывном случае» по той причине, что ряд, если уж сходится на $[-l, l]$, то в силу собственной периодичности он сходится на всей оси — к функции, периодически продолженной с $[-l, l]$ на $(-\infty, \infty)$. Периодическое же продолжение $f(x)$ обязательно порождает разрывы, если $f(-l) \neq f(l)$. Поэтому на краях промежутка $[-l, l]$ сходимость «специфическая». Зато, если $f(-l) = f(l)$, то среднее арифметическое частичных сумм

$$\frac{S_0(x) + S_1(x) + \dots + S_n(x)}{n+1}$$

сходится равномерно к $f(x)$ на всей оси (*теорема Фейера*). Затем возникает проблема равномерной сходимости в обычном смысле. Затем — изучается характер убывания коэффициентов. Потом — особенности. Далее — почленное интегрирование и дифференцирование. Теория начинает занимать тома, но это имеет свои плюсы.

Здесь стоит приоткрыть завесу над внутренней кухней математики. Там слишком много говорят о важности нужных и особенно ненужных теорем. Для открытия финансирования это действительно ценно — но в том нет никакой беды, ибо так устроен мир. Более того, в том заключена мудрость, поскольку от науки требуется не столько эффективность, которая возникает сама собой, сколько незатухающее тление. Главная проблема — чем-то занять учёных. Не так важно чем, но — по специальности. Пусть доказывают ерунду, публикуются и спорят. Это поддерживает квалификацию и не даёт увлечься коммерцией. В противном случае учёные выведутся, и станет ясно, что «были нужны», а новых нигде более взять.

Поскольку мир устроен мудро, математики постоянно заняты. В результате каждая дисциплина содержит кое-что

«лишнее». С другой стороны, присмотришься к человеку — масса необязательных деталей. Но уберешь их — и человека нет.

Что касается рядов Фурье, то они, конечно, очень сильно разрослись. Тому, однако, есть серьезная причина — особая роль **гармонических колебаний**, каковыми называют обыкновенные синусоиды

$$x(t) = A \sin(2\pi\nu t + \delta), \quad (11.22)$$

где A — амплитуда, ν — частота, $2\pi\nu = \omega$ — круговая частота, δ — сдвиг по фазе. Но о гармониках (11.22) лучше говорить в другом контексте.

11.6 Интеграл Фурье

Где-то, может быть, некоторые детали «гуляют сами по себе». Но в нашей Вселенной всякая частность цепляется за другую, и цепочки уходят в бесконечность. Потому «однажды начавши, трудно кончить» не только «вкусать хорошую пиццу и чесать, где чешется», но и многое другое. Затеяв разговор о рядах, невозможно умолчать об *интегралах Фурье*, которые являются логическим продолжением темы и дают представление о берегах. При этом, разумеется, не стоит лезть на берег и углубляться в дебри.

Итак. Ряд Фурье — это *обязательно периодическая* функция. При увеличении периода $2l$ в (11.20) спектр $f(t)$ все плотнее заполняется гармониками с круговыми частотами:

$$\omega_1 = \frac{\pi}{l}, \quad \omega_2 = \frac{2\pi}{l}, \quad \dots, \quad \omega_n = \frac{n\pi}{l}, \quad \dots,$$

отстоящими друг от друга на $\Delta\omega = \frac{\pi}{l}$.

В предположении абсолютной интегрируемости $f(t)$,

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty, \quad (11.23)$$

предельный переход в (11.20) при $l \rightarrow \infty$ заменяет суммирование интегрированием:

$$\int_0^{\infty} (a_{\omega} \cos \omega x + b_{\omega} \sin \omega x) d\omega, \quad (11.24)$$

где a_ω , b_ω получаются предельным переходом

$$\frac{a_n}{\Delta\omega} \rightarrow a_\omega, \quad \frac{b_n}{\Delta\omega} \rightarrow b_\omega \quad \text{при} \quad \Delta\omega \rightarrow 0 \quad (\text{т. е. при } l \rightarrow \infty),$$

что даёт

$$a_\omega = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \cos \omega t dt, \quad b_\omega = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \sin \omega t dt. \quad (11.25)$$

Коэффициент a_0 при $l \rightarrow \infty$ пропадает в силу (11.23).

Подстановка (11.25) в (11.24) после использования формулы косинуса разности двух углов приводит к

$$f(x) = \frac{1}{\pi} \int_0^{\infty} \int_{-\infty}^{\infty} f(t) \cos \lambda(t-x) dt d\lambda \quad \text{☹} \quad (11.26)$$

что, собственно, и называют *интегралом Фурье*⁷.

Предыдущая пара формул (11.24)–(11.25) — прозрачнее понятию. Свёрнутая в клубок формула (11.26) — маскирует содержание, но компактна.

Манипуляции типа (11.24)–(11.26) чаще используются на территории \mathbb{C} , порождая популярное *преобразование Фурье*⁸

$$\hat{f}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt. \quad (11.27)$$

Тут, правда, возникает впечатление, что шило меняется на мыло, поскольку одна функция f замещается другой, \hat{f} . Однако \hat{f}

⁷Для справедливости (11.26) необходимы дополнительные требования типа *условия Дини*. Если в (11.26) $f(x)$ слева заменяется на $\frac{f(x+0) + f(x-0)}{2}$, то формула становится верной в довольно свободных предположениях.

⁸Обратное преобразование: $f(t) = \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega t} d\omega$. Замкнутый круг делится иногда в другом месте, и тогда прямое и обратное преобразования Фурье выглядят несколько по-другому,

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{i\omega t} dt, \quad f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{-i\omega t} d\omega.$$

в некоторых отношениях, особенно с точки зрения дифференцирования, получается скомпонованной рациональнее. В итоге достигается выигрыш при решении линейных дифференциальных уравнений, и, что более важно, возникают новые категории мышления [2, т. 2, гл. 8].

11.7 Преобразование Лапласа

Затронув преобразование Фурье, т. е. сказав «а», приходится говорить «бэ». Преобразование Лапласа



$$\hat{f}(p) = \int_0^{\infty} e^{-pt} f(t) dt, \quad (11.28)$$

где $p = \sigma + i\omega$, — усовершенствует методику, основанную на преобразовании (11.27). Функцию $f(t)$ при этом называют *оригиналом*, а $\hat{f}(p)$ — *изображением*, что записывают обычно как

$$\boxed{f(t) \doteq \hat{f}(p)} \quad \text{либо} \quad \boxed{\hat{f}(p) = \mathcal{L}\{f(t)\}}.$$

Традиционно предполагается кусочная непрерывность $f(t)$. Точнее, на любом ограниченном участке допускается не более чем конечное число разрывов первого рода⁹. Разумеется также, $f(t) \equiv 0$ при $t < 0$, и показатель роста ограничен ($|f(t)| < Ke^{\alpha t}$) для некоторых α , $K > 0$ и всех $t \geq 0$.

Выгоды преобразования Лапласа по сравнению со стандартным преобразованием Фурье заключаются в том, что при больших $\operatorname{Re} p = \sigma$ метод (11.28) справляется с быстрорастущими функциями, тогда как (11.27) упирается в расходящиеся интегралы. В этом, собственно, и заключается роль действительной части σ , которую можно выбирать по своему усмотрению, исходя из потребностей задачи¹⁰. В частности, при $\sigma = 0$ получается преобразование Фурье функции $f(t)$, если $f(t) = 0$ при $t < 0$.

⁹ Разрыв первого рода характеризуется существованием конечных пределов слева и справа.

¹⁰ Правильнее даже сказать, что σ вообще заботы не требует. Выбирать и фиксировать значения σ не надо. Важно лишь понимать, что при достаточно больших σ интеграл (11.28) сходится.

Успех преобразования Лапласа при изучении линейных систем определяется двумя обстоятельствами. Во-первых, линейностью:

$$\lambda f(t) + \mu g(t) \equiv \lambda \hat{f}(p) + \mu \hat{g}(p).$$

Во-вторых, заменой дифференцирования умножением. Точнее говоря,

$$\begin{aligned} \dot{f}(t) &\equiv \int_0^{\infty} e^{-pt} \dot{f}(t) dt = \int_0^{\infty} e^{-pt} df(t) = \\ &= e^{-pt} f(t) \Big|_0^{\infty} - \int_0^{\infty} p e^{-pt} f(t) dt = p \hat{f}(p) - f(0), \end{aligned} \quad (11.29)$$

где при взятии интеграла по частям действительная часть p выбирается настолько большой, чтобы модуль $e^{-pt} f(t)$ при $t = \infty$ обращался в ноль.

Естественно, предпочтение отдается задачам, в которых $f(0) = 0$, и тогда $\boxed{\dot{f}(t) \equiv p \hat{f}(p)}$. Есть и другие удобные свойства.

• *Изображение интеграла:*

$$\int_0^t f(\tau) d\tau \equiv \frac{1}{p} \hat{f}(p). \quad (11.30)$$

◀ В силу (11.29) для $g(t) = \int_0^t f(\tau) d\tau$, $g(0) = 0$, имеем

$$f(t) = g'(t) \equiv p \hat{g}(p),$$

т. е. $f(p) = p \hat{g}(p) \Rightarrow (11.30)$. ▶

• *Дифференцирование изображения:* $\hat{f}'(p) \equiv -t f(t)$.

• *Интегрирование изображения:* $\frac{f(t)}{t} \equiv \int_p^{\infty} \hat{f}(p) dp$, где путь

интегрирования целиком лежит в области существования изображения $\hat{f}(p)$.

- **Запаздывание:**

$$f(t - a) \equiv e^{-pa} \hat{f}(p) \quad (a > 0).$$

- ◀ Поскольку $f(t - a) = 0$ при $t < a$, то

$$f(t - a) \equiv \int_a^{\infty} f(t - a) e^{-pt} dt = \int_0^{\infty} f(\tau) e^{-p(\tau+a)} d\tau = e^{-pa} \hat{f}(p). \quad \blacktriangleright$$

- **Свертка:**

$$\int_0^t f(\tau) g(t - \tau) d\tau \equiv \hat{f}(p) \hat{g}(p). \quad (11.31)$$

◀ Здесь и выше обходятся стороной вопросы ограниченности скорости роста оригиналов на бесконечности. В (11.31) подразумевается, что при условии существования изображений $\hat{f}(p)$ и $\hat{g}(p)$ преобразование (11.28) свертки

$\int_0^t f(\tau) g(t - \tau) d\tau$ существует и равно правой части (11.31)¹¹. Если без этих хлопот, то обоснование сводится к выкладке

$$\begin{aligned} \int_0^t f(\tau) g(t - \tau) d\tau &\equiv \int_0^{\infty} f(\tau) d\tau \int_{\tau}^{\infty} e^{-p\xi} g(t - \tau) d\xi = \\ &= \int_0^{\infty} f(\tau) e^{-p\tau} d\tau \int_0^{\infty} g(\xi) e^{-p\xi} d\xi = \hat{f}(p) \hat{g}(p). \quad \blacktriangleright \end{aligned}$$

Примеры и упражнения

$$f(\lambda t) \equiv \frac{1}{\lambda} \hat{f}\left(\frac{p}{\lambda}\right), \quad \lambda > 0.$$

$$t^n \equiv \frac{n!}{p^{n+1}}.$$

$$\sin \omega t = \frac{e^{i\omega t} - e^{-i\omega t}}{2i} \equiv \frac{1}{2i} \left(\frac{1}{p - i\omega} - \frac{1}{p + i\omega} \right) = \frac{\omega}{p^2 + \omega^2}.$$

¹¹ Иначе говоря, если значения действительной части в $p = \sigma + i\omega$ можно взять такими, что преобразование (11.28) существует как для $f(t)$, так и для $g(t)$, — то подходящее σ существует и для свертки (11.31).

$$\begin{aligned}\frac{\sin t}{t} &= \operatorname{arctg} p, \\ t \sin \omega t &= \frac{2p\omega}{(p^2 + \omega^2)^2}, \\ e^{\lambda t} t^n &= \frac{n!}{(p - \lambda)^{n+1}}, \\ \theta(t) &= \frac{1}{p}, \quad \text{где } \theta(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0. \end{cases}\end{aligned}$$

• Изображение $\hat{f}(p)$ функции $f(t)$ является аналитической функцией в области $\operatorname{Re} p > \alpha$, где α — оценка степени роста $f(t)$, т. е. такая постоянная, что $|f(t)| < K e^{\alpha t}$ для некоторого $K > 0$ и всех $t \geq 0$.

• Если $p \rightarrow \infty$ в секторе $|\arg p| < \frac{\pi}{2} - \varepsilon$ ($\varepsilon > 0$), то

$$\lim_{p \rightarrow \infty} p \hat{f}(p) = f(0).$$

А если существует предел $f(\infty) = \lim_{t \rightarrow \infty} f(t)$, то

$$\lim_{p \rightarrow 0} p \hat{f}(p) = f(\infty),$$

где $p \rightarrow 0$ внутри того же сектора $|\arg p| < \frac{\pi}{2} - \varepsilon$.

Обращение преобразования Лапласа. Оригинал по изображению (11.28) вычисляется по формуле Меллина

$$f(t) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} e^{pt} \hat{f}(p) dp, \quad (11.32)$$

где интегрирование идет вдоль вертикальной прямой $\operatorname{Re} p = \sigma$, причём значение σ в (11.32) — не менее показателя степени роста функции $f(t)$.

◀ Рецепт (11.32) сводится к обращению преобразования Фурье. Применяя к экспоненциально убывающей функции $e^{-\sigma t} f(t)$ сначала преобразование (11.27), а потом его обращение, имеем с учётом $f(\omega) \equiv 0$ при $\omega < 0$

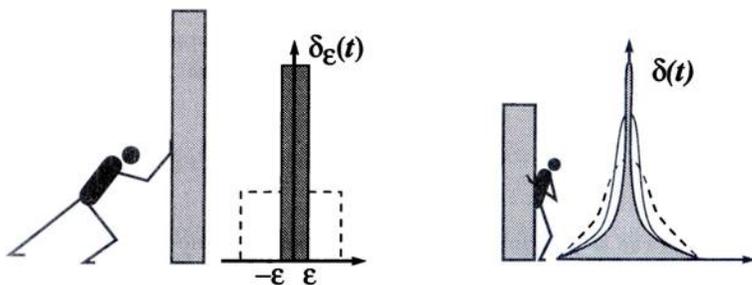
$$\begin{aligned}e^{-\sigma t} f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} e^{-\sigma\omega} f(\omega) e^{-i\tau(t-\omega)} d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\tau t} d\tau \int_0^{\infty} e^{-(\sigma+i\tau)\omega} f(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\tau t} \hat{f}(p) d\tau,\end{aligned}$$

где $p = \sigma + i\tau$. Деление на $e^{-\sigma t}$ приводит к

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(\sigma+i\tau)t} \hat{f}(p) d\tau = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{pt} \hat{f}(p) dp. \quad \blacktriangleright$$

11.8 Дельта-функция

О дельта-функции необходимо сказать главное на первых страницах высшей школы, потому что она, $\delta(t)$, то и дело появляется там и сям в сопровождении абсурдных комментариев. Феномен принадлежит теории обобщённых функций [2, т. 5], и за пределами этой территории $\delta(t)$ производит иногда странное впечатление. Дабы «по дороге в баню» не отсылать читателя «в Сочи» по адресу [2, т. 5, глава 8], напомним основные моменты здесь, чуть-чуть замахиваясь на все обобщённые функции.



Изначально $\delta(t)$ определялась как предел единичных импульсов¹² $\delta_\epsilon(t)$, например прямоугольной формы, при стремлении к нулю ширины импульса, $\epsilon \rightarrow 0$. Трудность заключалась в противоречивости обстоятельств. В пределе получалась функция, равная бесконечности в нуле и нулю — в остальных точках. Однако ситуации, в которых возникала потребность в чем-то подобном, всегда сводились к вариантам, когда $\delta(t)$ стояла под интегралом. То есть $\delta(t)$ нужна была не как функция, а как *нечто*, обеспечивающее при интегрировании определенный эффект. Но тогда и обыкновенный предел не нужен был.

¹²Характеризуемых условием $\int_{-\infty}^{\infty} \delta_\epsilon(t) dt = 1$.

Хватало сходимости интеграла

$$\int_{-\infty}^{\infty} \delta_\varepsilon(t) \varphi(t) dt \rightarrow \varphi(0) \quad \text{при } \varepsilon \rightarrow 0,$$

что позволяло определить $\delta(t)$ как особый предел $\delta_\varepsilon(t) \rightarrow \delta(t)$ в смысле

$$\int_{-\infty}^{\infty} \delta_\varepsilon(t) \varphi(t) dt \rightarrow \int_{-\infty}^{\infty} \delta(t) \varphi(t) dt.$$

Иначе говоря, под $\delta(t)$ достаточно было понимать «ничто», действующее на функции φ по правилу:

$$\langle \delta, \varphi \rangle = \int_{-\infty}^{\infty} \delta(t) \varphi(t) dt = \varphi(0). \quad (11.33)$$

Но (11.33) — это функционал. Надо лишь подходящим образом подобрать пространство вспомогательных функций φ .

Соотношение (11.33) служит прообразом общего определения. Обобщённые функции f описываются как функционалы

$$\langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(t) \varphi(t) dt, \quad \varphi \in \mathbb{D}, \quad (11.34)$$

на множестве \mathbb{D} финитных функций¹³, непрерывно дифференцируемых любое число раз. Последовательность $\{f_n\}$ полагается сходящейся к обобщённой функции f , если

$$\langle f_n, \varphi \rangle \rightarrow \langle f, \varphi \rangle$$

для любой функции $\varphi \in \mathbb{D}$.

11.8.1 Всякий линейный непрерывный функционал $\langle f, \varphi \rangle$ на \mathbb{D} называется обобщённой функцией¹⁴.

¹³ Непрерывная функция $\varphi(t)$ называется финитной, если область, в которой $\varphi(t) \neq 0$, — ограничена. Замыкание области, где $\varphi(t) \neq 0$, именуется носителем $\varphi(t)$ и обозначается как $\text{supp } \varphi$. Разные функции из \mathbb{D} могут быть отличны от нуля на разных областях (у каждой свой носитель).

¹⁴ Не каждый функционал на \mathbb{D} может быть представлен в виде (11.34). Если это возможно, обобщённую функцию $f(t)$ называют регулярной.

Дифференцирование определяется *правилом переброски производной*, $\langle f', \varphi \rangle = -\langle f, \varphi' \rangle$,

$$\begin{aligned} \langle f', \varphi \rangle &= \int_{-\infty}^{\infty} f'(t)\varphi(t)dt = \\ &= \int_{-\infty}^{\infty} \varphi(t)df(t) = \varphi(t)f(t)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(t)\varphi'(t)dt = -\langle f, \varphi' \rangle, \end{aligned}$$

каковое в ситуации (11.34) возникает само по себе в результате интегрирования по частям с учётом финитности функций φ .

На $\delta(t)$ общая часть теории проецируется следующим образом. *Дельта-функция* декларируется как функционал, действующий в \mathbb{D} по правилу $\langle \delta, \varphi \rangle = \varphi(0)$. Производная $\delta'(t)$ действует в результате как

$$\langle \delta', \varphi \rangle = -\langle \delta, \varphi' \rangle = -\varphi'(0). \quad \text{🧘}$$

Следствием *правила переброски производной* является соотношение

$$\theta'(t) = \delta(t),$$

где $\theta(t)$ — *функция Хэвисайда*¹⁵.

Замена переменной при интегрировании указывает на справедливость соотношений

$$\int_{-\infty}^{\infty} \delta(t-a)\varphi(t)dt = \varphi(a), \quad \int_{-\infty}^{\infty} \delta(at)\varphi(t)dt = \frac{1}{|a|}\varphi(0),$$

а также

$$\delta[f(t)] = \sum_{f(t_k)=0} \frac{\delta(t-t_k)}{|f'(t_k)|}.$$

В частности,

$$\delta(t^3 - 1) = \frac{\delta(t-1)}{3}, \quad \delta(t^2 - 1) = \frac{\delta(t-1)}{2} + \frac{\delta(t+1)}{2}.$$

¹⁵ Единичный скачок: $\theta(t \geq 0) = 1$ и $\theta(t < 0) = 0$.

Преобразование Лапласа $\hat{f}(p) = \int_0^{\infty} e^{-pt} f(t) dt$ обобщённых

функций было бы естественно рассматривать как значение функционала f на функциях e^{-pt} . Но $e^{-pt} \notin \mathbb{D}$, однако \mathbb{D} можно безболезненно переопределить, например, как совокупность бесконечно дифференцируемых функций, для которых $t^k \varphi^{(l)}(t) \rightarrow 0$ при $t \rightarrow \infty$ и любых целых $k, l \geq 0$. В этом случае обычные свойства преобразования Лапласа (раздел 11.7) остаются без изменения. В частности,

$$\langle \delta, e^{-pt} \rangle = e^{-pt} \Big|_{t=0} = 1, \quad \text{т. е.} \quad \boxed{\hat{\delta}(p) = 1.}$$

Обычным образом определяются изображения производных $\delta^{(n)}(t)$,

$$\widehat{\delta^{(n)}}(p) = \langle \delta^{(n)}, e^{-pt} \rangle = p^n,$$

запаздывания, $\delta(t - \tau) \hat{=} e^{-p\tau}$, и т. п.

Глава 12

Неподвижные точки

*Подход «на пальцах»
раскрепощает и вдохновляет.*

12.1 Проблемы существования решения

Исследование любой модели начинается, как правило, с выяснения существования решения у той или иной системы уравнений $f(x) = 0$, которую всегда можно привести к виду¹

$$x = T(x). \quad (12.1)$$

Решение (12.1) называют *неподвижной точкой* оператора T .

Существование решения уравнения (12.1) обеспечивается, к примеру, *принципом сжимающих отображений (теорема 6.7.2)*. За пределами этого результата мало что известно в рамках традиционного математического образования. Топологические методы исследования неподвижных точек считаются излишне сложными. До некоторой степени на слуху следующая теорема².

12.1.1 Теорема Брауэра. *Любое непрерывное отображение $T(x)$ шара B в себя имеет неподвижную точку $x \in B$.*

¹Форма (12.1) иногда возникает сама собой, иногда — искусственно, с прицелом, например, вычислительного характера.

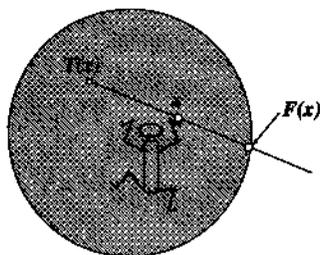
²Как правило, без доказательства.

Далее будет изложена общая идеология, в рамках которой результаты типа 12.1.1 тривиальны. Но на теорему Брауэра интересно взглянуть и с другой точки зрения.

Пластичный шар B , ограниченный сферой S , можно сплющить и распластать по S . Это означает существование непрерывного отображения $F(x)$ шара на сферу. Факт, конечно, банальный. Потребуем, однако, чтобы при такой деформации точки S оставались на месте. Другими словами, зафиксируем границу S и, деформируя внутренность B , пробуем распластать шар по граничной сфере. Интуитивно ясно, что, не разрывая шар, сделать это невозможно. Так что «не существует непрерывного отображения шара B в граничную сферу S , которое все точки S оставляет на месте»³.

Негативный характер утверждения не умаляет его ценности, ибо надлежащая переформулировка приводит к позитивному результату 12.1.1.

◀ Действительно, в предположении противного, $x \neq T(x)$ для любого $x \in B$. Соединим x и $T(x)$ отрезком прямой и продолжим его за точку x до пересечения со сферой S в точке $F(x)$. Понятно, что $F(x)$ — непрерывное отображение B в S , оставляющее сферу S на месте. Противоречие. ▶



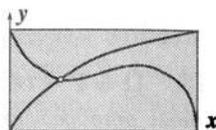
Л. Э. Я. Брауэр, 1881–1966

Несмотря на интуитивно прозрачное доказательство, попытки строго расставить точки над i уведут здесь весьма далеко. В связи с этим теорему Брауэра расценивают обычно как гроссмейстерский трюк. Кое-что по части разрешимости $f(x) = 0$ устанавливается легко. Начиная с теоремы Больцано — Коши⁴. В той же нише есть чуть более сложные результаты. Скажем, факт пересечения нелинейных диагоналей, непрерывных кри-

³Тем не мене строгое обоснование этого факта нетривиально.

⁴Если значения непрерывной функции $f(x)$ на концах отрезка $[a, b]$ имеют разные знаки, то $f(c) = 0$ в некоторой точке $c \in [a, b]$.

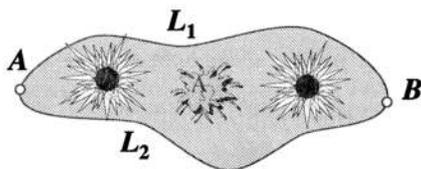
вых, соединяющих противоположные вершины прямоугольника,



(12.2)

Имест смысл доказать (12.2) в качестве упражнения. Феномен, правда, выглядит сбоку бантиком. Но к нему сводится красивая фольклорная задача⁵, вроде бы совсем из другой оперы.

Из пункта A в пункт B ведут две дороги L_1 и L_2 , по которым две точки из A в B могут двигаться, оставаясь все время в поле зрения друг друга



Вопрос в том, могут ли точки двигаться, — одна из A в B , другая из B в A , — оставаясь все время невидимы друг другу?

◀ Пусть x обозначает расстояние от A до точки на L_1 вдоль L_1 . Соответственно, y — расстояние от A до точки на L_2 вдоль L_2 . Одна кривая на рисунке (12.2) описывает движение точек из A в B , когда они остаются все время в поле зрения друг друга. Движению точек, одной из A в B , другой из B в A , — отвечает некоторая траектория, изображенная другой кривой на (12.2). В силу непрерывности кривые пересекаются. В пересечении точки видны друг другу. Решение задачи — отрицательно. ▶ Банальность (12.2) превращается таким образом в изюминку.

⁵Поскольку «ничто не было названо именем первооткрывателя» — об авторстве задач и теорем чаще всего трудно говорить. Достаточно сказать, что формула Эйнштейна $E = mc^2$ открыта Хэвисайдом, закон Воля — Мариотта — Гуком, преобразования Лоренца — Фогтом.

12.2 Вращение векторного поля

Далее мы обрисуем не вполне строгую, зато геометрически наглядную картину теории вращения векторных полей.

Итак, пусть Ω — некоторая область в \mathbb{R}^n , $\dot{\Omega} = \Gamma$ — её граница, а $f(\Gamma)$ — образ границы при непрерывном отображении f . В этом случае f называют также оператором⁶ и говорят, что f задаёт в \mathbb{R}^n векторное поле $f(x)$.

Допустим, речь идет о существовании решения уравнения

$$f(x) = 0 \quad (12.3)$$

в области Ω , т. е. о принадлежности нуля образу $f(\Omega)$. Другими словами, покрывает образ $f(\Omega)$ точку 0 или нет? Если да, то элемент $x^* \in \Omega$, переходящий в нуль, $f(x^*) = 0$, называют нулём векторного поля $f(x)$.

В случае, когда некая «благообразная» область  слегка деформируется преобразованием f во что-нибудь типа

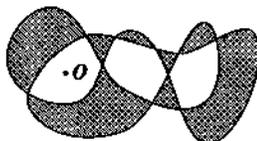
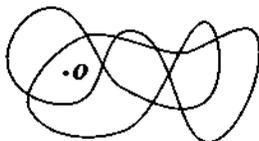


или даже



, вопрос о включении $0 \in f(\Omega)$

сводится к выяснению расположения нуля внутри образа границы $f(\Gamma)$ или — снаружи. В общем случае Γ может переходить в довольно хитроумные поверхности $f(\Gamma)$. В частности, окружность может перейти во что-нибудь типа



и тут уже не ясно, где внутренность $f(\Gamma)$. Но как бы там ни было, разрешимость (12.3) в Ω по значениям $f(x)$ на границе $\dot{\Omega}$

⁶ Отображение, оператор, функция — используются как синонимы.

поддаётся выяснению. Путь к этому не очень длинный, если не злоупотреблять деталями.

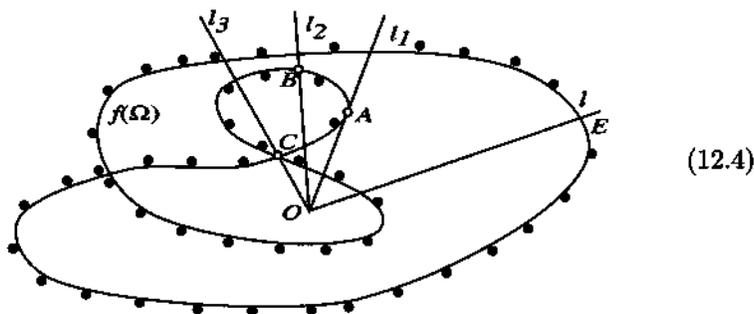
Итак. Будем считать Γ и f гладкими (в противном случае можно перейти к гладким аппроксимациям), и у поверхности $f(\Gamma)$ определим внешнюю и внутреннюю стороны. Для этого возьмём маленький кусочек Δ поверхности Γ — у него внутренняя сторона та, которая соприкасается с Ω , — и на фрагменте $f(\Delta) \subset f(\Gamma)$ стороны определим стандартным образом:

В *регулярной* точке $x \in \Delta$ локальный базис $\{e_1, \dots, e_n\}$ выбирается так, чтобы вектор e_1 был направлен по внешней нормали, а векторы e_2, \dots, e_n лежали в касательном пространстве T_x к Δ в точке x . Нормаль e'_1 к поверхности $f(\Delta)$ в точке $f(x)$ называют внешней, если базис

$$\{e'_1, \dots, e'_n\} = \{e_1, f'(x)e_2, \dots, f'(x)e_n\}$$

одинаково ориентирован с базисом⁷ $\{e_1, \dots, e_n\}$.

Далее разобьём $f(\hat{\Omega})$ на маленькие участки — *накладывающиеся друг на друга* — и, начиная с исходного Δ , продолжим определение внешней стороны на соседние участки, согласуя ориентацию на пересечениях, пока не определится внешняя сторона на всей поверхности $f(\hat{\Omega})$. В плоском случае этому сопоставить можно прокатывание шарика по одной из сторон, как на ниже-следующем рисунке.



Считая поле $f(x)$ невырожденным на $\hat{\Omega}$, т. е. $f(x) \neq 0$ при

⁷ Два базиса $\{e_1, \dots, e_n\}$ и $\{e'_1, \dots, e'_n\}$ называются одинаково ориентированными, если $e'_i = \sum a_{ij} e_j$ и $\det[a_{ij}] > 0$.

$x \in \dot{\Omega}$, предположим также (для удобства рассуждений), что поверхность $f(\dot{\Omega})$ связна и не имеет слипшихся участков⁸.

Выпустим из нуля произвольный луч l и посчитаем, сколько раз l протыкает $f(\dot{\Omega})$ *изнутри наружу* — допустим, γ_{\oplus} раз, и сколько — *снаружи внутрь* — пусть γ_{\ominus} раз⁹. Величину

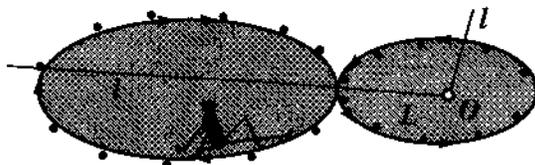
$$\gamma(f, \dot{\Omega}) = \gamma_{\oplus} - \gamma_{\ominus} \quad (12.5)$$

назовем *вращением векторного поля f на $\dot{\Omega}$* .

Для *корректности* определения требуется независимость $\gamma(f, \dot{\Omega})$ от выбора l . Очевидно, величина $\gamma(f, \dot{\Omega})$, как функция l , локально постоянна — не меняется при малых шевелениях луча. Изменений можно ожидать, если при повороте луч пересекает точку самопересечения C или касания A . Геометрически ясно, что при вращении l против часовой стрелки и переходе, скажем, критического положения l_1 , появляется пара точек протыкания. Причём, если в одной из них луч протыкает $f(\dot{\Omega})$ *изнутри наружу*, то в другой — наоборот. Поэтому точки протыкания появляются или исчезают полярными парами, и γ (12.5) не меняется.

Если $f(\dot{\Omega})$ имеет слипшиеся участки, то надо считать, что в соответствующем месте луч протыкает поверхность столько раз, сколько там слиплось слоев — и все остается по-прежнему. Это же должно быть сказано в отношении лучей, протыкающих поверхность в точках самопересечения $f(\dot{\Omega})$ — луч l_2 на рис. (12.4). В рассматриваемом случае получается $\gamma(f, \dot{\Omega}) = 2$.

В ситуации, когда окружность преобразуется в «восьмёрку»,

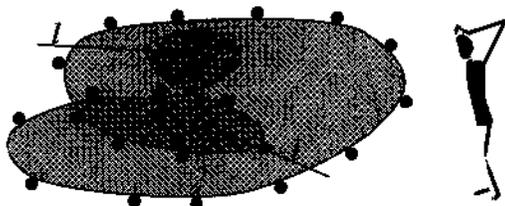


⁸Все эти требования снимаются рутинными приемами.

⁹Если у $\dot{\Omega}$ несколько компонент связности , манипуляции проводятся с каждой компонентой, и засчитываются в общую копилку.

$\gamma(f, \hat{\Omega}) = -1$, если O лежит в области L .

В случае отображения $\hat{\Omega}$ в более замысловатую кривую



возникают следующие варианты: $\gamma(f, \hat{\Omega}) = 1$, если O лежит в «внешней» области; $\gamma(f, \hat{\Omega}) = 2$, если O — в области E ; наконец, $\gamma(f, \hat{\Omega}) = 0$, если $\cap \in D$.

12.2.1 Теорема. Если $0 \in \Omega$, то вращение тождественного отображения $I(x) \equiv x$ на $\hat{\Omega}$ равно 1.

Это, пожалуй, единственная ситуация, где определение внешней стороны у $I(\hat{\Omega})$ не вызывает вопросов, поскольку $\hat{\Omega}$ остается на месте. Луч сначала протыкает $\hat{\Omega}$ изнутри, потом — снаружи, наконец последний раз — изнутри.

Все вышесказанное подразумевает пространства \mathbb{R}^n размерности $n \geq 2$. Для определенной «свободы передвижения» полезно доопределить вращение в одномерном случае.

Любая ограниченная область $\Omega \in \mathbb{R}^1$ представляет собой интервал (α, β) . Если $f(\alpha) < 0$, $f(\beta) > 0$, то $\gamma(f, \hat{\Omega}) = 1$. В случае

$$f(\alpha) > 0, \quad f(\beta) < 0 \quad \Rightarrow \quad \gamma(f, \hat{\Omega}) = -1.$$

Если же $f(\alpha)$ и $f(\beta)$ одного знака, то $\gamma(f, \hat{\Omega}) = 0$.

12.3 Гомотопия векторных полей

Определили вращение мы визуально, но вычислять-то $\gamma(f, \hat{\Omega})$ потребуются инструментально. Соответствующим инструментом для этого служат *деформации*, или *гомотопические переходы*.

12.3.1 Определение. *Невырожденные¹⁰ на $\hat{\Omega}$ отображения f_0 и f_1 называются гомотопными на $\hat{\Omega}$, если существует такая невырожденная непрерывная деформация (гомотопия) $H(x, t)$,*

$$H : \hat{\Omega} \times [0, 1] \rightarrow \mathbb{R}^n \setminus \{0\},$$

что¹¹

$$H(x, 0) \equiv f_0(x), \quad H(x, 1) \equiv f_1(x).$$

Операция гомотопического перехода соответствует непрерывное деформирование поверхности $f_0(\hat{\Omega})$ в поверхность $f_1(\hat{\Omega})$. Невырожденность гомотопии, $H(x, t) \neq 0$, означает, что деформируемая поверхность не пересекает точку 0.

Геометрически ясно, что невырожденная деформация не может поменять величину вращения $\gamma(\cdot)$, поскольку — движется ли луч, деформируется ли поверхность — точки протыкания исчезают и возникают парами. Разумеется, пока деформируемая поверхность не пересекает нуля, т. е. $H(x, t) \neq 0$. Поэтому справедлив следующий важный результат.

12.3.2 Теорема. *Гомотопные векторные поля имеют одинаковые вращения.*



Становится ясно, что бесконечное разнообразие практических ситуаций может быть разбито на классы. Представителями классов могут служить, например, $f(x)$, отображающие $\hat{\Omega}$ на единичную сферу S с центром в нуле. Деформация $f(\hat{\Omega})$ на S с последующим разглаживанием складок — порождает многослойную поверхность, окутывающую сферу $|\gamma(f, \hat{\Omega})|$ раз¹². Отрицательный знак $\gamma(f, \hat{\Omega})$ связан с выворачиванием $\hat{\Omega}$ наизнанку.

«Число окутываний» — и есть вращение $\gamma(f, \hat{\Omega})$ с точностью до знака. Если на плоскости в нуль вбить гвоздь, то число окутываний нуля деформацией контура $H(\hat{\Omega}, t)$ поменять невозможно, не нарушая условия $H(x, t) \neq 0$.

¹⁰ Не обращающиеся в нуль.

¹¹ Построение гомотопии от f_0 к f_1 равносильно продолжению отображения $H : X \times \{0\} \times \{1\} \rightarrow \mathbb{R}^n \setminus \{0\}$ на пространство $X \times [0, 1]$.

¹² Если f отображает S в S , то $\gamma(f, S)$ — это так называемая топологическая степень отображения f .



Несмотря на элементарный характер утверждения 12.3.2, оно играет принципиальную роль, поскольку служит основным инструментом для вычисления вращений векторных полей, позволяя переходить с помощью гомотопии от изучаемых полей к более простым, часто стандартным, вращениям которых известно.

12.3.3 Самый распространенный на практике вид деформаций – линейная гомотопия:

$$H(x, t) = (1 - t)f(x) + tg(x),$$

связывающая отображения f и g .

Линейная гомотопия может оказаться вырожденной лишь в том случае, когда найдется точка $x \in \Omega$, в которой векторы $f(x)$ и $g(x)$ противоположно направлены, $f(x) = \lambda g(x)$, $\lambda < 0$,

$$f(x) = -\frac{t}{1-t}g(x).$$

Поэтому в случае **непротивоположной направленности** в любой точке $x \in \Omega$ отображения $f(x)$ и $g(x)$ **линейно гомотопны** и, в соответствии с теоремой 12.3.2, имеют одинаковые вращения.

В связи с теоремой 12.3.2 закономерно возникает вопрос о справедливости обратного утверждения. Обязательно ли гомотопны поля с одинаковым вращением? В общем случае это не так. Если же Ω – шар, положительный ответ на поставленный вопрос даёт знаменитая **теорема Хопфа**.

12.3.4 Теорема. *Невырожденные на сфере $S \subset \mathbb{R}^n$, $n \geq 2$, векторные поля с одинаковым вращением – гомотопны.* \triangleright

Теорема 12.3.4 не участвует в формировании излагаемого далее аппарата, но показывает, что «аппарат» достигает максимума возможного. В том смысле что за рамками соответствующей «деформационной идеологии» не остается каких-либо других топологических инвариантов (помимо вращения), из которых можно было бы вытащить дополнительную информацию.

Выше подразумевался гомотопический переход от f_0 к векторному полю f_1 , деформирующий поверхность $f_0(\dot{\Omega})$ в $f_1(\dot{\Omega})$, с помощью гомотопии

$$f_t(x) = H(x, t).$$

Однако деформация может осуществляться не только за счёт изменения оператора, но и за счёт изменения области. Как бы деформация ни происходила, для сохранения величины вращения существенно лишь, чтобы деформируемая поверхность не пересекала точку 0. Вот точная формулировка результата. Пусть Ω_t ($t \in [0, 1]$) описывает непрерывный переход от области Ω_0 к Ω_1 . Результирующую деформацию $f_t(\dot{\Omega}_t)$ назовем невырожденной, если $f_t(x) \neq 0$ при любых $t \in [0, 1]$, $x \in \dot{\Omega}_t$.

12.3.5 Теорема. Если деформация $f_t(\dot{\Omega}_t)$ невырождена, то

$$\gamma(f_0, \dot{\Omega}_0) = \gamma(f_1, \dot{\Omega}_1).$$

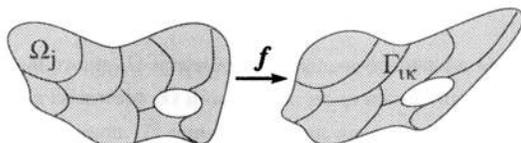
Это простое обобщение теоремы 12.3.2 иногда оказывается полезным, позволяя при вычислении вращения переходить не только к более простому полю, но и к более простой области.

12.4 Ядро теории

12.4.1 Теорема. Пусть векторное поле $f(x)$ определено и невырождено на границах областей $\Omega, \Omega_1, \dots, \Omega_m$, причём Ω_j попарно не пересекаются и $\dot{\Omega} = \bigcup \dot{\Omega}_j$. Тогда

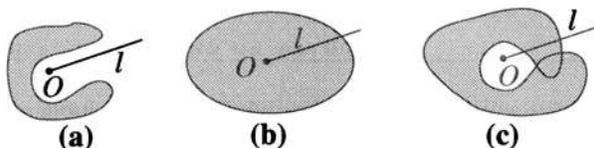
$$\gamma(f, \dot{\Omega}) = \sum_{j=1}^{\infty} \gamma(f, \dot{\Omega}_j). \quad (12.6)$$

◀ Результат очевиден в контексте определения $\gamma(\cdot)$ методом «протыкающего луча». Если Ω_i и Ω_k граничат друг с другом, то граничат и $f(\Omega_i)$ с $f(\Omega_k)$ по некоторому участку Γ_{ik} . Протыканию лучом Γ_{ik} , например, изнутри при подсчёте $\gamma(f, \Omega_i)$ будет соответствовать протыкание Γ_{ik} снаружи при подсчёте $\gamma(f, \Omega_k)$.



Поэтому в сумме (12.6) справа все протыкания внутренних перегородок взаимно сократятся. ▶

Будем говорить, что невырожденное на $\hat{\Omega}$ векторное поле $f(x)$ имеет *ахиллесово направление*, если $f(x)$ не принимает значений какого-то направления, как на рис. (а).



Иначе говоря, $f(x) \neq \lambda h$ для всех $x \in \hat{\Omega}$, $\lambda > 0$ и некоторого ненулевого $h \in \mathbb{R}^n$. На рисунках (b) и (c) направления $f(x)$ исчерпывают полный «телесный» угол.

12.4.2 Теорема. Если векторное поле $f(x)$ на $\hat{\Omega}$ имеет ахиллесово направление, то $\gamma(f, \hat{\Omega}) = 0$.

◀ Для доказательства достаточно в качестве l взять луч, задаваемый вектором h ,

$$l = \{x : x = \lambda h, \lambda \in [0, \infty)\},$$

он поверхность $f(\hat{\Omega})$ вообще не протыкает. ▶

Если поле $f(x)$ на $\hat{\Omega}$ принимает все направления, т. е. параметрическое уравнение $f(x) = \lambda h$ имеет решение $x \in \hat{\Omega}$ при любом векторе $h \neq 0$ и некотором $\lambda > 0$, то для $\gamma(f, \hat{\Omega}) \neq 0$ этого, конечно, недостаточно. Однако:

12.4.3 Если поле $f(x)$ на $\hat{\Omega}$ принимает все направления только по одному разу¹³, то $\gamma(f, \hat{\Omega}) \neq 0$, точнее: $|\gamma(f, \hat{\Omega})| = 1$.

12.4.4 Лемма. Пусть поле $f(x)$ невырождено на замыкании области Ω , т. е. $f(x) \neq 0$ при любом $x \in \bar{\Omega}$. Тогда $\gamma(f, \hat{\Omega}) = 0$.

◀ Разобьем Ω на непересекающиеся области Ω_j настолько малых размеров, что все векторы на каждом замыкании $\bar{\Omega}_j$ составляют друг с другом острый угол. В этом случае на любой границе $\hat{\Omega}_j$ поле f имеет азимутально направление. Поэтому (теорема 12.4.2) все $\gamma(f, \hat{\Omega}_j) = 0$. Окончательный вывод следует из теоремы 12.4.1. ▶

Из леммы 12.4.4 вытекает фундаментальный результат:

12.4.5 Теорема. Если $\gamma(f, \hat{\Omega}) \neq 0$, то уравнение $f(x) = 0$ разрешимо в Ω , т. е. в Ω существует нуль векторного поля $x^* \in \Omega$, $f(x^*) = 0$.

Таким образом, любая теорема об отличии от нуля вращения векторного поля f может трактоваться как принцип разрешимости уравнения $f(x) = 0$.

Отметим в заключение справедливость полезного дополнения к теореме 12.4.1, вытекающего из объединения теоремы 12.4.1 с леммой 12.4.4.

12.4.6 Утверждение теоремы 12.4.1 справедливо без предположения $\bar{\Omega} = \bigcup \bar{\Omega}_j$, но при условии невырожденности f на $\bar{\Omega} \setminus \bigcup \bar{\Omega}_j$.

Что касается практического вычисления $\gamma(f, \hat{\Omega})$, то главный акцент приходится обычно на поиск гомотопии изучаемого поля к какому-либо довольно простому типу $f(x)$. В крайнем случае «станцией назначения» служит линейное поле.

¹³Т. е. $f(x) = \lambda h$ при любом $h \neq 0$ имеет единственное решение $x \in \hat{\Omega}$.

12.4.7 Вращение линейного поля. Пусть $0 \in \Omega$ и A — линейное невырожденное преобразование. Тогда

$$\gamma(A, \dot{\Omega}) = \text{sign det } A, \quad (12.7)$$

причём $\gamma(A, \dot{\Omega}) = \text{ind}(A, 0)$.

◀ Если B шар с центром в нуле настолько малого радиуса, что $\overline{B} \subset \Omega$, то с учётом п. 12.4.6 имеем

$$\gamma(A, \dot{\Omega}) = \gamma(A, \dot{B}),$$

поскольку Ax невырождено на $\overline{\Omega} \setminus B$. Образ $A\dot{B}$ сферы \dot{B} также является сферой, вообще говоря, в другой норме. Следовательно, любой луч протыкает поверхность $A\dot{B}$ ровно один раз, откуда $|\gamma(A, \dot{\Omega})| = 1$. Знак вращения определяется ориентацией, что и даёт 12.7. ▶

12.5 Теоремы существования

12.5.1 Теорема. Пусть $0 \in \Omega$ и на границе, т. е. для любого $x \in \dot{\Omega}$ выполняется одно из условий:

- векторы x и $f(x)$ не направлены противоположно;
- $\|f(x) - x\| < \|f(x)\| + \|x\|$;
- скалярное произведение $(f(x), x) > 0$.

Тогда уравнение $f(x) = 0$ разрешимо в Ω .

◀ Из непротивоположной направленности x и $f(x)$ следует¹⁴

$$\gamma(f, \dot{\Omega}) = \gamma(f(x), \dot{\Omega}) = 1,$$

но тогда главный принцип существования нуля векторного поля (теорема 12.4.5) гарантирует разрешимость $f(x) = 0$. Два других условия влекут за собой непротивоположную направленность x и $f(x)$. ▶

Исследуемые уравнения часто имеют специальный вид

$$x = T(x). \quad (12.8)$$

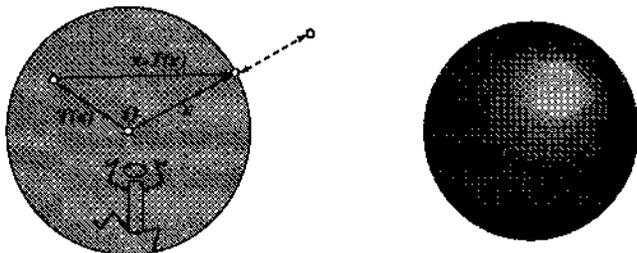
¹⁴См. теоремы 12.2.1, 12.3.2.

В этом случае решение (12.8) называют *неподвижной точкой* оператора T . Чтобы воспользоваться здесь результатом 12.5.1, достаточно перейти к векторному полю $f(x) = x - T(x)$. В итоге, например, можно гарантировать существование у T неподвижной точки, если

$$(Tx, x) < (x, x), \quad x \in \dot{\Omega}.$$

В юрисдикцию п. 12.5.1 попадает и *теорема Брауэра* (12.1.1). Её имеет смысл переформулировать с некоторым уточнением.

12.5.2 Теорема. *Оператор T , переводящий в себя замкнутый шар B , всегда имеет неподвижную точку $x^* \in B$. При этом $\gamma(I - T, \dot{B}) = 1$.*



◀ Без ограничения общности центр шара B можно считать расположенным в нуле. Допустим противное: T не имеет неподвижной точки в B . Тогда поле $x - T(x)$ невырождено на сфере \dot{B} и, в силу $T(B) \subset B$, векторы $x \in \dot{B}$ и $x - T(x)$ — непротивоположно направлены, иначе точка $T(x)$ лежала бы на продолжении радиус-вектора $x \in \dot{B}$, что противоречит $T(B) \subset B$. ▶

Из доказательства видно, что непротивоположная направленность x и $x - T(x)$ обеспечивается и в более свободных предположениях. Поэтому справедлив следующий результат.

12.5.3 Пусть Ω — произвольная область, $0 \in \Omega$, и оператор T на границе $\dot{\Omega}$ удовлетворяет условию $T(x) \neq \lambda x$ при $\lambda > 1$. Тогда у T существует неподвижная точка $x^* \in \bar{\Omega}$.

12.6 О теореме Брауэра

Хотя $\gamma(f, \hat{\Omega})$ может принимать любые целые значения, в практических задачах преобладает ситуация¹⁵

$$\gamma(f, \hat{\Omega}) = 1 \quad (12.9)$$

на областях Ω типа шара. Сюда попадает и теорема Брауэра 12.5.2. Обратим внимание на лёгкое, но существенное обобщение.

12.6.1 Шар B в теореме 12.5.2 можно заменить любым множеством A , гомеоморфным замкнутому шару.

◀ Действительно, пусть оператор f переводит в себя множество A , гомеоморфное замкнутому шару B , и G — соответствующий гомеоморфизм. В этом случае оператор $G^{-1}TG$ отображает в себя B и по теореме 12.5.2 имеет неподвижную точку $x^* \in B$, т. е.

$$G^{-1}TG(x^*) = x^* \Rightarrow TG(x^*) = G(x^*).$$

Следовательно, $y^* = G(x^*) \in A$ — неподвижная точка оператора T . ▶

Ключом к решению практических задач чаще всего оказывается либо сама теорема 12.5.2, либо утверждения, в которых теорема Брауэра выступает в роли скрытого механизма. Вот показательный пример.

12.6.2 Лемма Кнастера — Куратовского — Мазуркевича. Пусть X произвольное множество в \mathbb{R}^n , $A(x)$ — точечно-множественное отображение (из X в \mathbb{R}^n) с компактными образами. Пусть для всякого конечного подмножества $\{x_1, \dots, x_k\} \subset X$ выпуклая оболочка точек $\{x_1, \dots, x_k\}$ содержится в объединении $\bigcup_{i=1}^k A(x_i)$. Тогда пересечение всех множеств $A(x)$ не пусто,

$$\bigcap_{x \in X} A(x) \neq \emptyset.$$

◀ В силу компактности образов достаточно показать, что не пусто пересечение любого конечного множества образов. Предположим противное, т. е. $\bigcap_{i=1}^k A(x_i) = \emptyset$ для некоторого множества точек $\{x_1, \dots, x_k\}$. Тогда

$$\bigcup_{i=1}^k Y_i = \mathbb{R}^n, \text{ где } Y_i \text{ обозначает дополнение к } A(x_i).$$

¹⁵ Видимо потому, что «Бог изощрен, но не злонамерен».

Пусть теперь $\rho_i(x)$ — разбиение единицы в \mathbb{R}^n , согласованное с покрытием $\{Y_i\}$, т. е. такой набор функций, что¹⁶

$$\sum_{i=1}^k \rho_i(x) \equiv 1, \quad \text{supp } \rho_i \subset Y_i.$$

Отображение $\varphi(x) = \sum_{i=1}^k \rho_i(x)x_i$, очевидно, преобразует в себя выпуклую оболочку $c\{x_1, \dots, x_k\}$ и, по теореме Брауэра, имеет неподвижную точку x^* . При этом $\rho_i(x^*) > 0$ для i из некоторого подмножества индексов $M \subset \{1, \dots, k\}$, и $\rho_i(x^*) = 0$ для $i \notin M$. Но тогда

$$x^* = \sum_{i \in M} \rho_i(x^*)x_i \in \text{co}\{x_i : i \in M\} \subset \bigcup_{i \in M} A(x_i).$$

Следовательно, $x^* \in A(x_i)$ для некоторого $i \in M$, т. е. $x^* \notin Y_i$, а значит, $\rho_i(x^*) = 0$, что приводит к противоречию. ►

12.7 P -отображения

Круги по воде от теоремы Брауэра расходятся весьма далеко, и, так или иначе, многие результаты являются её следствиями. Вот своеобразная изюминка.

12.7.1 Теорема Карамардиана. Пусть непрерывный оператор $f(x) = \{f_1(x), \dots, f_n(x)\}$ определен при всех $x \geq 0$ и является P -отображением, т. е. для любого ненулевого вектора $x \geq 0$ можно указать такой индекс j , что

$$x_j \cdot f_j(x) > 0.$$

Тогда неравенство $f(x) > 0$ положительно разрешимо¹⁷.

Иными словами, уравнение $y = f(x)$ при некотором $y > 0$ имеет решение $x \geq 0$. Вкус и аромат скучному на вид утверждению придаёт следующая содержательная интерпретация.

Назовем переменные x_i действиями, y_i — результатами. Например, x_i — количество i -го лекарства, y_i — i -й показатель

¹⁶Здесь $\text{supp } \rho$ обозначает носитель функции $\rho(x)$, т. е. множество тех x , для которых $\rho(x) > 0$.

¹⁷Под $f(x) > 0$ подразумевается $f_i(x) > 0$ для всех $i = 1, \dots, n$.

здоровья (температура, давление). Одинаковые номера за действиями закрепляют «свой» результаты. Понятно, что при наличии побочных влияний совместные стереотипные действия могут приводить к отрицательным последствиям, в результате чего излечение больного может оказаться невозможным.

Однако пусть связи $y_i = f_i(x)$ определяются P -отображением, т. е. любая совокупность положительных действий не может дать отрицательные результаты по всем номерам, для которых $x_i > 0$. Хотя бы один результат будет ожидаемым. Тогда теорема 12.7.1 гарантирует возможность так подобрать действия (лекарства) $x_i > 0$, что все результаты будут строго положительными (все показатели улучшатся, несмотря на побочные влияния)¹⁸.

◀ Доказательство теоремы 12.7.1. Применим лемму 12.6.2, которая, заметим, остается справедливой при замене предположения о компактности образов $A(x)$ требованием их открытости и ограниченности. В качестве X возьмём n вершин симплекса $\Delta = \{x : \|x\| = 1, x \geq 0\}$:

$$x_i = \{0, \dots, 0, 1, 0, \dots, 0\},$$

и положим

$$A(x_i) = \{x : f_i(x) > 0, x \in \Delta\}.$$

Из определения P -отображения следует, что выпуклая оболочка любого подмножества вершин симплекса Δ покрывается соответствующим объединением $\bigcup A(x_i)$, что позволяет (лемма 12.6.2) гарантировать непустоту пересечения всех $A(x_i)$. ▶

12.8 Алгебраическое число нулей

Нуль $x_0 \in \Omega$ векторного поля $f(x)$ называют *изолированным*, если в достаточно малой окрестности x_0 нет других нулей. Вращение поля f на сферах достаточно малого радиуса с центром в x_0 называют *индексом* x_0 и обозначают $\text{ind}(f, x_0)$.

Корректность такого определения вытекает из одинаковости вращений f на сферах достаточно малого радиуса. Действительно, если шары B_1 и B_2 не содержат других нулей, кроме x_0 , и $B_1 \subset B_2$, то из теоремы 12.4.6 следует

$$\gamma(f, \dot{B}_1) = \gamma(f, \dot{B}_2).$$

¹⁸Об универсальных P -отображениях см. Опойцев В.И. Нелинейная статистика. М.: Наука, 1986.

12.8.1 Теорема об алгебраическом числе нулей. Пусть векторное поле $f(x)$ невырождено на $\dot{\Omega}$ и имеет в Ω лишь изолированные нули x_j . Тогда

$$\gamma(f, \dot{\Omega}) = \sum_j \text{ind}(f, x_j), \quad (12.10)$$

где суммирование идет по всем нулям $x_j \in \Omega$ поля f .

◀ Доказательство сразу вытекает из п. 12.4.6, если области Ω_j определить как шары достаточно малых радиусов с центрами в точках x_j . ▶

Теорема 12.8.1 имеет разнообразные применения. Гарантирует, например, существование неизвестного нуля, если сумма индексов известных нулей отлична от $\gamma(f, \dot{\Omega})$. В различных вариантах позволяет также устанавливать единственность решения. Скажем, пусть $\gamma(f, \dot{\Omega}) = 1$, и производная (матрица Якоби) $f'(x)$ невырождена на Ω . Этого, оказывается, достаточно для существования и *единственности* решения у $f(x) = 0$ на Ω .

◀ Действительно, если x_0 — нуль поля f , то $f(x)$ и $f'(x_0)(x - x_0)$ противоположно направлены в достаточно малой окрестности x_0 , ибо

$$f(x) = f'(x_0)(x - x_0) + o(\|x - x_0\|).$$

В силу невырожденности $f'(x)$ детерминант $\det f'(x)$ сохраняет знак, поэтому индексы всех нулей обязаны быть одинаковы,

$$\text{ind}(f, x_0) = \text{sign det } f'(x_0) = 1.$$

А это, благодаря условию $\gamma(f, \dot{\Omega}) = 1$, допускает не более одного нуля. ▶

12.9 Индексы на бесконечности

Если поле f невырождено при больших по норме $x \in \mathbb{R}^n$, то говорят, что *особая точка* ∞ поля f изолирована. При этом теорема 12.4.1 и лемма 12.4.4 гарантируют, что вращение $f(x)$ на сферах достаточно большого радиуса одно и то же. Это общее вращение называют *индексом на бесконечности* и обозначают $\text{ind}(f, \infty)$.

Если поле $f(x)$ имеет лишь изолированные нули x_j и определен индекс f на бесконечности (т. е. $f(x)$ невырождено при достаточно больших по норме $x \in \mathbb{R}^n$), то число нулей x_j конечно, и в силу п. 12.4.6 имеет место равенство

$$\sum_j \text{ind}(f, x_j) = \text{ind}(f, \infty).$$

Допустим, для некоторой матрицы $f'(\infty)$

$$\lim_{\|x\| \rightarrow \infty} \frac{\|f(x) - f'(\infty)x\|}{\|x\|} = 0. \quad (12.11)$$

В этом случае поле f называют *асимптотически линейным*, а матрицу $f'(\infty)$ — *производной на бесконечности* оператора f .

12.9.1 Пусть поле f асимптотически линейно и $\det f'(\infty) \neq 0$. Тогда особая точка ∞ поля f изолирована и

$$\text{ind}(f, \infty) = \text{sign det } f'(\infty).$$

◀ Доказательство совсем просто. По условию $\|f(x) - f'(\infty)x\| = o(\|x\|)$. Но так как матрица $f'(\infty)$ невырождена, это влечёт за собой справедливость неравенств

$$\|f(x) - f'(\infty)x\| \leq \|f'(\infty)x\|$$

на сферах достаточно больших радиусов, что гарантирует непротивоположную направленность полей $f(x)$ и $f'(\infty)x$. ▶

12.10 Накрытия и гомеоморфизмы

Как уже отмечалось в разделе 6.8, для *глобальной обратимости отображения* F к его локальной обратимости достаточно добавить условие

$$\lim_{\|x\| \rightarrow \infty} \|F(x)\| = \infty. \quad (12.12)$$

Странно, что этот результат остаётся за пределами общего математического образования. Странно по нескольким причинам. С одной стороны, сей факт важен и полезен, ибо освещает немало прикладных задач. При этом достаточно прост, существенное обобщение достигается малой кровью. Наконец, результат несёт на себе немалый педагогический потенциал, закладывая ценные стереотипы мышления.

12.10.1 Теорема. Для глобальной обратимости непрерывного отображения F , действующего в \mathbb{R}^n , необходимо и достаточно выполнение двух условий: (12.12) и локальной обратимости F в любой точке \mathbb{R}^n .

◀ *Необходимость* очевидна. *Достаточность*:
Во-первых, в силу локальной обратимости F функция

$$\text{ind}[F(x) - F(u), u]$$

локально постоянна в достаточно малой окрестности любой точки $u \in \mathbb{R}^n$. Переход из u в любую другую точку v можно представить в виде последовательности сколь угодно малых шагов вдоль отрезка

$$\lambda u + (1 - \lambda)v, \quad \lambda \in [0, 1].$$

Поэтому функция $\text{ind}[F(x) - F(u), u]$ постоянна в \mathbb{R}^n .

Далее. В силу (12.12) деформация

$$H(x, \tau) = F(x) - \tau F(0) - (1 - \tau)u$$

невыврождена на сферах S достаточно большого радиуса, при любом $u \in S$. Поэтому вращения полей $F(x) - F(0)$ и $F(x) - u$ на S равны при любом $u \in S$. А так как все индексы одинаковы, то по теореме 12.8.1 об алгебраическом числе нулей число решений уравнения $F(x) = u$ как функция u есть константа, отличная от нуля.

Допустим, что $F(x) = u$ имеет $m > 1$ решений. Пусть, например, $u = 0$ соответствует решению x_1, \dots, x_m . При движении по любому лучу vt ($v \in \mathbb{R}^n$, $t \in (0, \infty)$) из каждой точки x_i выходит непрерывная кривая $x_i(t)$ — такая, что $F(x_i(t)) = vt$. В силу локальной обратимости F кривые $x_i(t)$ не могут пересекаться.

Отнесем теперь к множеству C_i все точки кривых $x_i(t)$, соответствующие всевозможным v и t . В силу той же локальной обратимости F множества C_i открыты и не пересекаются. Но $\bigcup C_i = \mathbb{R}^n$. Это противоречит связности \mathbb{R}^n . ▶

В более общей ситуации речь идет о существовании функции $x = G(y)$, неявно задаваемой уравнением

$$\Phi(x, y) = 0.$$

Здесь сохраняется полная аналогия с предыдущим случаем.

12.10.2 Теорема¹⁹. Для глобальной разрешимости $\Phi(x, y) = 0$ относительно x необходимо и достаточно выполнение двух условий:

- Уравнение $\Phi(x, y) = 0$ локально разрешимо.
- $\|x_k\| \rightarrow \infty \Rightarrow \|y_k\| \rightarrow \infty$, если $\Phi(x_k, y_k) = 0$. ▽

Поскольку никогда не знаешь, что где и когда выстрелит, вот несколько опорных вех. Для тех, кто изволит добровольно бросить взгляд.

○ Пусть $f : X \rightarrow Y$, где X, Y линейно связны, причём любая замкнутая кривая в Y может быть стянута в точку. Тогда чтобы f было гомеоморфизмом X на Y необходимо и достаточно выполнение двух условий:

¹⁹ Доказательство более общего утверждения есть в [2, т. 13].

- (а) f является локальным гомеоморфизмом;
 (б) прообраз любого компакта из Y компактен в X . \triangleright

○ Пусть $\text{ind}(f, \infty) \neq 0$, и выполняется (12.12). Тогда уравнение $f(x) = y$ разрешимо при любом $y \in \mathbb{R}^n$. \triangleleft В силу (12.12) на сферах достаточно большого радиуса поля $f(x)$ и $f(x) - y$ непротивоположно направлены. \blacktriangleright

○ Пусть $f(x)$ — локальный гомеоморфизм и образ $f(\mathbb{R}^n)$ замкнут в \mathbb{R}^n . Тогда f покрывает \mathbb{R}^n . \triangleright

○ Пусть локальный гомеоморфизм $G: \text{int}\mathbb{R}_+^n \rightarrow \text{int}\mathbb{R}_+^n$ удовлетворяет условию²⁰

$$\|LG(x)\| \rightarrow \infty \text{ при } \|L(x)\| \rightarrow \infty.$$

Тогда G — гомеоморфизм $\text{int}\mathbb{R}_+^n$ на $\text{int}\mathbb{R}_+^n$. \triangleright

12.11 Параметрические уравнения

На практике нередко встречаются уравнения

$$G(x, \lambda) = 0 \tag{12.13}$$

со скалярным параметром λ . Здесь теория вращения работает в сочетании с дополнительными ухищрениями.

12.11.1 Пусть

$$\gamma[G(x, \lambda_1), \dot{\Omega}] \neq \gamma[G(x, \lambda_2), \dot{\Omega}].$$

Тогда при некотором $\lambda \in (\lambda_1, \lambda_2)$ на $\dot{\Omega}$ существует решение уравнения (12.13).

\triangleleft В противном случае поля $G(x, \lambda_1)$ и $G(x, \lambda_2)$ соединял бы невырожденный гомотопический мост

$$H(x, t) = G[x, t\lambda_1 + (1-t)\lambda_2]. \quad \blacktriangleright$$

Ещё одна популярная идея — функционализация параметра. В (12.13) вместо λ подставляют некоторый функционал $\lambda(x)$, в результате чего $G(x, \lambda) = 0$ переходит в

$$G(x, \lambda(x)) = F(x) = 0,$$

²⁰ Где $L(x) = \{\ln x_1, \dots, \ln x_n\}$.

и если вращение $\gamma(F, \dot{\Omega})$ оказывается ненулевым, то (12.13) имеет решение $x^* \in \Omega$ при $\lambda^* = \lambda(x^*)$.

От идеи до исполнения путь, конечно, неблизок. По шучьему велению не получается, потому что выбор функционала $\lambda(x)$ — самостоятельная проблема. Но она все же менее сложна и более конкретна чем расплывчатая задача (12.13), не привязанная ни к какому конструктивному замыслу.

Когда уравнение с параметром имеет специальный вид

$$f(x) = \lambda x,$$

его решения, по аналогии с линейным случаем, называют *собственными векторами оператора f* .

12.11.2 Лемма. В нечётномерном пространстве в случае $0 \in \Omega$ невырожденный на $\dot{\Omega}$ оператор f всегда имеет на $\dot{\Omega}$ собственный вектор, т. е. $f(x) = \lambda x$ при некоторых $\lambda \neq 0$ и $x \in \dot{\Omega}$.

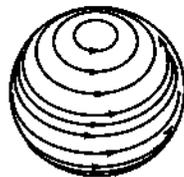
◀ Если $\gamma(f, \dot{\Omega}) \neq 1$, то поле f не гомотопно тождественному и, следовательно, при некоторых $x \in \dot{\Omega}$ и $t \in (0, 1)$

$$tf(x) + (1-t)x = 0, \quad \text{т. е.} \quad f(x) = \frac{t-1}{t}x.$$

Если же $\gamma(f, \dot{\Omega}) = 1$, то поле f не гомотопно полю $-I(x)$, и тогда при некоторых $x \in \dot{\Omega}$ и $t \in (0, 1)$

$$tf(x) - (1-t)x = 0, \quad \text{т. е.} \quad f(x) = \frac{1-t}{t}x. \quad \blacktriangleright$$

12.11.3 В нечетномерном пространстве сферу невозможно причесать (без макушки), т. е. задать на сфере невырожденное касательное поле.



◀ Это сразу становится ясным, если центр сферы расположить в нуле и применить результат 12.11.2 к касательному полю $f(x)$. ▶

Глава 13

Проблемы обучения¹

*Час, затраченный на понимание,
экономит год жизни.*



Учить математику на 90% надо так же, как остальное. Ибо успех обеспечивают общие принципы, а 10% приходится на математическую специфику. Но «принципы» системой образования игнорируются. Школа учит английскому 10–15 лет — результаты катастрофические. Тогда как Миша Цацкин женился на китаянке, прожил год в большой китайской семье, и теперь болтает на китайском как на родном.

С математикой — такая же история. Основная беда в том, что знание в препарированном виде пытаются протолкнуть через бутылочное горлышко подсознания. Не пролезает. Нутро человека заточено под естественные способы обучения. А *система* придерживается искусственного оплодотворения.



¹ О проблемах обучения см. oschool.ru

13.1 Кто мы есть и как мы учимся

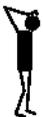
Любая аудитория делится на три категории: *A*, *B* и *C*. Первые скватывают на лету, вторые переосмысливают², третьи ничего не понимают, как ни объясняй.

Сразу важно оговориться, все группы одинаково ценны для мировой цивилизации. Не по причине лицемерной позиции, а по сути. Потому что в условиях, скажем, землетрясения или рисования — они проявляются иначе. Не говоря о том, что гениальные решения чаще рождаются в недрах *C*.



Конечно, мы баланс немного нарушаем в пользу *C*, но это для компенсации их невыгодной роли в школьные годы. Плюс к тому, выразив восхищение, спокойно можно писать, что они — не понимают. Теперь в этом нет ничего обидного. Из них потом такие Менделеевы получаются, что группе *A* остается сожалеть, что скватывала на лету.

Однако разговор сейчас о другом. Группа *C* ничего не понимает, и не поймет — но Система смириться с этим не хочет. В результате процесс затягивается до тошноты. Остановиться нельзя, ибо *C* ещё не поняли, а *B* и *A* перестали понимать — ибо когда скобки в $(x - y)(x + y)$ раскрываются полгода, ум начинает заходить за разум.



*Выход из положения кажется тривиальным. Надо объяснить коротко и просто. Ориентироваться на *A* и *B*. Группе *C* хуже не будет. Да и перед *A + B* не надо слишком расширяться. Тогда процесс вернётся в естественное русло. Нормализуются акценты, пропорции, интонация.*

Конечно, вроде бы возникает противоречие. Если «коротко и просто», то чем наполнять занятия? Понятно же, что общее время обучения сокращать нельзя. Инкубационный процесс, так сказать. Набивший оскомину пример девятимесячного вынашивания плода здесь как раз к месту.

²То есть сначала понимают частично или неправильно. Потом возвращаются назад и переосмысливают. И далее по кругу.

Но «коротко и просто» легко вписывается в инкубационное ожидание. Масса времени уходит на перечисление. Быстрое скватывание на деле не так замечательно, как выглядит. Выясняется, что быстро — не обязательно правильно. Да ещё выветривается. Поэтому скватывание на лету не означает завершения процесса. В группе *B* положение ещё хуже. Переосмысливание не обходится без ошибок и цементирует заблуждения. Поэтому объяснять приходится снова и снова.



Получается в итоге сказка про белого бычка. Группу *C* перестали принимать в расчёт, но потребность в объяснениях едва ли не увеличилась. Однако объяснения теперь нужны совсем другие. Не для заполнения отведённого времени, а для получения результата. Внешне всё по-прежнему переливается из пустого в порожнее, но знание растёт и крепнет под мягким корректирующим воздействием.

Если кому-то кажется, что в *Школе/Университете* так и есть — думайте дальше. На словах похоже, но о наличии пропасти легко судить по результатам «на стороне», которые имеют очень простое объяснение. Процесс вне *Школы* вытаскивается из ловушки «одни делают вид, что учат, другие — что учатся» и ставится на нормальные рельсы. Объяснения, решение задач, комментарии — все делается с единственной целью достижения результата. Поэтому объяснения, не меняя своего имени, меняют суть. Исчезает направленность «в никуда». Без науги, вскользь, не преследуя цель сразу достичь максимального эффекта, не выходя за пределы 5–10 минут, выделяя главное и оставляя возможность неправильно понять детали, возвращаясь раз за разом назад, но в другой аранжировке, в лёгкой и даже в легкой форме — такие объяснения несколько отличаются от того, что могло бы одобрить министерство образования. Но они дают результат! И было бы неплохо поставить на те же рельсы образование целиком.



Теперь посмотрим на *A, B, C* изнутри, дабы понять, кто мы есть, или кого мы учим. Представители *A, B, C* разбросаны в диапазоне, о широте которого можно судить по крайним точкам.

• Слева пребывает тот, кого стандартными методами учить невозможно. В чистом виде — это уникал, гений. Ни одну строчку

он не может дочитать до конца, потому что спотыкается на каждом слове. Слова завораживают и уносят в другой мир. Верблюдов до сих пор не видел, хотя пять раз бывал в зоопарке. Но так вышло, что всё внимание поглощали воробьи на взводе. Фокус его интереса цепляется за любую деталь, и воображение уносит неведомо куда, с любой стартовой площадки.

• На правом краю устроился совсем другой индивид. Ему не до подробностей. Взор устремлён за горизонт. Мелкие «почему» его не интересуют, детали не отвлекают. Тратит энергию и время готов только по большому счёту.

Второго учить традиционными методами очень легко. Вы ещё не успели объяснить, как он призывает двигаться дальше. Вот уравнение, вот решение. Что ещё объяснять?

А как первого из группы С можно учить? Никак³. Вдумайтесь. Что посоветовать можно? — Тому, кто учит, — смиритесь и будьте деликатны. — Тому, кто учится, — не расстраивайтесь, такова жизнь. Живите, как получается.

А ля гер ком алягер, что поделаешь. Проблему усугубляет русский менталитет. Запад стремится понять КАК, мы всегда хотим докопаться ПОЧЕМУ. Как бы там ни было, но и системы образования наши так ориентированы. А учить умению КАК, между прочим, гораздо проще, чем разумению ПОЧЕМУ. Бесконечно проще. Потому что у КАК всегда есть исчерпывающий ответ⁴, а ПОЧЕМУ уводит в бездонную пропасть. Одно цепляется за другое, и нет этому конца. Отсюда и томление русской души, и необыкновенные взлёты на фоне мучительных провалов русского образования. Тактически иногда целесообразнее начинать с вопроса КАК. Запад предпочитает этим и заканчивать, тщательно заделывая брешь в пространство ПОЧЕМУ⁵. Мы же расширяем эту брешь и культивируем стремление туда. В результате часть населения не удаётся вернуть обратно. Зато другая — вдохновлена и не может успокоиться.

³Разве что индивидуально, подыгрывая и аккомпанируя. Утопия.

⁴В одном из вариантов: НИКАК, — т. е. нет способа.

⁵Известно: кто знает КАК — исполняет, кто знает ПОЧЕМУ — руководит. Но пусть, думает Запад, «почему-калки» пробивают себе дорогу сами. Иначе с остальными — такая морока. Как у нас.

13.2 О взаимодействии с подсознанием

Проблема взаимодействия сознания с подсознанием в обучении краеугольная (см. oschool.ru). Слабое звено здесь в том, что осознав что-либо «математическое» умом, человек останавливает процесс. И *Бегемотик*⁶, который учится медленнее, остаётся необученным. В других областях источником у него служат чувства. А тут формулы — *без сопровождения разумом* теряют смысл, превращаются в информационный мусор. Поэтому выход из положения остаётся один: *продолжать процесс обдумывания, воображения, проигрывания в голове.*

Затея, конечно, выглядит по-дурацки. Вы уже всё поняли, но продолжаете переливать из пустого в порожнее. Однако процедуру можно организовать иначе. Персонифицируйте подсознание и начинайте обучать его, разыгрывая из себя учителя. Процесс невероятно полезный. Если бы *Бегемотик* не существовал, его бы стоило выдумать. Ибо лучший способ изучить предмет — научить другого. И подсознание — идеальный слушатель, который всегда под рукой. Разумеется, исполнять сие желательно с лёгкой иронией, дабы компенсировать иногда возникающее ощущение идиотизма. Но будьте уверены, что освоив такой внутренний диалог, вы добьётесь больших успехов не только в математике, но и в других сферах бытия.



Обучая подсознание, необходимо стремиться к простоте, в поисках которой приходится сталкиваться с определёнными сложностями. Потому что сообщества профессионалов образуют анклав, куда не хотят пускать посторонних. Недаром Бернард Шоу писал, что *всякая профессия — это заговор против непосвящённых.*

Экономисты отгораживаются от толпы заумными теориями. А хирурги в не столь давние времена принимали в пытки обезболивающие препараты, считая, что при обезболивании любой дурак сможет делать операции. Математики тоже не лыком шиты. Ставя барьеры строгости обоснования результатов, они не только решают благородную задачу очищения теории от разной шелухи и создания надёжного фундамента, но и ставят палки в колеса новичкам, которым бы хотелось понять

⁶Или Крокодильчик, или Змей Горыныч, или Мапа/Сапа. Подсознание мы называем *Бегемотиком* в качестве общего знаменателя.

сначала кое-что «на пальцах». Но не тут-то было. Вспоминается жена, которая спрашивала мужа, накрыть стол для гостей так, чтобы ещё пришли, или чтобы больше не приходили? Как математики «стол накрывают» сами знаете.

О других особенностях воспитания *Бегемотика* см. oschool.ru.

13.3 Гипнотический вирус, будь он неладен

Анатомия непонимания не так проста. Конечно, если историю России можно изучать, не зная, что натворили греки до нашей эры, то даже за арифметику, не владея ремеслом раскрытия скобок, нечего браться. В математике всё жутко переплетено. И если что упущено — так оно потом аukaется, не переставая.

Ещё хуже другое. Ничего не упущено — а непонимание как та ложка дёгтя. Всё выучено, усвоено до автоматизма. Но какая-то заноза остаётся. И ноет, и болит, и какая-то пелена застилает мозги.

Жизнь превращается в каторгу и сводится к написанию шаргалок, которые — очень ненадёжная вещь. «Шаг вправо, шаг влево» — и вы в ступоре, и вас экзаменатор берёт за фук. Не говоря о том, что любую новую тему вы всё равно не понимаете, ибо для понимания нужны инструменты в голове, а не писульки в рукаве. Вызубренное не лучше шаргалки. Ибо «шаг вправо, шаг влево» — и вы опять в ступоре. Потому что знаете КАК, но не знаете ПОЧЕМУ.



Разумеется, изучая предмет, надо ходить кругами. Много раз по одному и тому же месту. Вы представьте, что переехали жить в новый город, и хотите узнать его. Вам достаточно будет проехать по территории? Конечно, нет. Город надо исходить вдоль и поперёк. Пешком и на трамвае, и днём и ночью, и в дождь и в снег, и в праздник и в будни. Поговорить с людьми. И порыскать «в поисках», и проклясть его лабиринты и тупики. Чтобы, как говорится, полюбить на целый век. И тогда лишь, однажды выходя из дома, вы увидите вдруг *Знакомый город*. С математикой такая же песня — и хождение кругами необходимо. Но без понимания диагноза «лечение» идёт всё-таки не

так успешно, как хотелось бы. А одна из главных причин заключена в действии «гипнотического вируса».



Гипноз, как принято считать, основан на проходе в глубины подсознания обманым или силовым путём⁷. И чем легче человек поддаётся гипнозу, тем труднее ему даётся математика, в которой слишком много механизмов, заклинивающих сознание и действующих по типу *тройной спирали Эриксона*⁸. Спираль Эриксона — это хитрый и вместе с тем очень простой трюк.

Рассказывается некая история, которая в середине обрывается, и начинает рассказываться вторая история, которая снова не доводится до конца, и повествование переключается на третью историю. Сознание вынуждено держать в памяти все эти половинчатые истории — и у него оказываются «заняты руки». Охрана снята, дорога к подсознанию свободна, слушатель в трансе⁹. В математике нечто подобное происходит постоянно. В результате многие попадают в состояние транса задолго до того, как то или иное рассуждение услышано до конца.

Состояние транса — это состояние, в котором мы все в той или иной степени пребываем. Это не гипнотический сон с каталептической атрибутикой, а бодрствование с некоторой заторможенностью сознания. Если заторможенность превышает определённый порог, адекватность поведения и восприятия нарушается. Это и происходит часто при изучении математики, что препятствует пониманию происходящего.

Например, доказывая что-то, мы предполагаем противное и ведём рассуждение к противоречию. Но о «предположении противного» надо помнить, что мешает рассуждению, загружая часть памяти, которую хорошо было бы использовать для другой надобности. Аналогично действует гипнотизёр, отвлекая ресурсы нашего мозга на какую-либо

⁷ Гипнотизёр действует обычно, маскируясь, пугая или отвлекая сознание, играющее роль щита.

⁸ Эриксон — знаменитый американский гипнотизёр.

⁹ И это не сказка, а психологический приём, простой как молоток и эффективный, как противозаконный «двадцать пятый кадр». Примерно так, кстати, действуют некоторые компьютерные вирусы. Сначала переполняют память компьютера, а потом с парализованным прибором делают, что хотят.

постороннюю работу, типа бессмысленного счёта, а сам втягивает в подсознание какую-нибудь «пейте кока колу».

Яркий пример: определение непрерывной функции. Напомним.

Функция $f(x)$ непрерывна в точке x_0 , если по любому $\varepsilon > 0$ можно указать такое δ , что $|f(x) - f(x_0)| < \varepsilon$ как только¹⁰ $|x - x_0| < \delta$.



Две строчки определения, а не даются. В голову не помещаются. В ту её часть, которая как раз предназначена для одновременного слежения за несколькими факторами. Двух-трёх обычно оказывается чересчур много.

Другой пример. Известная задача, в которой математика на уровне детского сада. *В мешок положили 3 чёрных колпака и 2 белых. Трёх участникам игры завязали глаза, и надели им на головы 3 чёрных колпака. Повязки с глаз сняли. Как они быстро догадались, какого цвета колпаки у них на головах?*



Обозначим игроков А, В, С. Каждый, например А, рассуждает так. Если на мне белый колпак, то В, видя перед собой белый и чёрный, легко понял бы, что на нём чёрный. Поскольку в противном случае С мгновенно бы сообразил, видя перед собой два — белых, что на нём чёрный колпак. Но С мгновенно не реагирует. Значит, на мне чёрный, — решил бы В. Но он пока молчит. Значит, на мне чёрный, — решает А.

Рассуждение плохо помещается в голову, и его многие даже не могут воспроизвести. Не любит наш интеллектуальный процессор удерживать одновременно несколько деталей. И его надо тренировать.

Это, пожалуй, единственный способ для развития и увеличения своего ментального пространства, чтобы оно было способно вмещать длинные цепочки рассуждений. Говоря условно, на языке компьютерных аналогий, необходимо наращивать свою «оперативную память» и учиться блокировать мысленные посторонние коловращения, закливающие мозговой процессор. Всё это поддаётся тренировке, но требует упорства и терпения.

¹⁰ Сначала с потолка появилось ε , для чего неясно, но о нём надо помнить, чтобы потом понять конец истории. Затем при аналогичных обстоятельствах возникло δ . В конце концов обе истории заканчиваются, но отключенное сознание уже ничего не слышит.

Обозначения

◀ и ▶ — начало и конец рассуждения, темы, доказательства

▷ — утверждение приводится без доказательства

◁ — утверждение легко может быть доказано

(?) — предлагает проверить или доказать утверждение в качестве упражнения, либо довести рассуждение до «логической точки»

(!) — предлагает обратить внимание

«в томм случае» — «в том и только том случае»

п. — пункт либо раздел

$A \Rightarrow B$, или $A \rightarrow B$ — из A следует B

$x \in X$ — x принадлежит X

$X \cup Y$, $X \cap Y$, $X \setminus Y$ — объединение, пересечение и разность множеств

$X \subset Y$ — X подмножество Y , в том числе имеется в виду возможность $X \subseteq Y$, т. е. между $X \subset Y$ и $X \subseteq Y$ различия не делается

\sim — отношение эквивалентности, определяемое контекстом

\emptyset — пустое множество

$\bar{\Omega}$ — замыкание Ω

$\dot{\Omega} = \partial\Omega$ — граница Ω

\mathbb{N} — множество натуральных чисел $\{1, 2, \dots\}$

\mathbb{Q} — множество рациональных чисел (дробей)

$\mathbb{R} = (-\infty, \infty)$ — вещественная прямая

\mathbb{R}^n — n -мерное евклидово пространство

\mathbb{C} — комплексная плоскость

\exists — квантор существования

\forall — квантор общности

S^{n-1} — единичная сфера в \mathbb{R}^n

I — тождественное отображение, $I(x) \equiv x$. В \mathbb{R}^n

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

\mathbb{R}_+^n — неотрицательный ортант

i — мнимая единица, $i^2 = -1$

$z = x + iy$ — комплексное число (КЧ),

$z = r(\cos \varphi + i \sin \varphi)$ — его тригонометрическая запись

$\mathbf{x} = \{x_1, \dots, x_n\}$ — вектор, x_i — его координаты

$\mathbf{x} \mathbf{y}$, $\mathbf{x} \times \mathbf{y}$ — соответственно скалярное и векторное произведение векторов \mathbf{x} и \mathbf{y} ; для скалярного произведения используется также эквивалентное обозначение (\mathbf{x}, \mathbf{y}) либо $\langle \mathbf{x}, \mathbf{y} \rangle$

$\frac{df(\mathbf{x})}{dx}$ — производная (скорость изменения по \mathbf{x}) функции $f(\mathbf{x})$ в точке \mathbf{x} , эквивалентное обозначение: $f'(\mathbf{x})$

для производной по времени вместо $x'(t)$ чаще используется \dot{x} , а для второй производной \ddot{x}

$\frac{\partial u}{\partial x}$ — частная производная функции u по переменной x

$\nabla f(\mathbf{x})$ — градиент функции $f(\mathbf{x})$, т.е. вектор

$$\left\{ \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right\},$$

который направлен по нормали к поверхности постоянного уровня функции $f(\mathbf{x})$ и численно равен скорости максимального роста $f(\mathbf{x})$ в точке \mathbf{x}

Литература

- [1] *Босс В.* Интуиция и математика. Изд. 5-е, М.: УРСС, 2015.
- [2] *Босс В.* Лекции по математике. Т. 1 — Анализ. Т. 2 — Дифференциальные уравнения. Т. 3 — Линейная алгебра. Т. 4 — Вероятность, информация, статистика. Т. 5 — Функциональный анализ. Т. 6 — Алгоритмы, логика, вычислимость. От Диофанта до Тьюринга и Гёделя. Т. 7 — Оптимизация. Т. 8. — Теория групп. Т. 9 — ТФКП. Т. 10 — Перебор и эффективные алгоритмы. Т. 11 — Уравнения математической физики. Т. 12 — Контрпримеры и парадоксы. Т. 13 — Топология. Т. 14 — Теория чисел. Т. 15 — Нелинейные операторы и неподвижные точки. Т. 16 — Теория множеств: от Кантора до Коэна М.: Книжный дом «Либраком»/URSS, 2004–2015.
- [3] *Гелбаум Б., Олмстед Дж.* Контрпримеры в анализе. М.: Мир, 1967.
- [4] *Колмогоров А. Н., Фомин С. В.* Элементы теории функций и функционального анализа. М.: УРСС, 2006.
- [5] *Кочин Н. Е.* Векторное исчисление и начала тензорного исчисления. М.: Наука, 1965.
- [6] *Постон Т., Стюарт И.* Теория катастроф и ее приложения. М., 1980.
- [7] *Федорюк М. В.* Обыкновенные дифференциальные уравнения. М., 2003.
- [8] *Фихтенгольц Г. М.* Курс дифференциального и интегрального исчисления в 3-х томах. М., 2015.

Предметный указатель

P -отображение 246
 ε -окрестность 25

Абсолютно сходящийся ряд 46
активные ограничения 182
амплитуда 103
аналитическое продолжение 205
ахиллесово направление 241

Базис 130, 153, 214
бесконечный предел 28
бесконечный ряд 44
бесконечно большая величина 28
бесконечно малая величина 28
бином Ньютона 20

Вырожденный случай 177
варианта 45
вектор аксиальный 159
вектор полярный 159
векторы коллинеарные 130, 153
векторы компланарные 130, 153
векторное поле 234
векторное произведение 157
включение 23
включение строгое 23
вращение векторного поля 236
вращение линейного поля 243

Гармоническая функция 195
гармонический ряд 44

гармоническое колебание 221
гессиан 175
гиперплоскость 127, 157
главная диагональ 128
гомеоморфизм 131
гомеоморфизм локальный 131
гомеоморфные пространства 131
гомотопический переход 237
гомотопные отображения 238
градиент 114
график функции 25

Двойственная задача 185
декартово произведение 24
дельта-функция 227, 229
детерминант 132
деформация 237
дивергенция 164
дифференцирование 52
дифференцирование оператора 136
длина дуги 86
дополнение 24
допустимое решение 182

Евклидово пространство 106

Замкнутое множество 25

- Игровая площадка** 9
 инъекция 130
 инвариантность формы дифференциала 58
 индекс на бесконечности 248
 индекс нуля 247
 интеграл 75
 интеграл Фурье 222
 интеграл Лебега 216
 интеграл Римана 78
 интеграл двойной 142
 интеграл кратный 146
 интеграл неопределённый 75
 интеграл несобственный 92
 интеграл определённый 77
 интегральная формула Коши 200
 интегрирование 75
 интегрирование по частям 83
 инфимум 39
- Канторова лестница** 71
 канторово множество 71
 касательная плоскость 115
 квадратируемая область 142
 коэффициент искажения объёма 132
 компактность 42
 комплексная плоскость 13
 комплексное число 13
 контурный интеграл 170
 конформное преобразование 194
 координаты 151, 153
 кривая Пеано 70
 криволинейный интеграл 168, 169
 критерий Коши 43
 критическая точка 175
 критическая точка невырожденная 176
 круговая частота 103
- Лагранжиан** 178
 лемма Больцано — Вейерштрасса 40
 лемма Гейне — Бореля 41
 лемма Кнастера — Куратовского — Мазуркевича 245
 лемма Вейерштрасса 30, 40
 лемма о вложенных отрезках 40
 линейный оператор 127
 линейная гомотопия 239
 линейная независимость 130, 153
 линейно зависимые векторы 129, 153
 логарифмический вычет 208
- Мажорирующий ряд** 46
 матрица 128
 матрица Якоби 136
 матрица кососимметрическая 129
 матрица невырожденная 132
 матрица симметрическая 129
 матрица транспонированная 129
 множитель Лагранжа 178
 момент инерции 161
 момент количества движения 160
 момент силы 160
 морсовское l -седло 176
- Натуральный ряд** 9
 неявная функция 139, 250
 невырожденная деформация 238
 нелинейное программирование 182
 неподвижная точка 137, 231
 непрерывная функция 35, 109
 непротивоположная направленность 239
 неравенство Коши — Буняковского 107

норма вектора 106, 135, 156
 норма евклидова 106
 норма кубическая 133
 норма матрицы 135
 норма октаэдрическая 133
 норма подчинённая 135
 носитель функции 228
 нуль векторного поля 234, 242, 247

О-малое 56

обезьянье седло 176
 обобщенная функция 228
 образ 130
 обратимость отображения 249
 обратная функция 138
 оператор 126
 оператор «набла» 167
 оператор Гамильтона 167
 оператор Лапласа 168
 определитель 132
 орт 151
 ортогональный базис 214
 ортогональность 107
 особые точки 207
 особая точка изолированная 208
 особая точка устранимая 208
 открытое множество 25
 отображение 126
 отображение «в» 130
 отображение «на» 130
 отображение инъективное 130
 отображение обратное 130
 отображение сюръективное 130

Пассивные ограничения 182

первообразная 75
 перестановки 18
 перестановки с повторениями 19
 площадь кривой 142
 плоскость 107
 подпоследовательность 40

полярные координаты 14, 155
 полус k -го порядка 208
 полином Бернштейна 69
 полиномы Лежандра 216
 полный дифференциал 112
 порядок малости 56
 последовательность Коши 43
 правило переброски производной 229
 предел 27
 предел функции 32, 107
 предельная точка 25
 преобразование 126
 преобразование Фурье 223
 преобразование Лапласа 223
 преобразование Лапласа обращение 226
 признак Даламбера 47
 признак Коши 47
 пример Брауэра 72
 принцип аргумента 209
 принцип максимума модуля 201
 принцип сжимающих отображений 138
 продолжение функции 25
 проекция на направление 154
 производные высших порядков 66
 производная 52, 191
 производная Фреше 136
 производная на бесконечности 249
 производная обратной функции 55
 производная сложной функции 54
 прообраз 130

Ряд Фурье 218

ряд Лорана 204
 ряд Тейлора 67
 ряд функциональный 210
 равенство Парсеваля 215

равномерная непрерывность 42, 109
равномерная сходимостъ интегралов 124
равномерная сходимостъ ряда 211
радиус сходимости 213
размещения 18
размерность 130
размерность пространства 153
расходящаяся последовательность 28
регулярная задача 178
регулярная точка 178
регулярность 181
ротор 171

Свойство Коши 73
сечение Дедекинда 38
сжимающее отображение 137
симметрическая разность 24
система координат ортогональная 151
система координат правая/левая 154
скалярное произведение 106, 155
скалярное произведение функций 214
собственный вектор 188
собственное значение 188
сопряжённое число 14
сочетания 18
степень отображения 238
степенной ряд 212
столбцовая норма 135
строчная норма 135
сужение функции 25
сумма Дарбу 78
супремум 39
сходимостъ 27
сходимостъ равномерная 121

Теорема Абеля 203
теорема Безу 21
теорема Больцано – Коши 36
теорема Брауэра 244
теорема Дарбу 66, 73
теорема Гаусса – Остроградского 164
теорема Кантора 42
теорема Каруша – Джона 182
теорема Коши 65
теорема Коши-плюс 66
теорема Лагранжа 65
теорема Римана 50
теорема Руше 209
теорема Стокса 171
теорема Штольца 48
теорема Вейерштрасса 36, 41
теорема Вейерштрасса (аппроксимационная) 69
теорема Вьета 23
теорема о среднем 117
теорема сравнения 46
точки второго рода 71
точки первого рода 71

Уравнение теплопроводности 166, 167
ускорение Кориолиса 163
условия Коши – Римана 191
условия дополняющей нежёсткости 183
условие Дини 220
условный экстремум 177

Финитная функция 228
формула Эйлера 15, 68
формула Грина 171
формула Меллина 226
формула Муавра 15
формула Стирлинга 85
формула Тейлора 67

фундаментальная последовательность 43
функция Хэвисайда 229
функция Римана 70
функция аналитическая 193
функция гладкая 174
функция голоморфная 203
функция регулярная 203
функция целая 207
функционализация параметра 251

Характеристическое уравнение 188

Циркуляция 170

Частичная сумма 44, 210
частота 221
частота круговая 221
число e 31
числовая последовательность 26, 27

Эквивалентность норм 134
экспонента 62

Школа Опойцева

*Час, затраченный
на понимание,
экономит
год жизни*

Математический анализ

Краткое и ясное изложение предмета



URSS

Валерий Иванович Опойцев —

доктор физико-математических наук, профессор.
Выделяется умением сложное объяснять просто.

Большой известностью пользуются его «Лекции по математике» (вышло 16 томов под псевдонимом В. Босс). Читайте также идущую нарасхват популярную книгу В. Босс. «Интуиция и математика».



В рамках недавно стартовавшего проекта **ШКОЛА ОПОЙЦЕВА** планируется выпуск учебных курсов для студентов и школьников. Все курсы сопровождаются видео-лекциями на сайте **oschool.ru** и на **youtube.com**

Данная книга — первая ласточка в ансамбле печатных изданий проекта. Изложение матанализа отличается краткостью и прозрачностью. Акцент делается на понимании существа дела, причем с заботой о новичках.

Из отзывов читателей:

«Чтобы усвоить предмет, надо освободить его от деталей, обнажить центральные конструкции. Эту тяжёлую работу автор берёт на себя».

«Дётся то, чего недостает. Общая картина, мотивация, взаимосвязи. И самое главное — легкость вхождения в любую тему».

«Содержание продумано и хорошо увязано. Доказательства ужасты до нескольких строчек. Virtuозное владение языком».

19529 ID 213615



9 785971 032953

Издательская группа

URSS 

Каталог изданий
в Интернете:
<http://URSS.ru>

E-mail: URSS@URSS.ru

117335, Москва, Телефон / факс
Нахимовский (многоканальный)
проспект, 56 +7 (499) 724 25 45

Отзывы о настоящем издании, а также обнаруженные опечатки присылайте по адресу URSS@URSS.ru. Ваши замечания и предложения будут учтены и отражены на web-странице этой книги на сайте <http://URSS.ru>